Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Chromatography A, 1065 (2005) 301–313
JOURNAL OF CHROMATOGRAPHY A
www.elsevier.com/locate/chroma

Capillary electrophoresis enhanced by automatic two-way background correction using cubic smoothing splines and multivariate data analysis applied to the characterisation of mixtures of surfactants

Virginia Bernabé-Zafón[a], José R. Torres-Lapasió[a], Silvia Ortega-Gadea[b], Ernesto F. Simó-Alfonso[a], Guillermo Ramis-Ramos[a,*]

[a] *Departament de Química Analítica, Universitat de València, Dr. Moliner 50, 46100 Burjassot, Spain*
[b] *Químicas Oro S.A., CV-35, km 13.5, 46184 Sant Antoni de Benaixever, Spain*

Received 8 July 2004; received in revised form 5 November 2004; accepted 30 December 2004
Available online 21 January 2005

Abstract

Mixtures of the surfactant classes coconut diethanolamide, cocamido propyl betaine and alkylbenzene sulfonate were separated by capillary electrophoresis in several media containing organic solvents and anionic solvophobic agents. Good resolution between both the surfactant classes and the homologues within the classes was achieved in a BGE containing 80 mM borate buffer of pH 8.5, 20% *n*-propanol and 40 mM sodium deoxycholate. Full resolution, assistance in peak assignment to the classes (including the recognition of solutes not belonging to the classes), and improvement of the signal-to-noise ratio was achieved by multivariate data analysis of the time–wavelength electropherograms. Cubic smoothing splines were used to develop an algorithm capable of automatically modelling the two-way background, which increased the sensitivity and reliability of the multivariate analysis of the corrected signal. The exclusion of significant signals from the background model was guaranteed by the conservativeness of the criteria used and the safeguards adopted all along the point selection process, where the CSS algorithm supported the addition of new points to the initially reduced background sample. Efficient background modelling made the application of multivariate deconvolution within extensive time windows possible. This increased the probability of finding quality spectra for each solute class by orthogonal projection approach. The concentration profiles of the classes were improved by subsequent application of alternating least squares. The two-way electropherograms were automatically processed, with minimal supervision by the user, in less than 2 min. The procedure was successfully applied to the identification and quantification of the surfactants in household cleaners.

Keywords: Alkyl benzene sulfonates; Alternating least squares; Cocamido propylbetaine; Coconut diethanolamide; Cubic smoothing splines; Household cleaners; Orthogonal projection approach

1. Introduction

The surfactant classes coconut diethanolamide (CDEA, non-ionic secondary amides) and cocamidopropyl betaine (CAPB, zwitterionic primary amides) are frequently found in toiletries and household cleaners. CDEA and CAPB are mild surfactants that do not irritate the skin and mucous membranes, have good characteristics as boosters and stabilizers of foam, and as thickening, anti-static and wetting agents, thus are used to reduce causticity and to improve other product properties in several industries. These surfactants are frequently mixed with alkyl benzene sulfonates (ABS, anionic) to improve detergency at a reduced cost. The respective generic formulae of CDEA, CAPB and ABS are:

$$R–CO–N(CH_2CH_2OH)_2$$

$$R–CO–NH(CH_2)_3N^+(CH_3)_2CH_2COO^-$$

$$R'–CH[(C_6H_4)SO_3^-]–R''$$

* Corresponding author. Tel.: +34 963543003; fax: +34 963544436.
E-mail address: ramis@uv.es (G. Ramis-Ramos).

0021-9673/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2004.12.079

where R are linear alkyl chains having from 7 up to 17 carbon atoms in CDEA and CAPB. In ABS, the *p*-phenylsulfonate group is attached to an intermediate point of a linear alkyl chain with 10–13 carbon atoms, both R′ and R″ having at least one carbon atom. Coconut fatty acids are used in the manufacturing of CDEA and CAPB, which show alkyl chain distributions similar to that found in the coconut oil. In this work, the homologues will be named according to the alkyl chain as follows: CDEA18:1 is the CDEA homologue with a double bond and with 18 carbon atoms (those in the alkyl chain plus the carbonyl), and C18:1 is the corresponding free fatty acid.

