Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Chromatography A, 1096 (2005) 165–176
JOURNAL OF CHROMATOGRAPHY A
www.elsevier.com/locate/chroma

Simultaneous resolution of overlapping peaks in high-performance liquid chromatography and micellar electrokinetic chromatography with diode array detection using augmented iterative target transformation factor analysis

P.V. van Zomeren *, H.J. Metting, P.M.J. Coenegracht, G.J. de Jong [1]

Department of Analytical Chemistry and Toxicology, University of Groningen, P.O. Box 196, NL-9700 AD Groningen, The Netherlands

Available online 31 August 2005

Abstract

In this paper, augmentation has been applied to data matrices, which originate from hyphenated methods that share the same mode of detection, but use different separation methods, HPLC-DAD and MEKC-DAD. A novel method, wavelength shift eigenstructure tracking (WET), has been proposed for the alignment between the wavelength scale of both detectors. WET proves to be suitable for the detection as well as correction of wavelength shift between both detectors. After correction of the wavelength scale, data obtained on both systems have been augmented and submitted to iterative target transformation factor analysis. Augmented curve resolution provides significantly better estimates of the chromatographic and electrophoretic profiles and spectra than the use of non-augmented curve resolution on HPLC and MEKC data separately. It is particularly useful when the pure fraction of a chromatographic peak is less than 0.10. Finally, the relative weight of MEKC versus HPLC in augmentation may be increased using intensity and noise normalisation. However, since noise normalisation and its accompanying decrease in signal-to-noise ratio leads to a loss of information, and, since intensity normalisation may cause a failure of the augmented curve resolution algorithm, benefits and drawbacks of normalisation should be weighed on a case-by-case basis.

Keywords: High performance liquid chromatography; Micellar electrokinetic chromatography; Diode array detection; Iterative target transformation factor analysis; Augmentation; Wavelength shift eigenstructure tracking

1. Introduction

Drug impurity profiling requires the use of separation methods with a very high degree of selectivity, since the compounds that need to be analysed are usually closely related. So, hyphenated methods are often used [1]. Application of hyphenated methods results in the acquisition of bilinear data matrices. Treatment of these matrices with chemometrical methods may enhance selectivity by resolution of overlapping peaks. An overview of the, so-called, curve resolution methods was given by Liang et al. [2]. Unfortunately, concentration profiles and spectra, which are obtained by curve resolution of a bilinear data matrix that corresponds to a single experiment, are potentially unreliable as a result of unresolved rotational and intensity ambiguities, unless selective information is present for all compounds [3]. Intensity ambiguities are caused by the scaling of the estimated concentration profiles and spectra by some unknown factor, while rotational ambiguities are caused by the fact that the estimated concentration profiles and spectra are mixtures of the true ones due to insufficient selective information. These ambiguities can be removed if several experiments, which have been performed under different conditions, are analysed simultaneously, and if the concentration profiles or the spectra in these different experiments are equal [4]. Relative intensity ambiguities among data matrices will always be solved when they are analysed simultaneously, since the unitary concentration profiles or spectra of the individual species are forced to be equal. However, rotational ambiguities will only be solved if every species has a selective region in at least one of the experiments [5]. Combination of multiple data matrices, augmentation, is only possible when individual matrices have at least one order in common. Rank analysis of column-wise and

* Corresponding author. Tel.: +31 50 3633336; fax: +31 50 3637582.
E-mail address: paulvanzomeren@yahoo.co.uk (P.V. van Zomeren).
[1] Present address: Department of Biomedical Analysis, University of Utrecht, P.O. Box 80082, 3508 TB Utrecht, The Netherlands.

0021-9673/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2005.08.047

row-wise augmented data matrices is an appropriate method to detect orders in common between data matrices. Furthermore, the total number of species can be estimated in this way [6].

Up to now, augmentation was mainly used for the simultaneous analysis of bilinear data matrices, which were obtained using the same hyphenated method under different experimental conditions. Examples are the application of augmented MCR-ALS or other deconvolution methods to excitation–emission fluorescence spectroscopy [7–9], spectrophotometric titrations [2–5,10–16], HPLC-DAD [17–21], and CE-DAD [22,23]. Since in drug impurity profiling several hyphenated methods are used for the analysis of a single sample, it would make sense to evaluate whether augmentation can be used for the simultaneous analysis of data from different hyphenated methods. As augmentation is only possible when data matrices have at least one order in common, these hyphenated methods would either have to share the same separation method and use different modes of detection, or, share the same mode of detection and use different separation methods. Examples of the former were revealed by Brereton et al. in the simultaneous treatment of HPLC-DAD and HPLC–MS data [24–25] and by Wang et al. in the combination of GC–FTIR and GC–MS [26].

Alignment and interpolation of the chromatographic profiles proves to be a decisive step in this type of augmentation. A large number of methods has been introduced for the alignment of chromatographic profiles with small retention time shifts [26–31]. In order to apply augmentation to bilinear data matrices from hyphenated methods, which share the same mode of detection but use different separation methods, the alignment of spectra is crucial. Some experience in the adjustment of spectral changes has been acquired in the field of calibration transfer. Two approaches can be used; either the calibration model is modified in order to provide similar results for data from the second instrument, or data from the second instrument are modified in such a way that they appear to have been measured on the first instrument [32]. Only the latter approach is appropriate for the current objective (matching spectra from HPLC and MEKC).

Full spectrum standardisation (FFS) is an approach, based on the assumption that differences between spectra are linear, contain no wavelength shifts and that the intensity shifts are equal at all wavelengths [33]. A comparable method, called single wavelength standardisation (SWS), also corrects for intensity differences that vary across the wavelength region by regressing the responses of both instruments at each channel. However, the requirements of linearity and the absence of wavelength shift are still maintained [32–34]. In order to adjust for wavelength shifts and non-linear behaviour, novel methods of calibration transfer have been developed. First of all, a patented method was introduced by Shenk and Westerhaus [35]. This method corrects the full spectral response function of an instrument by performing a wavelength index conversion, which is followed by a spectral intensity transformation at each wavelength [36]. Subsequently, direct standardisation (DS) was proposed by Wang et al. [37]. This method directly relates the response of a sample measured with one instrument to its response obtained on another instrument. A suitable transformation matrix is calculated using linear regression. Unfortunately, DS does not adept well to locally occurring wavelength shifts and non-linearity [38,39]. To improve DS, Wang et al. introduced piece-wise direct standardisation (PDS) [37]. This method is based on the consideration that the spectral information given at a certain wavelength under the calibration conditions is contained in a small spectral region of neighbouring wavelengths under the prediction conditions [40]. PDS is capable of taking into account wavelength shifts, small non-linearities, and different intensity shifts across the wavelength range [33]. It is one of the most widely used methods for calibration transfer [33,34,38,41–51]. Lately, modified versions of PDS were introduced; continuous piecewise direct standardisation (CPDS) [52] and wavelet hybrid direct standardisation (WHDS) [53]. Finally, the use of artificial neural networks (ANN) was proposed. Since spectral windows on the second instrument are matched to the centre points of the corresponding spectral windows on the first instrument and non-linear transfer functions are used, the problems of wavelength shift and non-linear behaviour are both handled [54,55].

All above-mentioned approaches share the same disadvantage; a representative set of compounds is required to perform alignment of the spectra. Since in drug impurity profiling the number and identity of the compounds is principally unknown, these approaches are not suitable. A standard-free calibration transfer method was introduced by Blank et al. [56]. This approach, which is known as finite impulse response (FIR) filtering, uses a representative spectrum from the master instrument to filter the spectra from the slave instrument. The main disadvantages of FIR are the inability to process the ends of spectra and to handle baseline discontinuities [56]. An improved algorithm has been proposed that makes the transfer more robust by avoiding transfer artefacts [57,58].