A variety of procedures for the evaluation of ABS in industrial and environmental samples, using HPLC [1–4], GC [5] and capillary electrophoresis (CE) with hydro-organic [6–13] and non-aqueous [14,15] background electrolytes (BGEs), in some cases with both homologue and positional isomer resolution [10,12,16], have been described. Ionic, non-ionic and zwitterionic surfactants have been separated in mixtures by thin-layer chromatography [17], near-infrared spectrometry [18] and HPLC with refractometric [19] and mass spectrometric [20–23] detection. An HPLC procedure capable of separating surfactants of all classes using gradient elution in two dimensions followed by evaporative light scattering detection has been also described [24]. The amido group absorbs moderately in the UV, although the signal-to-noise ratios are low due to the very short wavelengths required (e.g., the molar absorptivities of CABP and CDEA at 210 nm are 8200 and 1200 $M^{-1}\,cm^{-1}$, respectively). However, alkyl-ethanolamides have been separated by HPLC with spectrophotometric detection at 215 nm [25], and alkylbetaines and CAPB have been evaluated in shampoos by cation-exchange HPLC with detection at 210 nm [26].

The selectivity of HPLC and CE separations can be highly enhanced by recording the full spectra of the solutes with a UV–vis diode-array detector, followed by multivariate data treatment of the resulting time–wavelength matrix. This approach has been widely applied to HPLC data, where a variety of chemometric tools focussed on the deconvolution of overlapped signals have been reported [27–29]. The use of chemometric methods for CE data analysis has been revised [30]. In CE, a multivariate curve resolution—alternating least squares (ALS) procedure was applied to the deconvolution of the overlapped peaks of dinitrotoluene isomers [31]. In a previous work, we described the resolution of mixtures of absorbent ABS and non-absorbent anionic surfactants using either instrumental or chemometric approaches [32]. The peak purity assay known as orthogonal projection approach (OPA) and the bilinear decomposition method ALS were used in the chemometric procedures. In this work, a CE method, capable of resolving mixtures of CDEA, CAPB and ABS, is reported. First, the BGE was optimised to resolve both the three surfactant classes and the homologues within the classes. Then, the resolution and the reliability of peak assignments to the classes (including the recognition of solutes not belonging to the classes) were largely improved by multivariate data analysis of the two-way electropherograms. The enhancement of resolution relied on the similarity of the spectra of the homologues within the same surfactant class regarding to the much greater spectral differences among the three classes.

However, with many samples the success of the deconvolution was seriously compromised by the severe fluctuations and drift showed by the background along both measuring orders. Algorithms based on adaptive Kalman filters [33], wavelet transforms [34] and principal component modelling [35,36] have been described for the multivariate background correction of diode-array HPLC and CE signals. Using principal component modelling, a procedure with the capability of performing background correction automatically, without prior knowledge about the problem, and without using standards or reference mixtures, has been also described [36]. In this work, we proposed the use of cubic smoothing splines (CSS), a class of non-parametric regression method capable of modelling irregular profiles of all kinds with excellent accuracy [37]. Using CSS, an algorithm capable of automatically establishing an accurate model of the two-way background was constructed. The model was used to obtain an improved signal by subtraction, as well as to ameliorate the reliability and sensitivity of the deconvolution process. CSS are rather popular in engineering, computational and environmental sciences, but have been scarcely used in analytical chemistry. Applications in titrimetry [38], stripping voltammetry [39] and HPLC [40] have been reported.

Efficient background modelling yielded a quality corrected two-way signal, which made possible the application of OPA within extensive time windows. This increased the probability of finding selective spectra of the best available quality for each solute class, thus incidentally improving the refined spectra and the reconstructed electropherograms of the individual classes which were obtained by the subsequent application of ALS. Application to household cleaning products was demonstrated.

2. Materials and methods

2.1. Apparatus, software, reagents and samples

A diode-array spectrophotometer (model HP 8453, Agilent Technologies, Waldbronn, Germany) provided with a standard 1 cm quartz cell, and a capillary electrophoresis system (model G1600A, Agilent Technologies) provided with a diode-array spectrophotometric detector and fused-silica capillaries (Composite Metal Services, Ilkley, UK) of 48.5 cm (40 cm effective length) × 50 μm i.d. (363 μm o.d.), were used. The signals, exported as ASCII files by the HPCE3D software (Agilent Technologies), were treated with a collection of home-made data analysis subroutines which were written in MATLAB 6.5 (The Mathworks, Natick, MA, USA).