In the current paper, augmentation has been applied to data matrices from hyphenated methods that share the same mode of detection, but use different separation methods, in this case, HPLC-DAD and MEKC-DAD. As none of the above-mentioned calibration transfer methods completely meets the requirements for spectral alignment, a novel method based on singular value decomposition (SVD) and cubic spline interpolation, has been developed. This approach, which is called wavelength shift eigenstructure tracking (WET), has been used to detect the wavelength shift that exists between both detectors. It is also suitable for the actual alignment by adjustment of the wavelength scale of one of the detectors. Since determination of the required wavelength correction is based on the actual measurements, separate analysis of a representative set of compounds on both systems is not necessary. After wavelength scale alignment, the resulting augmented data matrices have been decomposed using an augmented curve resolution method, in this case, iterative target transformation factor analysis (ITTFA). Results have been compared to those obtained using the same curve resolution method without augmentation. The effect of normalisation of the intensity and/or noise level between HPLC and MEKC data

on the performance of augmentation as well as the influence of the degree of peak overlap in HPLC and MEKC data have been investigated.

2. Theory

2.1. Augmented iterative target transformation factor analysis (ITTFA)

Augmented ITTFA does not differ fundamentally from normal ITTFA. Only a few minor modifications of the algorithm are necessary to accommodate for the fact that augmented chromatographic profiles contain multiple peaks, while normal chromatographic profiles are unimodal. These multiple peaks all correspond to the same compound under various experimental conditions. The algorithm used in this paper is based on a paper by Vandeginste et al. [59]. In the first step, principal component analysis (PCA) is performed on the complete data matrix. Since augmentation does not change the number of chemical species present, this method provides an estimate of the total number of compounds. This number also indicates which of the obtained abstract spectra and abstract augmented chromatographic profiles are significant. Subsequently, an iterative target testing procedure is used to change the abstract augmented chromatographic profiles into the true augmented chromatographic profiles. First of all, a Varimax rotation [60] is used. This type of rotation tries to obtain a few large loadings and as many near-zero loadings as possible. In other words, it attempts to make the chromatographic profiles unimodal. Therefore, it is applied separately to each part of the augmented chromatographic profile that originates from a single chromatographic run. After the rotation, the separated parts are recombined to form augmented start targets. Subsequently, these start targets are processed, one at a time. A rotation matrix is calculated based on the first target. Then, the obtained rotation matrix is used to calculate a predicted target. Finally, this predicted target is evaluated and adapted. Elements of the predicted target which are below a certain threshold are set to zero, and minor peaks which are separated from the main peak by one or more zeros are eliminated. For obvious reasons, this part of the algorithm is performed separately for each part of the augmented predicted target that originates from a single chromatographic run. After adaptation the target is resubmitted to the iterative procedure. Iteration is stopped, when a maximum number of iterations is reached, adaptation of the target is not possible, or the predicted and adapted target are sufficiently similar. The iterative procedure is repeated for the remaining start targets. Finally, the accepted targets are scaled to unit length to obtain augmented chromatographic profiles. Linear regression is used to calculate the spectra based on these augmented chromatographic profiles and the augmented data matrix.

2.2. Wavelength shift eigenstructure tracking (WET)

A primary consideration in the alignment of spectra between different spectrophotometers is the question whether a discrete or continuous wavelength shift is required. A discrete shift involves the transfer of one of the wavelength scales by an integer multiple of the sample interval. Essentially, the data obtained on one of the spectrophotometers is moved one or more positions in the spectral direction. On the other hand, a continuous wavelength shift involves the transfer of one of the wavelength scales by a fraction of the sample interval. Since, in this case, the target wavelength will not match the next position in the spectral direction, interpolation of the data is necessary. Although linear interpolation is a straightforward approach that could prove to be suitable, a more subtle method was proposed by Gong et al. [61]. They used cubic splines for the interpolation of chromatograms. The same approach would be suitable for spectra without modification. Spline functions are defined as piecewise polynomials of degree n. The pieces join in so-called knots and fulfil continuity conditions for the function itself and for the first $n-1$ derivatives. The number of knots (m) is equal to the number of polynomial pieces that are used to build the spline function. Consequently, a cubic spline is a continuous function, which is compiled using third order polynomial pieces, with continuous first and second derivatives [62]:

$$y = \sum_{j=1}^{m} a_j + b_j x + c_j x^2 + d_j x^3 \quad (1)$$

A second concern is the development of a suitable approach towards determination of the actual extent of the wavelength shift. In the field of remote sensing, cross correlogram spectral matching (CCSM) was developed by van der Meer and Bakker [63]. This method plots the correlation between two spectra as a function of wavelength shift. A maximum is observed when the most appropriate wavelength shift is reached. Unfortunately, the method is not suitable for the current purpose, since it can only handle vectors. Furthermore, it cannot be used to manage multiple components at the same time. Eigenstructure tracking analysis (ETA) was introduced by Cuesta Sánchez et al. [64] to create local rank maps of bilinear data matrices. A small selection of the bilinear data matrix, the so-called window, which contains all spectral points and only a few (odd number) time points, is moved along the matrix in the time direction, one point at a time. Each window is submitted to SVD. Subsequently, the natural logarithms of the eigenvalues are plotted as a function of time. The number of eigenvalues, which deviate significantly from the noise level, is an indication of the local rank. A combination of the concepts of CCSM and ETA has led to the development of WET. Given the fact that two bilinear data matrices include spectral information for the same compounds, the resulting augmented matrix should have a rank equal to the number of compounds when the wavelength scales are aligned perfectly. Therefore, in WET augmented matrices are formed comprising of matrices with different degrees of wavelength shift. Each augmented matrix is submitted to SVD. Then, the natural logarithms of the eigenvalues are plotted as a function of the degree of wavelength shift. A minimum in the number of significant eigenvalues is observed when the most appropriate wavelength shift is reached. The augmented matrix which produced this minimum is suitable for further processing.

3. Experimental

3.1. Chemicals

The benzodiazepines, nitrazepam, clonazepam, and lorazepam, were donated by several pharmaceutical companies. Since their molecular structures are closely related, they have rather similar UV-spectra [65] and comparable chromatographic behaviour. Therefore, a mixture of these compounds is difficult to analyse, which makes them excellent candidates to investigate the potential of selectivity enhancement by combined interpretation of data from HPLC and MEKC. HPLC-grade acetonitrile and methanol were obtained from Labscan (Dublin, Ireland), and, phosphoric acid (85%), sodium dodecyl sulphate (SDS) and sodium hydroxide from Merck (Darmstadt, Germany). All chemicals were of analytical grade. An ELGA Maxima ultra pure water system (Salm & Kipp, Breukelen, the Netherlands) was used to prepare de-ionised water. A solution containing 200 mg/l of each of the benzodiazepines was prepared in a mixture of methanol and water. The elution strength of the solvent was kept as close as possible to the elution strength of the mobile phases. This solution was used for HPLC. Solutions containing 62.5 or 125 mg/l of each of the benzodiazepines were prepared in separation buffer. These solutions were used for MEKC.

3.2. HPLC system

An HP 1090 liquid chromatograph (Agilent, Waldbronn, Germany) was used, equipped with a diode array detection system. System control, data acquisition, and data analysis were performed using ChemStation Release 04.02. An ODS Hypersil C_{18} column with dimensions 250 mm × 4 mm and particle size 5 μm was supplied by Agilent (Waldbronn, Germany). Mobile phases containing water (H_2O), methanol (MeOH), and acetonitrile (MeCN) were used. Flow rate was kept constant at 0.8 ml/min. A built-in autosampler was used to inject 5 μl portions of the sample solution. The syringe was rinsed three times with 20 μl of methanol between injections. Spectra were obtained from 220 to 350 nm using a 2 nm resolution and 1.3 Hz sample rate. All experiments were performed at ambient pH (5.5–7.0, not buffered) and temperature (20–25 °C). Details on the mobile phase composition are provided in Table 2 under Section 4.

3.3. MEKC system

Experiments were carried out using a Hewlett-Packard 3D capillary electrophoresis system (Agilent, Waldbronn, Germany), equipped with a diode-array detection system and ChemStation Release 05.01 for system control, data acquisition and data analysis. All samples were injected hydrodynamically by applying a pressure of 50 mbar for 2 s. Electrophoresis was performed at a constant potential of 15 or 25 kV, and the temperature of the capillary was maintained at 25 °C. An uncoated fused-silica capillary of 58.5 or 62.0 cm (effective length 50.0 or 53.5 cm) × 75 μm internal diameter was used. Before use the capillary was rinsed with 1 M sodium hydroxide (15 min), followed by de-ionised water (15 min) and separation buffer (30 min). Between runs the capillary was flushed with buffer for 2 min. The same separation buffer, containing 7.0 mM SDS and 15.0% acetonitrile in 10 mM phosphate buffer pH 7.0, was used for all samples. Spectra were obtained from 220 to 350 nm using a 2 nm resolution and a 1.25 Hz sample rate. Details on the capillary length, sample concentration, and separation potential are provided in Table 3 under Section 4.

3.4. Software

The obtained data were saved in ASCII-format using an HP ChemStation macro. For each run, one matrix with the measured absorbance as a function of wavelength and time, and two vectors with wavelength and time labels, respectively, were stored. Subsequently, these data were loaded into Matlab 5.3 (Mathworks, Natick, MA, USA) for further processing. All steps of the method were integrated into one graphical user interface controlled Matlab 5.3 algorithm. In the first step of the algorithm ETA was used to detect zero concentration windows in the HPLC data. This information was used to correct background fluctuations (baseline offset and drift); the response in the zero concentrations windows was interpolated using linear regression, for each wavelength separately, and subtracted from the data matrix. This process was repeated for the MEKC data. Subsequently, WET was employed to determine the appropriate degree of wavelength shift. After manual selection, WET was applied again, on the corrected matrix, to confirm that the chosen wavelength shift was suitable. As an option, the response and the noise level were normalised in both data matrices. Normalisation of the response was performed by multiplication of the MEKC data matrix with a constant in order to obtain the same total peak area as for the HPLC data matrix, while normalisation of the noise was performed by adding an amount of normally distributed noise to the HPLC data matrix in order to obtain the same average higher order eigenvalues in the ETA plot as for the MEKC data matrix. Both normalisation steps were optional. Finally, the matrices were augmented and submitted to curve resolution by ITTFA.

4. Results and discussion

4.1. Simultaneous curve resolution of HPLC and MEKC peaks

An HPLC chromatogram was obtained for a sample containing nitrazepam, clonazepam, and lorazepam. The applied mobile phase contains 40% (v/v) of water and 60% (v/v) of acetonitrile. Although these conditions permit a very fast analysis, the obtained separation is completely inadequate; only one peak is observed in the chromatogram (Fig. 1a). A selection of the obtained data matrix, from 3.0 to 5.0 min (0.0125 min interval) and from 221.5 to 349.5 nm (2 nm interval), was submitted to ITTFA. Although three compounds are detected and resolved, the obtained chromatographic profiles show artefacts in the form of negative parts (Fig. 1b).

Conditions are chosen in order to ensure that spectra in HPLC and MEKC are as closely related as possible within the limits set

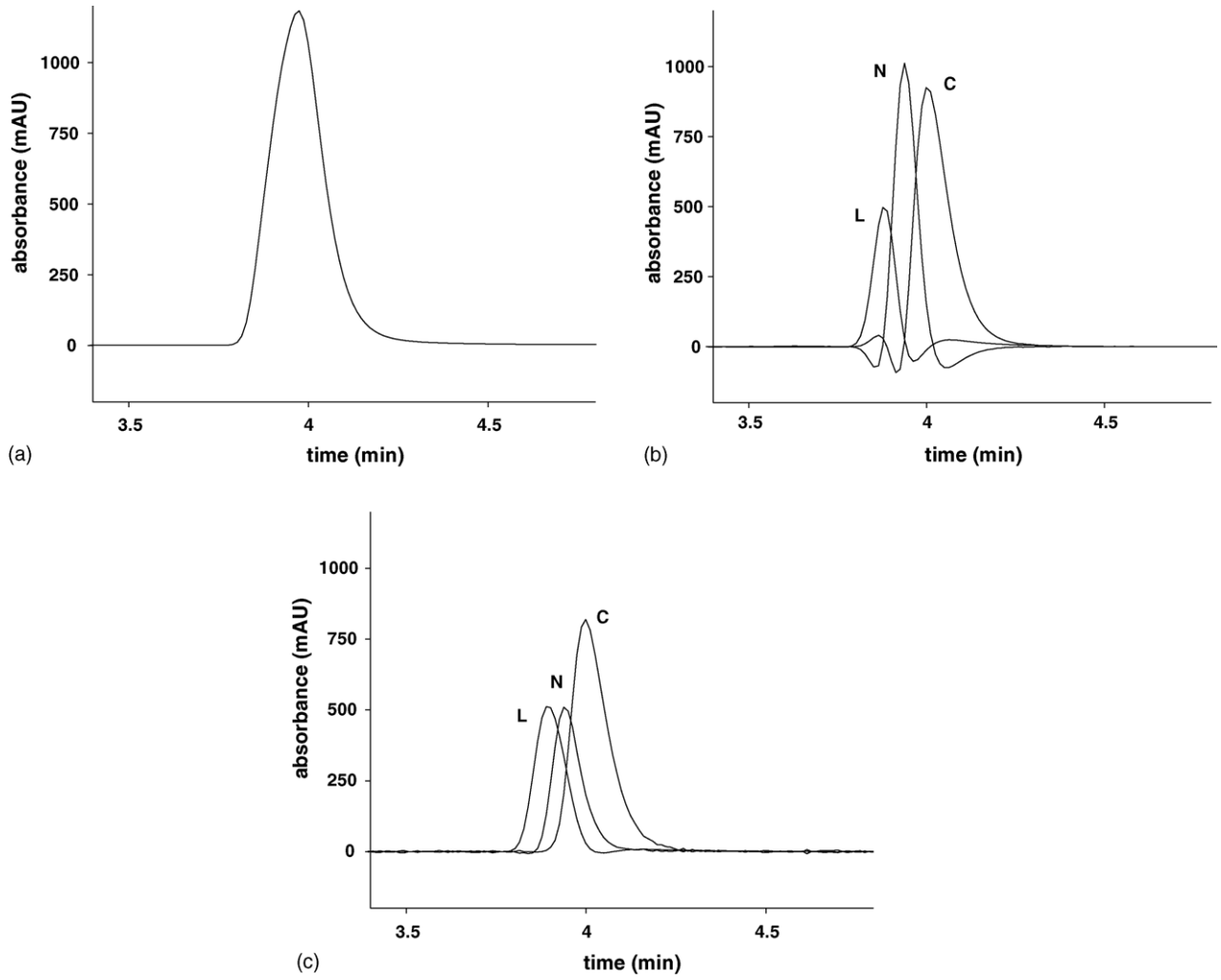

Fig. 1. HPLC chromatogram obtained using a mobile phase containing 40% H_2O and 60% MeCN (a), corresponding concentration profiles provided by the application of ITTFA to the HPLC data only (b), and augmented ITTFA to the HPLC and MEKC data simultaneously (c).

by these separation methods. Both methods are applied at neutral pH and using the same organic modifier (acetonitrile) with different concentrations (ca. 60% (v/v) for HPLC and 15% (v/v) for MEKC). Compounds are also forced as much as possible into the aqueous phase by use of a low micelle concentration. Furthermore, conditions are selected in order to obtain peaks that are sufficiently wide for the diode array detector to collect enough spectra. While the sample frequency of the detector is adequate for normal MEKC peaks with single-wavelength UV mode, it is not sufficient for such peaks with DAD mode. Consequently, the peaks in the electropherogram are wider, and, hence the plate numbers are lower than normally observed in MEKC.

An electropherogram was acquired for a sample containing nitrazepam, clonazepam, and lorazepam. The applied separation potential is 25 kV, and the effective length of the capillary is 50.0 cm. At pH 7.0 the benzodiazepines of interest are all uncharged. As a result, their separation depends on the presence of micelles in the separation buffer. Although the concentration of sodium dodecyl sulphate (7.0 mM) seems below the critical micelle concentration, taking into account the relatively high acetonitrile content of the separation buffer (15%), micelles are still assumed to be present. However, their concentration is relatively low and they are much smaller [66]. Consequently, these conditions provide a poor separation; only one peak is observed in the electropherogram (Fig. 2a). A selection of the data matrix, from 7.6 to 10.6 min (0.0133 min interval) and from 220.5 to 348.5 nm (2 nm interval), was submitted to ITTFA. Three compounds are detected and resolved, but the obtained electrophoretic profiles show the same artefacts that are also observed in the chromatographic profiles (Fig. 2b).