The following analytical grade reagents were used: methanol, acetonitrile (ACN), *n*-propanol (PrOH), *n*-butanol

(Scharlab, Barcelona, Spain), dipentylamine (DPA), sodium deoxycholate (SDC), sodium dodecylsulfate (SDS), dioctyl sulfosuccinate (DOSS), Brij 35 and caprylic (C8:0), capric (C10:0), lauric (C12:0), oleic (C18:1), linoleic (C18:2) (Fluka, Buchs, Switzerland), myristic (C14:0), palmitic (C16:0) (Sigma–Aldrich, Steinheim, Germany), boric, acetic, phosphoric and 3,4,5-trimethoxybenzoic (TMBA) acids. Deionized water (Barnstead deionizer, Sybron, Boston, MA, USA) was also used.

The samples (provided by Químicas Oro S.A., Sant Antoni de Benaixever, Spain) were: AMIDET B-112 (92.8% CDEA, taken from two different batches), BETADET HR (ca. 30% CAPB) (Kao, Barcelona, Spain), and 4-dodecylbenzenesulfonic acid (technical mixture of ABS homologues, ∼90% purity, Fluka, Buchs, Switzerland). CAPB and ABS solutions were prepared in water, and PrOH–water (1:4, v/v) was used for CDEA. Mixtures of 3 mg ml^{-1} CDEA, CAPB and ABS stock solutions were used in the optimisation studies. The respective average molecular masses (used to calculate molar absorptivities) were 301 and 356 g mol^{-1} for CDEA and CAPB (calculated from the respective homologue distributions observed on the electropherograms) and 326.5 g mol^{-1} for ABS (from the Fluka catalog).

2.2. Procedures

The recommended BGE is an 80 mM borate hydro-organic buffer (from boric acid adjusted to pH 8.5 with 0.1 M NaOH) containing 40 mM SDC and 20% PrOH. New capillaries were flushed at 100 kPa with 1 and 0.1 M NaOH and water at 60 °C (10 min each). Daily before use and between consecutive runs, the capillary was conditioned with 0.1 M NaOH (5 min), water (2 min) and the running buffer (3 min). All solutions were filtered through a 0.45 μm pore-size nylon filter (Albet, Barcelona). Hydrodynamic injection (5 kPa × 2 s) was used, and separations were performed at 25 °C under either +15 kV(alkaline media) or −15 kV (acidic media). Monitoring was performed at 200 nm with 450 nm as reference (20 and 80 nm bandwidths, respectively), and the full diode-array-detector signals were also stored for further off-line data analysis. After each working session, the capillaries were flushed with 0.1 M NaOH (5 min), water and air (2 min each).

Standards of the CDEA and CAPB homologues are not commercially available, thus, the fatty acids were liberated by boiling the solutions several minutes at pH 10, and analysed using pure fatty acids as standards. Since CDEA is not soluble in water, in this case boiling was performed in the presence of ca. 20% PrOH, with reposition of the evaporated alcohol when precipitation was observed. The CE separation conditions for the free fatty acids were adapted from literature [9,41]; thus, the hydrolysed samples were injected hydrodynamically (5 kPa × 3 s) in a capillary (80.5 cm × 50 μm i.d.) filled up with a BGE constituted by 5 mM TMBA, 7 mM DPA, 10 mM Brij 35 and 60% ACN. Separation of the free fatty acids was performed at 45 °C under +20 kV, using indirect detection at 380 nm with 245 nm as reference (80 and 10 nm bandwidths, respectively). The free fatty acid distributions found in the hydrolysates were useful to support peak identification of the homologues of the corresponding surfactant classes.