The spectra obtained from HPLC (Fig. 3a) and MEKC (Fig. 3b) data deviate somewhat from the pure HPLC spectra (Fig. 3d). This is most prominent for the compounds that do not possess selective regions; in the HPLC chromatogram nitrazepam overlaps completely with lorazepam (first peak) and clonazepam (last peak), while, in the MEKC electropherogram clonazepam overlaps completely with nitrazepam (first peak) and lorazepam (last peak). Although the occurrence of artefacts in chromatographic profiles and spectra depends on the use of constraints, the magnitude of the artefacts is also determined by the degree of peak overlap [67].

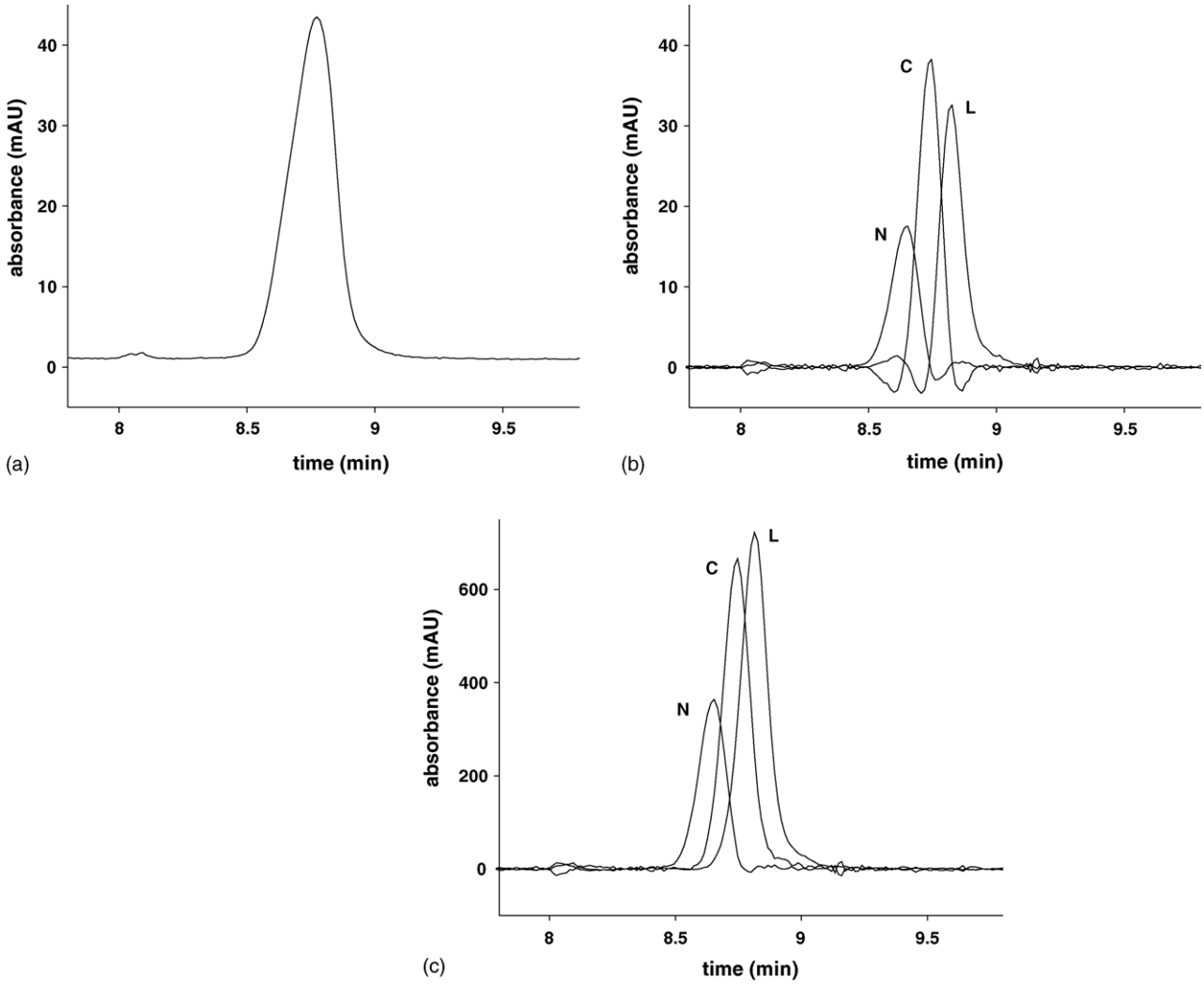

Fig. 2. MEKC electropherogram obtained using a capillary of 58.5 cm (effective length 50.0 cm) for a sample containing 62.5 mg/l of each of the benzodiazepines (a), corresponding concentration profiles provided by the application of ITTFA to the MEKC data only (b), and augmented ITTFA to the HPLC and MEKC data simultaneously (c).

Wavelength shift eigenstructure tracking (WET) was applied to the HPLC and MEKC matrices in order to determine the required degree of wavelength shift. Three eigenvalues remain constant independent of the wavelength shift, while the remaining eigenvalues show a minimum at a shift of 1 nm (Fig. 4a). This clearly indicates that the combined matrices contain spectral information from three different compounds and that the wavelength scales in both matrices differ by 1 nm. The observed degree of wavelength shift is correct, since it is known that the HPLC data have been obtained at 221.5, 223.5, . . ., 349.5 nm, while the MEKC data have been obtained at 220.5, 222.5, . . ., 348.5 nm. Although it is possible to select wavelength range as well as sample frequency, the exact wavelengths can not be chosen. After correction, WET was applied again, to confirm that the chosen wavelength shift is adequate (Fig. 4b). First of all, the spectra from HPLC and MEKC now match very closely; apart from the wavelength shift no other differences are present in the spectra. This can be concluded from the fact that at the minimum in the WET plot only three eigenvalues deviate significantly from the noise level, which corresponds to the three compounds. Secondly, the WET plot before wavelength scale correction (Fig. 4a) shows artefacts (bumpy plots of higher order eigenvalues), which are absent after correction (Fig. 4b). These artefacts are more intense when linear interpolation is used instead of cubic spline interpolation. The frequency of this bumpy pattern is once every 2 nm, which is equal to the sample interval of the diode array detector. Furthermore, the tops appear to coincide with wavelengths that do not require interpolation. Therefore, it is likely that this artefact is related to the interpolation. The second WET plot does not show any artefacts. Since one of the matrices involved in the construction of the second plot WAS obtained by correction of the wavelength scale, all data are influenced by interpolation.

The intensity and noise level were normalised between both data matrices. Next, these matrices were augmented and submitted to curve resolution by ITTFA. Three compounds are detected and resolved. The obtained chromatographic profiles (Fig. 1c) and electrophoretic profiles (Fig. 2c) do not show any signs of artefacts. Furthermore, the obtained spectra (Fig. 3c) are almost identical to the pure HPLC spectra (Fig. 3d). For all compounds,

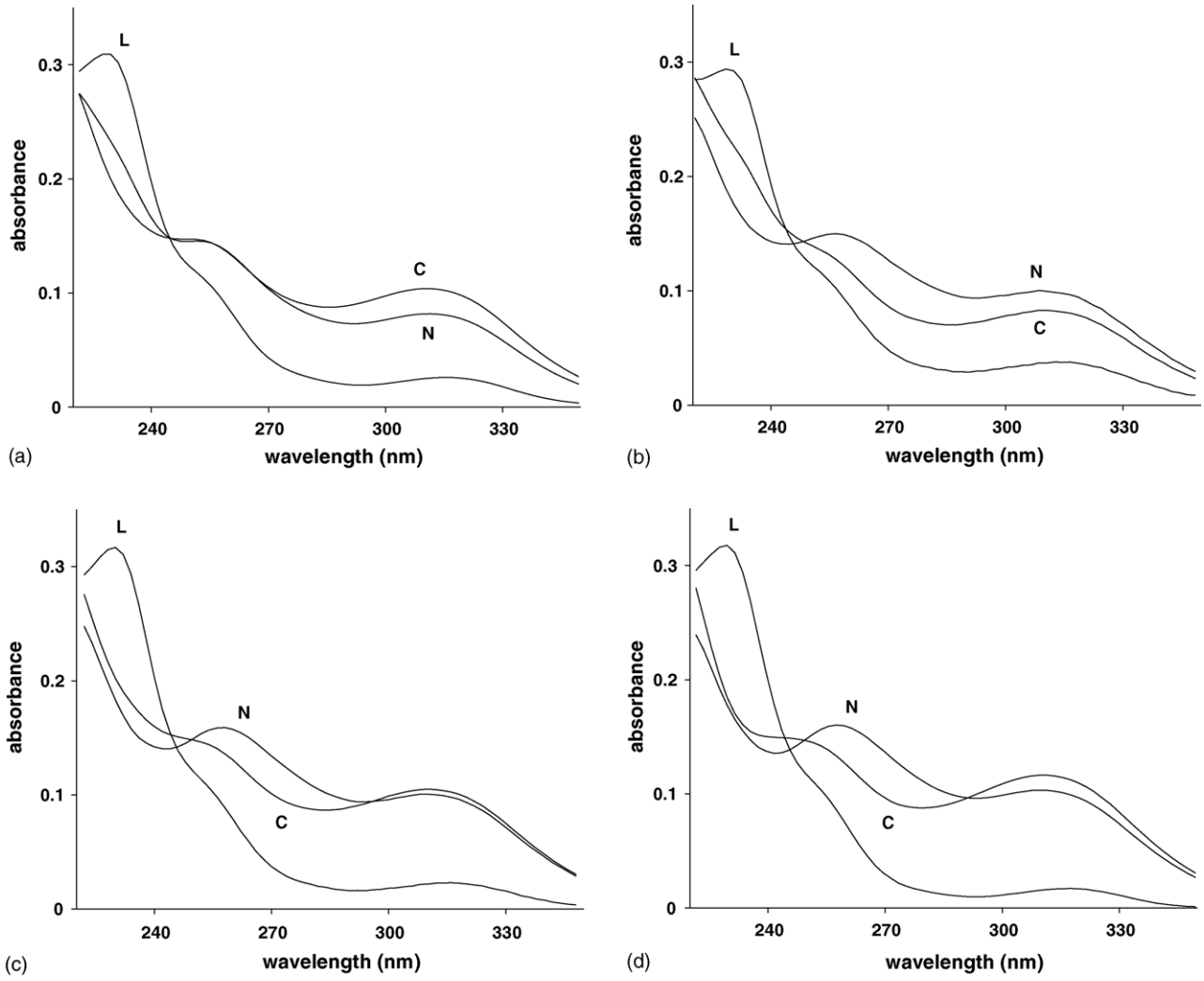

Fig. 3. Spectra provided by the application of ITTFA to the HPLC data only (a), spectra provided by the application of ITTFA to the MEKC data only (b), spectra provided by the application of augmented ITTFA to the HPLC and MEKC data simultaneously (c), and pure spectra obtained by the analysis of standards under the same HPLC conditions (d).

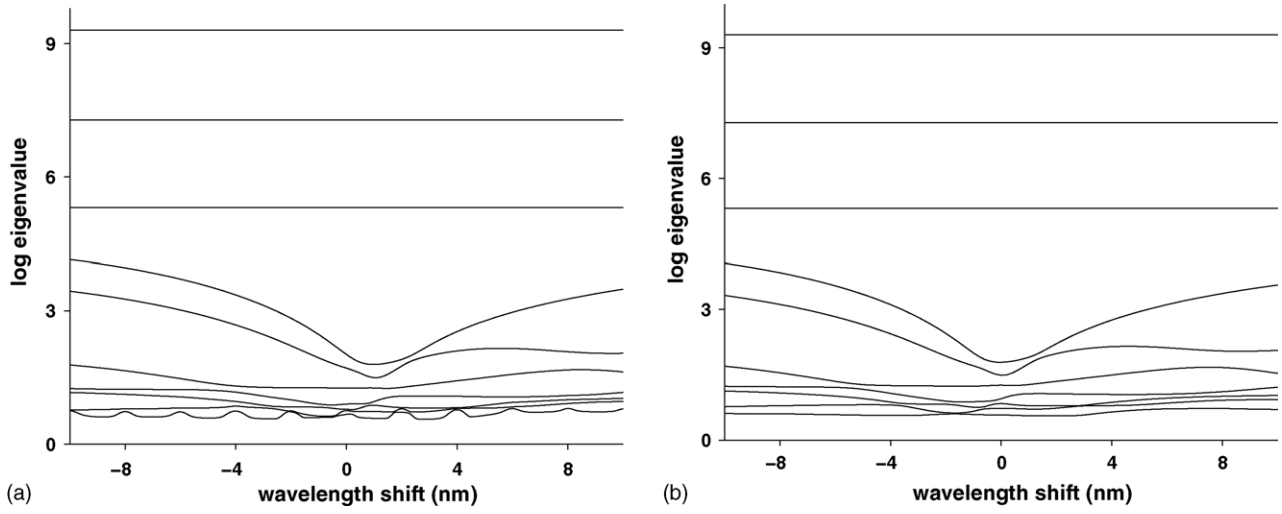

Fig. 4. Wavelength shift eigenstructure tracking (WET) plot obtained for the HPLC and MEKC data before (a) and after (b) wavelength scale correction.

Table 1
Correlation between pure and resolved spectra

Compound	Single ITTFA for HPLC data	Single ITTFA for MEKC data	Augmented ITTFA for both data sets
Nitrazepam	0.9243	0.9911	0.9972
Clonazepam	0.9853	0.9636	0.9909
Lorazepam	0.9993	0.9982	0.9997

the correlation between the resolved spectra from augmented curve resolution and the pure spectra is higher than the correlation between the resolved spectra from non-augmented curve resolution and the pure spectra (Table 1). Moreover, after augmented curve resolution, this correlation is very good (>0.99) for all compounds. On the other hand, after non-augmented curve resolution, the compounds with the most severe degree of overlap, nitrazepam for HPLC and clonazepam for MEKC, show the lowest correlation.

4.2. Evaluation of intensity and noise normalisation

In Fig. 5, ETA-plots are shown for augmented HPLC and MEKC data from the example. Spectra 1–162 were obtained by HPLC, and spectra 163–387 were obtained by MEKC. Before normalisation, the noise levels of both data are approximately equal (Fig. 5a), which is shown by comparable higher order eigenvalues. On the other hand, the signal intensity is significantly lower for MEKC than for HPLC (Fig. 5a), which is shown by a lower first eigenvalue. Thus, normalisation of the intensity may be indicated to balance the weight of HPLC and MEKC in the resolution process. Since the signal-to-noise ratio of the MEKC data is not affected by intensity normalisation, the noise level of the MEKC data is increased compared with that of the HPLC data (Fig. 5b). Hence, normalisation of the noise level may be appropriate. However, the addition of normally distributed noise leads to a decrease in the signal-to-noise ratio of HPLC data (Fig. 5c). Since a decrease in signal-to-noise ratio leads to a loss of information, this approach may not be favourable.

A series of eight HPLC data matrices were obtained using different compositions of mobile phase (Table 2). Additionally, a sequence of eight MEKC data matrices were acquired using different capillaries, sample concentrations, and separation potentials (Table 3). Augmented curve resolution was applied to all possible HPLC and MEKC data combinations (in pairs) to obtain chromatographic or electrophoretic profiles and spectra. The retention or migration times and the peak widths at half height of the benzodiazepines were calculated from the profiles. Subsequently, the elution order was determined based on the retention or migration times of the individual compounds (Tables 2 and 3). It is assumed that all peaks are symmetrical and Gaussian-shaped with a peak width at the baseline of 6σ ($\pm 3\sigma$). Subsequently, the fraction of the area under the peak, which does not show overlap with other compounds, was calculated by integration of the Gaussian distribution using a so-called error function. This fraction of the area is defined as the pure part of the peak (Tables 2 and 3). All normalisation scenarios were tested.