3. Results and discussion

3.1. Optimisation of the CE separation conditions

Optimisation was performed both in alkaline (80 mM borate buffer of pH 8.5, and 50 mM DPA buffer of pH 10) and acidic (10 and 20 mM phosphoric acid) media. As indicated in Table 1, in both cases several organic solvents (added to inhibit micellisation) and anionic solvophobic agents at different concentrations were tried. In each case, solutions of

Table 1
Optimisation of the BGE in alkaline and acidic media[a]

BGE composition	R_1	t	Comments
80 mM Borate (pH 8.5), 20% PrOH and 20, 40 or 75 mM SDC	–	–	Optimal in alkaline media (40 mM SDC)
80 mM Borate (pH 8.5), 20% PrOH and 100 mM SDC	0	↑↑	CDEA/CAPB peaks delayed up to 45 min
80 mM Borate (pH 8.5), 20% PrOH and 40 mM fatty acid[b] or 30 mM DOSS	↓	↑	Efficiency decreases
80 mM Borate (pH 8.5), 75 mM SDC and 30 or 40% PrOH	↓	↑	ABS peaks delayed
80 mM Borate (pH 8.5), 75 mM SDC and 40 or 50% ACN or MeOH	↓	↑	Noise increases
25 mM HOAc, 50 mM DPA (pH 10), 75 mM SDC or 50 mM SDS and 50 or 70% CAN	↓↓	0	Borate not soluble in 50 and 70% ACN
80 mM Borate (pH 8.5), 15% PrOH and 40 mM SDS	↓	↑↑	R_2 decreases largely
80 mM Borate (pH 8.5), 40 mM SDS and 30 or 20% PrOH	↑	↑	R_2 decreases largely
10 mM Phosphoric acid (pH 2), 20% PrOH and 10 mM DOSS	–	–	Optimal in acidic media
10 mM Phosphoric acid (pH 2), 20% PrOH and 30 mM DOSS or 10 mM SDS	↓	↑	Noise increases
10 mM Phosphoric acid (pH 2), 20% PrOH and 40 mM SDS	↓↓	↓	
10 mM Phosphoric acid (pH 2), 30% PrOH and 10 mM SDS	↑	↑↑	
20 mM Phosphoric acid (pH < 2), 20% PrOH and 40 mM SDS	↓↓	↓	R_2 decreases

[a] R_1: resolution within classes for CDEA and CAPB; R_2: resolution within classes for CDEA and CAPB with respect to ABS; t: analysis time for a mixture of the three solute classes. Other symbols: increases (↑), decreases (↓), effect not significant (0) with respect to the optimal BGE; SDC is not soluble in acidic media.

[b] The lauryc, miristic and palmitic acids were also assayed as solvophobic agents.