Table 2
HPLC mobile phase compositions and separations for nitrazepam (N), clonazepam (C), and lorazepam (L)

Run	Mobile phase composition			Separation			
	H_2O (%)	MeOH (%)	MeCN (%)	Peak order	Pure part of N peak	Pure part of C peak	Pure part of L peak
1	40	0	60	L–N–C	0.00	0.03	0.03
2	60	0	40	L–N–C	0.01	1.00	0.03
3	56	8	36	N–C–L	1.00	0.06	0.14
4	51	0	49	N–L–C	0.00	0.95	0.00
5	45	0	55	N–L–C	0.00	0.19	0.00
6	46	3	51	N–L–C	0.04	0.02	0.00
7	42	2	56	N–L–C	0.00	0.00	0.00
8	60	1	39	N–L–C	0.39	1.00	0.64

Table 3
MEKC conditions and separations for nitrazepam (N), clonazepam (C), and lorazepam (L)

Run	Conditions		Separation			
	Column/sample[a]	Potential (kV)	Peak order	Pure part of N peak	Pure part of C peak	Pure part of L peak
1	I	25	N–C–L	0.11	0.00	0.04
2	I	15	N–C–L	0.21	0.00	0.25
3	I	15	N–C–L	0.22	0.00	0.33
4	I	15	N–C–L	0.14	0.00	0.37
5	II	25	N–C–L	0.14	0.00	0.93
6	II	25	N–C–L	0.20	0.00	0.97
7	II	25	N–C–L	0.20	0.00	0.95
8	II	15	N–C–L	0.27	0.09	1.00

[a] A column with a total length of 58.5 cm and an effective length of 50 cm is used and samples containing 62.5 mg/l of each of the benzodiazepines (I), or a column with a total length of 62 cm and an effective length of 53.5 cm is used and samples containing 125 mg/l of each of the benzodiazepines (II).

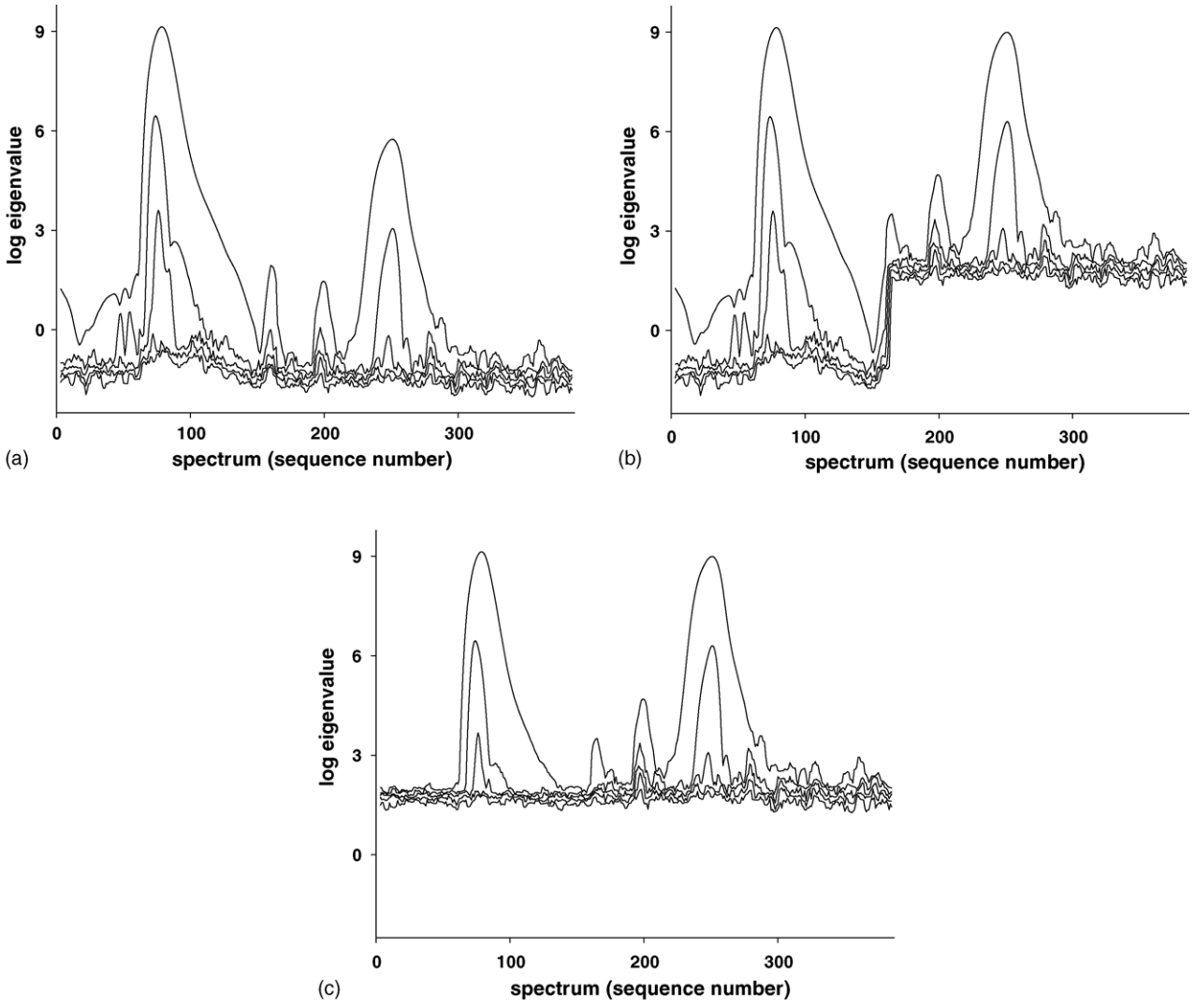

Fig. 5. Eigenstructure tracking analysis (ETA) plot obtained for the augmented HPLC and MEKC data without normalisation (a), with intensity normalisation only (b), and with intensity as well as noise normalisation (c).

Correlation between the obtained and the pure spectra of individual compounds is used as a criterion to determine whether normalisation is successful. The use of intensity normalisation to increase the influence of the MEKC data on the resolution process is predominantly advantageous in situations where a compound, which shows strong peak overlap in HPLC, is reasonably well separated in MEKC. For example, lorazepam is reasonably to very well separated under most MEKC conditions, while it is poorly separated under most HPLC conditions. As expected, the use of intensity normalisation provides the best resolution in 82.8% of the cases, while no normalisation is most successful in only 17.2% of the cases. Furthermore, a combination of intensity and noise normalisation works much better (65.6%) than intensity normalisation alone (17.2%). Whether this is attributable to removal of noise discontinuities or suppression of the HPLC data due to a decreased signal-to-noise ratio, is unknown. On the other hand, clonazepam is not separated well under any MEKC conditions, while it is reasonably to very well separated under some HPLC conditions. Interestingly, here, no normalisation provides the best results in 68.8% of the cases, while intensity normalisation alone or in combination with noise normalisation is most successful in 15.6% of the cases, each. It should be remarked that for HPLC conditions with a very good separation of clonazepam (e.g. runs 2, 4, and 8) no normalisation scores much better (91.7%), while for HPLC conditions with a less favourable separation (e.g. run 7) intensity normalisation works better (87.5%). Finally, it is noticed that intensity normalisation may induce a failure of the augmented curve resolution algorithm (close to singularity of the cross product matrix in the linear regression step). This may be the result of the amplification of certain baseline artefacts in the MEKC data. Consequently, it is important to weigh the benefits and drawbacks of normalisation.