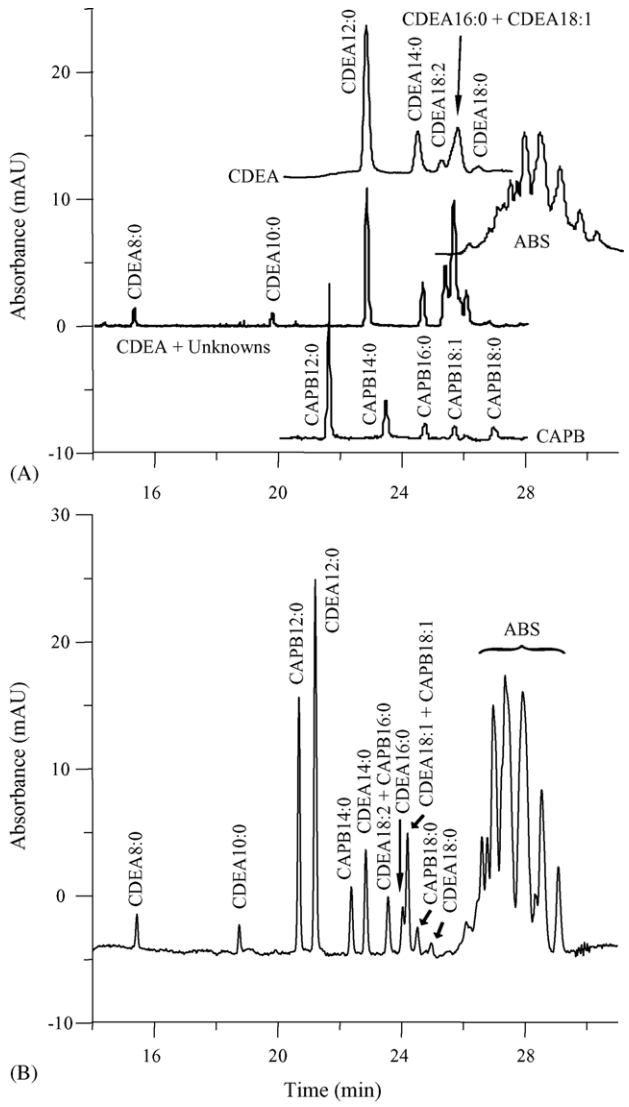

Fig. 1. Electropherograms of CDEA without and with unknown impurities, CAPB and ABS (A), and a mixture of the three surfactant classes in the absence of the CDEA impurities (B) (1000 μg ml^{-1} each) obtained with the optimal BGE in alkaline medium (80 mM borate buffer of pH 8.5, 20% PrOH and 40 mM SDC); EOF at ca. 11.3 min.

CDEA, CAPB and ABS, and mixtures of two and the three surfactant classes were injected.

In the alkaline medium, the best separations of the homologues of CDEA and CAPB mixtures were achieved in the presence of SDC with a 20% PrOH. Electropherograms of the three surfactant classes, injected both individually and in a mixture, and obtained with this BGE are presented in Fig. 1. Separation of the non-ionic CDEA and zwitterionic CAPB homologues was attributed to solvophobic association with SDC. These solutes and ABS move upstream, the cathodic EOF carrying them all towards the detector. Peak identification was made according to the fatty acid composition found in the hydrolysates (obtained as indicated in Section 2.2), and it was also supported by the regular decrease of the electrophoretic mobility of the homologues as the number of carbon atoms increased (plots not shown). Further, both CDEA and CAPB showed a large amount of the lauric acid derivative and progressively lower amounts of the C14:0, C16:0, C18:1, C18:0 and C18:2 derivatives, as occurs in coconut oil [42]; CDEA additionally showed the C10:0 and C8:0 derivatives, also present in coconut oil. However, the conclusive peak assignments of the figure were only possible after deconvolution of the electropherograms of both the mixtures and the raw materials. Apparently with aging, and maybe due to formation of hydroperoxides, the CDEA samples also showed peaks of unknown impurities, whose spectra (with additional maxima ca. 245 nm) were quite different from that of the CDEA homologues. As shown in Fig. 1A, the peaks of the impurities appeared within the 24–26 min region, thus overlapping with either the last peaks of CDEA and CAPB or the first ABS peaks.

In addition, total overlapping of the CDEA18:2 and CAPB16:0 peaks, and total or partial overlapping of CDEA16:0, CDEA18:1 and CAPB18:1 peaks, were observed. It was not expected to resolve the 20 positional isomers that are present in the ABS samples [10,12], but at least to observe the several groups that are resolved in most HPLC and CE separations. Up to nine groups of ABS homologues were in fact recognized in the electropherograms of Fig. 1. The resolution between the CDEA and CAPB peaks improved upon substitution of SDC by SDS; however, the ABS peaks migrated faster with SDS, thus extensively overlapping with the peaks of the two other surfactant classes. Therefore, in comparison to the most common solvophobic agent, SDS, the different selectivity among classes provided by SDC yielded the best separation of both the three surfactant classes and the homologues within the classes.

The best separations in an acidic medium were achieved with a BGE containing 10 mM phosphoric acid (pH 2), 10 mM DOSS and 20% PrOH. All the CDEA and CAPB peaks were well resolved, showing the derivatives of the C18:0–C12:0 fatty acids within the 16–24 min region, but giving rise to much longer times for the lighter homologues (e.g. 55 min for CDEA8:0). Owing to the EOF suppression by the low pH, negative polarity was required to observe the peaks. Thus, separation of the CDEA homologues in acidic medium was exclusively due to solvophobic association with the anionic surfactant added to the BGE. Instead of this, a mixed mechanism should be attributed to the CAPB homologues, which were in the cationic form due to protonation of the carboxylate groups; however, the net positive charge of the quaternary ammonium was not enough to delay the CAPB homologues respecting most CDEA homologues. Further, the ABS peaks appeared between 15 and 20 min thus overlapping most CDEA and CAPB peaks (electropherograms not shown). For this reason, as well as due to the excessively long analysis time required to observe the CDEA lighter homologues, the optimal alkaline medium was selected. In Table 1, the separations achieved in other alkaline and acidic conditions are compared with reference to the electropherograms obtained in the respective optimal BGEs.

Time \ Wavelength	Electropherogram at λ_1 : $\mathbf{c_1}$	Electropherogram at λ_j : $\mathbf{c_j}$	Electropherogram at λ_n : $\mathbf{c_n}$	Mean electropherogram: $\mathbf{c_{n+1}}$	St. dev. electropherogram: $\mathbf{d_s}$
Spectrum at t_1: $\mathbf{s_1}$	x_{11}	x_{1j}	x_{1n}	$\bar{x}_1$	d_{s1}
...	...	...	...	...	...
Spectrum at t_i: $\mathbf{s_i}$	x_{i1}	x_{ij}	x_{in}	$\bar{x}_i$	d_{si}
...	...	...	...	...	...
Spectrum at t_m : $\mathbf{s_m}$	x_{m1}	x_{mj}	x_{mn}	$\bar{x}_m$	d_{sm}
Spectrum of the maxima: $\mathbf{s_M}$	x_{M1}	x_{Mj}	x_{Mn}		$\bar{x}_{ds} \pm d_{ds}$
Backgr. st. dev. spectrum: $\mathbf{d_0}$	d_{01}	d_{0j}	d_{0n}		

Fig. 2. Data matrix **X** constituted by *m* spectra or *n* electropherograms, and other vectors and scalars used.

3.2. Examination of the solute spectral features in order to apply multivariate deconvolution

After optimisation of the working conditions, several CDEA and CAPB peaks still remain overlapped, and peaks of CDEA impurities when present also overlapped with either CDEA homologues or ABS peaks. As explained next, full peak resolution, assistance in peak assignment to the solute classes (including recognition of unknowns) and improvement of the signal-to-noise ratio were accomplished by multivariate data treatment techniques. Electropherograms of CDEA, CAPB, ABS and their mixtures were processed. For each injection, a matrix, **X**, including *m* spectra at *n* wavelengths was stored. As indicated in Fig. 2, **X** can be considered as constituted by *n* vectors $\mathbf{c_j}$ ($1 \le j \le n$), each one containing an electropherogram at a given λ_j wavelength (one per detector diode), or either by *m* vectors $\mathbf{s_i}$ ($1 \le i \le m$), each one containing a spectrum at a given t_i migration time. Other vectors and scalars also used in the proposed algorithms and calculations performed below were: (a) the spectrum formed by the maximal value of each $\mathbf{c_j}$ vector, $\mathbf{s_M}$, or spectrum of the maxima; (b) the points selected as true background points for each particular $\mathbf{c_j}$ vector, which constituted the experimental estimations of the respective background electropherograms, $\mathbf{c_{j0}}$; (c) the standard deviations along the $\mathbf{c_{j0}}$ vectors, constituting the background standard deviation spectrum, $\mathbf{d_0}$; and (d) the mean and standard deviation of each $\mathbf{s_i}$ vector, respectively constituting the mean and standard deviation electropherograms, $\mathbf{c_{n+1}}$ (or $\mathbf{c_j}$ where $j = n + 1$) and $\mathbf{d_s}$. The mean and standard deviation of the elements of $\mathbf{d_s}$, $\bar{x}_{ds}$, and d_{ds}, were also used.

The deconvolution of mixtures of surfactant classes that is proposed in this work is based on the assumption that a single reference spectrum can be used to describe all the peaks of a single surfactant class, being the spectral differences between the homologues of a given class negligible in relation to the much larger differences between classes. The overall spectra of the three surfactant classes of concern, obtained with a spectrophotometer, are shown in Fig. 3 (dotted lines with the *Y*-axis in molar absorptivity units at the right side of A–C parts). The spectra of the alkyl amides and that of aromatic ABS differed largely, but those of the secondary (CDEA) and primary (CAPB) amides showed smaller differences. The similarity between pairs of spectra can be estimated using linear correlation coefficients; thus, in the 200–260 nm wavelength range: $r = 0.75$, 0.74 and 0.95 for the CDEA/ABS, CAPB/ABS and CDEA/CAPB pairs of overall spectra, respectively.