4.3. Evaluation of augmented versus non-augmented curve resolution

First of all, non-augmented curve resolution was applied to each of the eight HPLC data matrices (Table 2). The correlation between the resolved and pure spectra was cal-

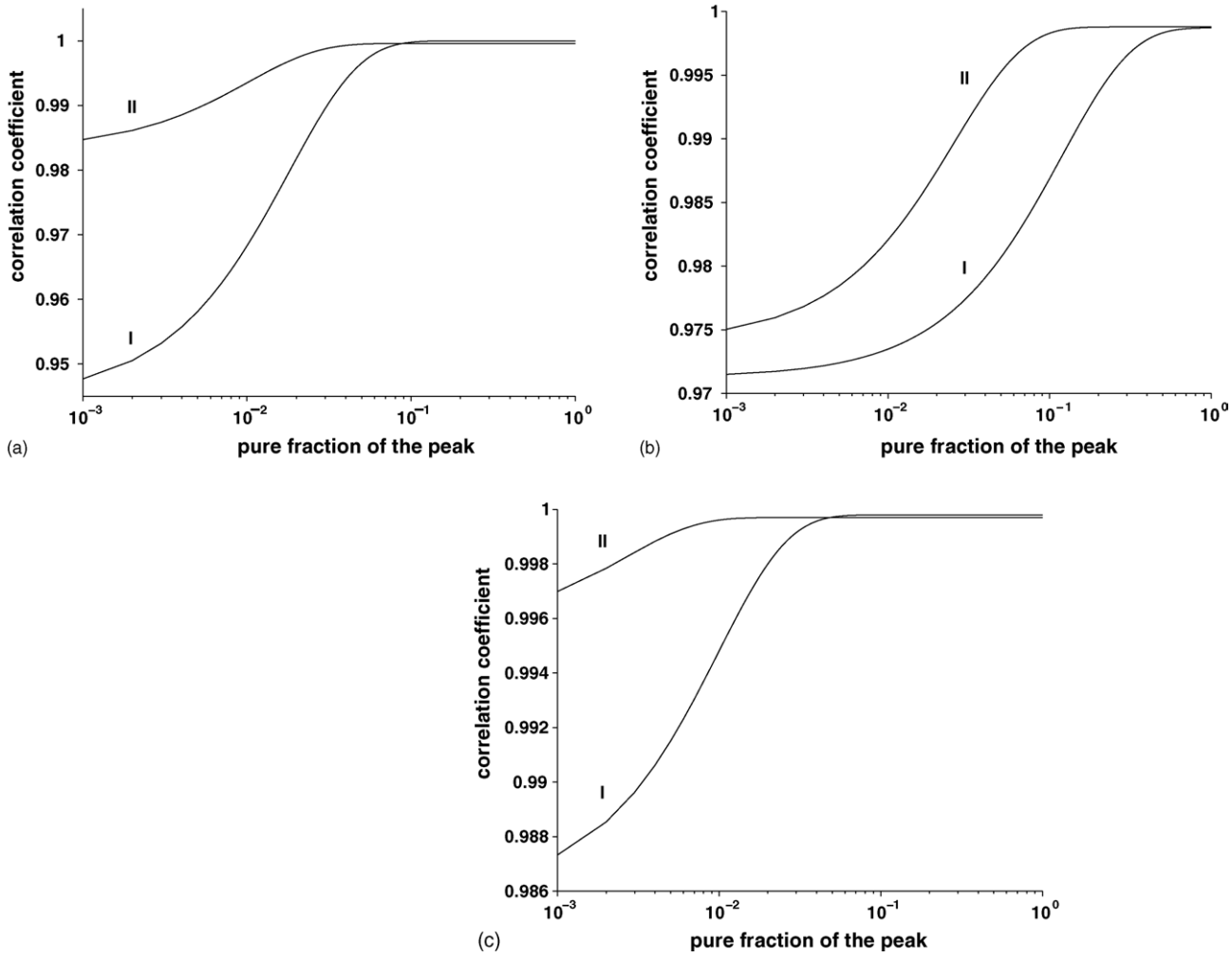

Fig. 6. Correlation between the resolved and the pure spectrum of nitrazepam (a), clonazepam (b), and lorazepam (c) as a function of the pure fraction of the chromatographic peak in HPLC. Separate curves are shown for non-augmented curve resolution (I) and augmented curve resolution (II).

culated and displayed as a function of the pure fraction of the corresponding chromatographic peak, for each compound separately (Fig. 6a–c, curve I). To enhance the clearness of the graphs, an exponential function was fitted through the points and a logarithmic scale was used for the pure fraction.

For nitrazepam each of the eight HPLC runs were combined with the fourth MEKC run (Table 3). Combination of HPLC and MEKC clearly increases the correlation between pure and resolved spectra for HPLC runs with a pure fraction of nitrazepam below 0.10 (Fig. 6a, curve II). Subsequently, for lorazepam all HPLC runs were combined with the eighth MEKC run (Table 3). Once more, combination of HPLC and MEKC increases the correlation between pure and resolved spectra for HPLC runs with a pure fraction of lorazepam below 0.10 (Fig. 6c, curve II). Finally, for clonazepam all HPLC runs were combined with the third MEKC run (Table 3). Although the combination of HPLC and MEKC increases the correlation between pure and resolved spectra (Fig. 6b, curve II), the behaviour of clonazepam is clearly different from that of nitrazepam and lorazepam.

For nitrazepam and lorazepam the correlation between pure and resolved spectra decreases for HPLC runs with a pure fraction below 0.10. Although this effect is observed for non-augmented as well as augmented curve resolution, it is much less pronounced for the latter. Conversely, for clonazepam, the decrease in correlation appears to be similar for non-augmented and augmented curve resolution. However, for non-augmented curve resolution the correlation decreases immediately if the peak is not pure, while this effect is only observed for augmented curve resolution for HPLC runs with a pure fraction below 0.10. Although the reason for this difference in behaviour is unknown, it is assumed that the peak order in MEKC plays a role. Clonazepam is the second peak in all MEKC runs (Table 3). Consequently, this compound is very poorly separated under all MEKC conditions. Therefore, combination of HPLC and MEKC data only has a limited effect on the resolution of clonazepam.

Augmentation of HPLC and MEKC data appears to be particularly useful when the pure fraction of an HPLC peak is less than 0.10. This seems logical, since curve resolution is not required for completely separated peaks, and, since non-augmented curve resolution is usually sufficient for peaks with moderate overlap.

5. Conclusions

First of all, it has been shown that WET is suitable for the detection of a wavelength shift between bilinear data matrices, which are obtained with different detectors of the same type. Furthermore, the method is also useful for the correction of this wavelength shift. Finally, the method provides an appropriate rank estimate as well; the number of eigenvalues that produce a constant horizontal line in the WET plot seems to reflect the total number of compounds that are present in these data matrices. Whether WET works well in the presence of spectral changes – other than a simple shift – should be subject of future research.

When curve resolution is used on HPLC and MEKC data separately, negative distortions may be observed in the obtained chromatographic and electrophoretic profiles in the case of severe overlap. The simultaneous deconvolution of chromatographic and electrophoretic profiles by means of augmented curve resolution helps to avoid these artefacts. Furthermore, the quality of the obtained spectra is significantly improved.

A combination of intensity and noise normalisation may be useful for compounds that are poorly separated in HPLC and better separated in MEKC, since it increases the weight of MEKC versus HPLC. However, since noise normalisation and its accompanying decrease in signal-to-noise ratio leads to a loss of information, and, since intensity normalisation may cause a failure of the augmented curve resolution algorithm, benefits and drawbacks of normalisation should be weighed on a case-by-case basis.

Augmentation of HPLC and MEKC data appears to be particularly useful when the pure fraction of an HPLC peak is less than 0.10. This seems logical, since curve resolution is not required for completely separated peaks, and, since non-augmented curve resolution is usually sufficient for peaks with moderate overlap.