For a given solute class, a spectrophotometer yields a single spectrum which is an average of the spectra of the homologues, thus not showing the spectral dispersion among them. To check the significance of the differences within each par-

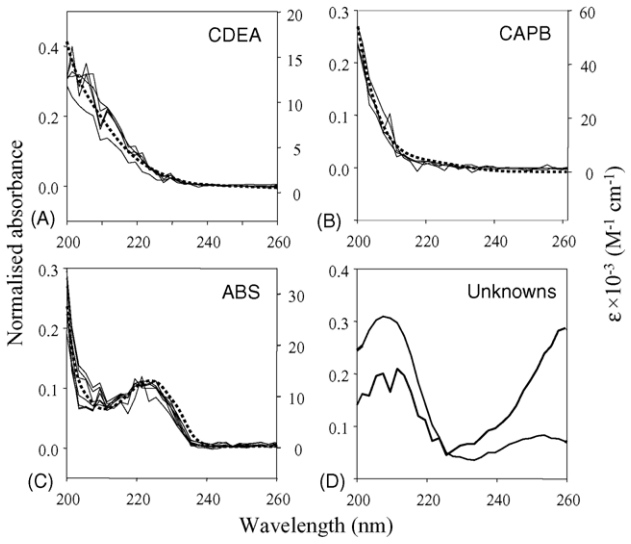

Fig. 3. Overlaid normalised spectra (divided by their respective modulus) extracted from the **X** matrices of CDEA, CAPB and ABS raw materials at the time locations of the respective homologue peaks, and spectra of two unknown solutes found in a CDEA sample. The dotted lines are the respective absorption spectra of CDEA, CAPB and ABS (obtained with the HP 8453 spectrophotometer using solutions prepared in the optimal BGE at pH 8.5) divided by the respective molar concentrations (35, 15 and 40 μM, respectively; molar absorptivity axes at the right).

ticular class, the spectra at the maxima of all peaks along the CDEA, CAPB and ABS electropherograms were extracted from the respective **X** matrices. Then, each spectrum was normalised by dividing the absorbances by its modulus, i.e. by the square root of the sum of the squares of the absorbances. Normalisation removes the influence of concentration in solutes exhibiting the same spectrum. When the solutes have a different spectrum, normalisation attenuates the differences caused by both the concentration and the molar absorptivity of the maxima, thus also helping in the appreciation of the spectral shape differences. The normalised spectra of the CDEA, CAPB and ABS homologues are shown overlaid in Fig. 3A–C. The spectra of the individual homologues and those measured with the spectrophotometer for the whole class were essentially the same; however, owing to the much smaller signals measured in the capillary, the former showed a higher noise. The spectra of the homologues were effectively closely similar within the classes, the inter-class differences being larger between the CDEA and CAPB homologues, and much larger between them and the ABS homologues. The small systematic differences which were observed between the successive groups of the ABS homologues could be explained as due to the different isomeric distribution within them.

Fig. 3D, shows the spectra of two unknown absorbent impurities which were found in aged CDEA samples. Experiments performed with mixtures of the three surfactant classes prepared using both new and aged CDEA samples will be used in the discussion that follows. The latter were useful to illustrate the capability of the algorithms to address more complex cases. In fact, at least four reference spectra extracted from **X** by using OPA were required to deconvolve the mixtures prepared in the presence of CDEA impurities, whereas three reference spectra were enough to account for the mixtures prepared without these impurities.

3.3. Automatic background modelling and correction

Background disturbances (low-frequency fluctuations and drift) generated by the individual diodes of the detector array were found to be of utmost relevance compromising the success in the multivariate treatment of the weak signals constituting the electropherograms. As shown in Fig. 4A, the background signal of the diodes consisted of both a high-frequency random noise and a slow drift with occasional low-frequency fluctuations. Further, since the diodes of an array are basically independent devices, the disturbances from consecutive diodes were scarcely correlated each other (see the insert in Fig. 4A). Thus, the background disturbances along both the time and wavelength ways of the electropherograms showed irregular and unpredictable shapes. The magnitude of the disturbances were significant enough to require careful correction, particularly below 220 nm, where they amounted to more than 10% of the largest solute peaks. It should be taken into account that all the significant information concerning the CDEA and CAPB homologues was obtained at