References

[1] S. Görög, Trac. Trends Anal. Chem. 22 (2003) 407.
[2] Y.Z. Liang, O.M. Kvalheim, R. Manne, Chemom. Intell. Lab. Syst. 18 (1993) 235.
[3] R. Gargallo, R. Tauler, A. Izquierdo-Ridorsa, Anal. Chem. 69 (1997) 1785.
[4] R. Tauler, A. Izquierdo-Ridorsa, E. Casassas, Chemom. Intell. Lab. Syst. 18 (1993) 293.
[5] R. Tauler, A. Izquierdo-Ridorsa, R. Gargallo, E. Casassas, Chemom. Intell. Lab. Syst. 27 (1995) 163.
[6] R. Tauler, A. Smilde, B. Kowalski, J. Chemom. 9 (1995) 31.
[7] R. Tauler, I. Marqués, E. Casassas, J. Chemom. 12 (1998) 55.
[8] A. Garrido Frenich, D. Picón Zamora, J.L. Martínez Vidal, M. Martínez Galera, Anal. Chim. Acta 449 (2001) 143.
[9] A. Garrido Frenich, D. Picón Zamora, M. Martínez Galera, J.L. Martínez Vidal, Anal. Bioanal. Chem. 375 (2003) 974.
[10] E. Casassas, R. Gargallo, I. Giménez, A. Izquierdo-Ridorsa, R. Tauler, Anal. Chim. Acta 283 (1993) 538.
[11] J. Saurina, S. Hernández-Cassou, R. Tauler, Anal. Chem. 67 (1995) 3722.
[12] R.M. Dyson, S. Kaderli, G.A. Lawrance, M. Maeder, A.D. Zuberbühler, Anal. Chim. Acta 353 (1997) 381.
[13] A. Izquierdo-Ridorsa, J. Saurina, S. Hernández-Cassou, R. Tauler, Chemom. Intell. Lab. Syst. 38 (1997) 183.
[14] A. de Juan, A. Izquierdo-Ridorsa, R. Gargallo, R. Tauler, G. Fonrodona, E. Casassas, Anal. Biochem. 249 (1997) 174.
[15] A. de Juan, A. Izquierdo-Ridorsa, R. Tauler, G. Fonrodona, E. Casassas, Biophys. J. 73 (1997) 2937.
[16] R. Gargallo, R. Tauler, A. Izquierdo-Ridorsa, Biopolymers 42 (1997) 271.
[17] R. Tauler, Chemom. Intell. Lab. Syst. 30 (1995) 133.
[18] S. Lacorte, D. Barceló, R. Tauler, J. Chromatogr. A 697 (1995) 345.
[19] A. Garrido Frenich, M. Martínez Galera, J.L. Martínez Vidal, D.L. Massart, J.R. Torres-Lapasió, K. de Braekeleer, J.-H. Wang, P.K. Hopke, Anal. Chim. Acta 411 (2000) 145.
[20] P.V. van Zomeren, A. Hoogvorst, P.M.J. Coenegracht, G.J. de Jong, Analyst 129 (2004) 241.
[21] A. Bogomolov, M. McBrien, Anal. Chim. Acta 490 (2003) 41.
[22] R.M. Latorre, J. Saurina, S. Hernández-Cassou, Electrophoresis 21 (2000) 563.
[23] S. Sentellas, J. Saurina, S. Hernández-Cassou, M. Teresa Galceran, L. Puignou, Electrophoresis 22 (2001) 71.
[24] S. Dunkerley, J. Crosby, R.G. Brereton, K.D. Zissis, R.E.A. Escott, Analyst 123 (1998) 2021.
[25] C. Bessant, R.G. Brereton, S. Dunkerley, Analyst 124 (1999) 1733.
[26] C.P. Wang, T.L. Isenhour, Anal. Chem. 59 (1987) 649.
[27] N.P.V. Nielsen, J.M. Carstensen, J. Smedsgaard, J. Chromatogr. A 805 (1998) 17.
[28] N.P.V. Nielsen, J. Smedsgaard, J.C. Frisvad, Anal. Chem. 71 (1999) 727.
[29] J. Forshed, I. Schuppe-Kostinen, S.P. Jacobsson, Anal. Chim. Acta 487 (2003) 189.
[30] P.H.C. Eilers, Anal. Chem. 76 (2004) 404.
[31] K.J. Johnson, B.W. Wright, K.H. Jarman, R.E. Synovec, J. Chromatogr. A 996 (2003) 141.
[32] R.N. Feudale, N.A. Woody, H. Tan, A.J. Myles, S.D. Brown, J. Ferré, Chemom. Intell. Lab. Syst. 64 (2002) 181.
[33] L. Nørgaard, Chemom. Intell. Lab. Syst. 29 (1995) 283.
[34] F. Sales, M.P. Callao, F.X. Rius, Chemom. Intell. Lab. Syst. 38 (1997) 63.
[35] J.S. Shenk, M.O. Westerhaus, Crop. Sci. 31 (1991) 1694.
[36] B. Steuer, H. Schultz, Phytochem. Anal. 14 (2003) 285.
[37] Y.D. Wang, D.J. Veltkamp, B.R. Kowalski, Anal. Chem. 63 (1991) 2750.
[38] L.A. Tortajada-Genaro, P. Campíns-Falcó, F. Bosch-Reich, Anal. Chim. Acta 446 (2001) 385.
[39] Y.D. Wang, M.J. Lysaght, B.R. Kowalski, Anal. Chem. 64 (1992) 562.
[40] E. Bouveresse, D.L. Massart, Vib. Spectrosc. 11 (1996) 3.
[41] L. Nørgaard, Talanta 42 (1995) 1305.
[42] A. Garrido Frenich, D. Picón Zamora, J.L. Martínez Vidal, M. Martínez Galera, Anal. Chim. Acta 477 (2003) 211.
[43] Y.D. Wang, B.R. Kowalski, Appl. Spectrosc. 46 (1992) 764.
[44] E. Bouveresse, D.L. Massart, Chemom. Intell. Lab. Syst. 32 (1996) 201.
[45] J. Lin, S.C. Lo, C.W. Brown, Anal. Chim. Acta 349 (1997) 263.
[46] S. Macho, A. Rius, M.P. Callao, M.S. Larrechi, Anal. Chim. Acta 445 (2001) 213.
[47] K.S. Park, Y.H. Ko, H. Lee, C.H. Jun, H. Chung, M.S. Ku, Chemom. Intell. Lab. Syst. 55 (2001) 53.
[48] F.D. Barboza, R.J. Poppi, Anal. Bioanal. Chem. 377 (2003) 695.
[49] J. Gislason, H. Chan, M. Sardashti, Appl. Spectrosc. 55 (2001) 1553.
[50] C.S. Chen, C.W. Brown, S.C. Lo, Appl. Spectrosc. 51 (1997) 744.
[51] F. Sales, A. Rius, M.P. Callao, F.X. Rius, Talanta 52 (2000) 329.
[52] F. Wülfert, W.Th. Kok, O.E. de Noord, A.K. Smilde, Anal. Chem. 72 (2000) 1639.
[53] H.W. Tan, S.D. Brown, J. Chemom. 15 (2001) 647.
[54] R. Goodacre, E.M. Timmins, A. Jones, D.B. Kell, J. Maddock, M.L. Heginbothom, J.T. Magee, Anal. Chim. Acta 348 (1997) 511.
[55] F. Despagne, D.L. Massart, M. Jansen, H. van Daalen, Anal. Chim. Acta 406 (2000) 233.
[56] T.B. Blank, S.T. Sum, S.D. Brown, Anal. Chem. 68 (1996) 2987.
[57] H.W. Tan, S.T. Sum, S.D. Brown, Appl. Spectrosc. 56 (2002) 1098.
[58] S.T. Sum, S.D. Brown, Appl. Spectrosc. 52 (1998) 869.
[59] B.G.M. Vandeginste, W. Derks, G. Kateman, Anal. Chim. Acta 173 (1985) 253.
[60] K.V. Mardia, J.T. Kent, J.M. Bibby, Multivariate Analysis, Academic Press, London, 1979, p. 268.

[61] F. Gong, Y.Z. Liang, Y.S. Fung, F.T. Chau, J. Chromatogr. A 1029 (2004) 173.
[62] S. Wold, Technometrics 16 (1974) 1.
[63] F. van der Meer, W. Bakker, Remote Sens. Environ. 61 (1997) 371.
[64] F. Cuesta Sánchez, J. Toft, O.M. Kvalheim, D.L. Massart, Anal. Chim. Acta 314 (1995) 131.
[65] S. Budavari, M.J. O'neil, A. Smith, P.E. Heckelman, J.F. Kinneary, The Merck Index, Merck Research Laboratories, Whitehouse Station, NJ, 1996, pp. 404, 953 and 1129.
[66] R.M. Seifar, J.M. Kraak, W.Th. Kok, Anal. Chem. 69 (1997) 2772.
[67] J.K. Strasters, H.A.H. Billiet, L. de Galan, B.G.M. Vandeginste, G. Kateman, J. Chromatogr. 385 (1987) 181.