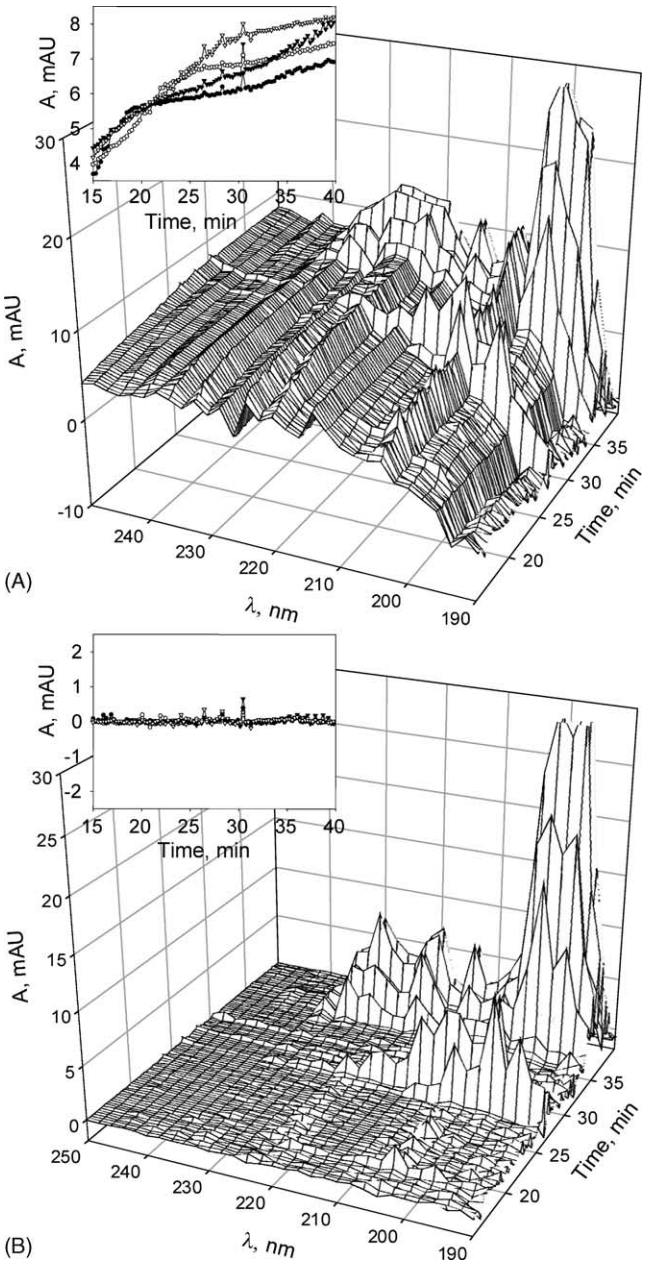

Fig. 4. Three-dimensional view of matrix **X** for a mixture containing CDEA, CAPB (300 μg ml^{-1} each) and ABS (900 μg ml^{-1}) before (A) and after (B) background correction using CSS; to improve the view, only one of every 20 spectra was plotted. The inserts show the two-dimensional electropherograms of the four diodes located at the longest wavelengths.

these short wavelengths. Moreover, when replicated electropherograms were compared, the differences among the background disturbances led to different spectra for the same solute class. Therefore, a model of the two-way background should be necessarily estimated for each particular electropherogram.

An initial estimate of the background matrix, $\mathbf{X_0}$, can be constructed by sampling either the time domain at each individual wavelength or the wavelength domain at each particular time value. However, when UV–vis spectra are recorded

in solution, the wide bands extended along large spectral regions, and the intense noise and generalised presence of absorption bands at the short wavelength end, prevent the adequate background sampling along the wavelength domain. On the contrary, as illustrated in Fig. 4A, background points bracketing the peak regions all along the time domain were present in all the electropherograms. Thus, to construct $\mathbf{X_0}$, the limits of the working domain along the time way were first established by using the standard deviation electropherogram, $\mathbf{d_s}$, which shows the time regions where significant ab-

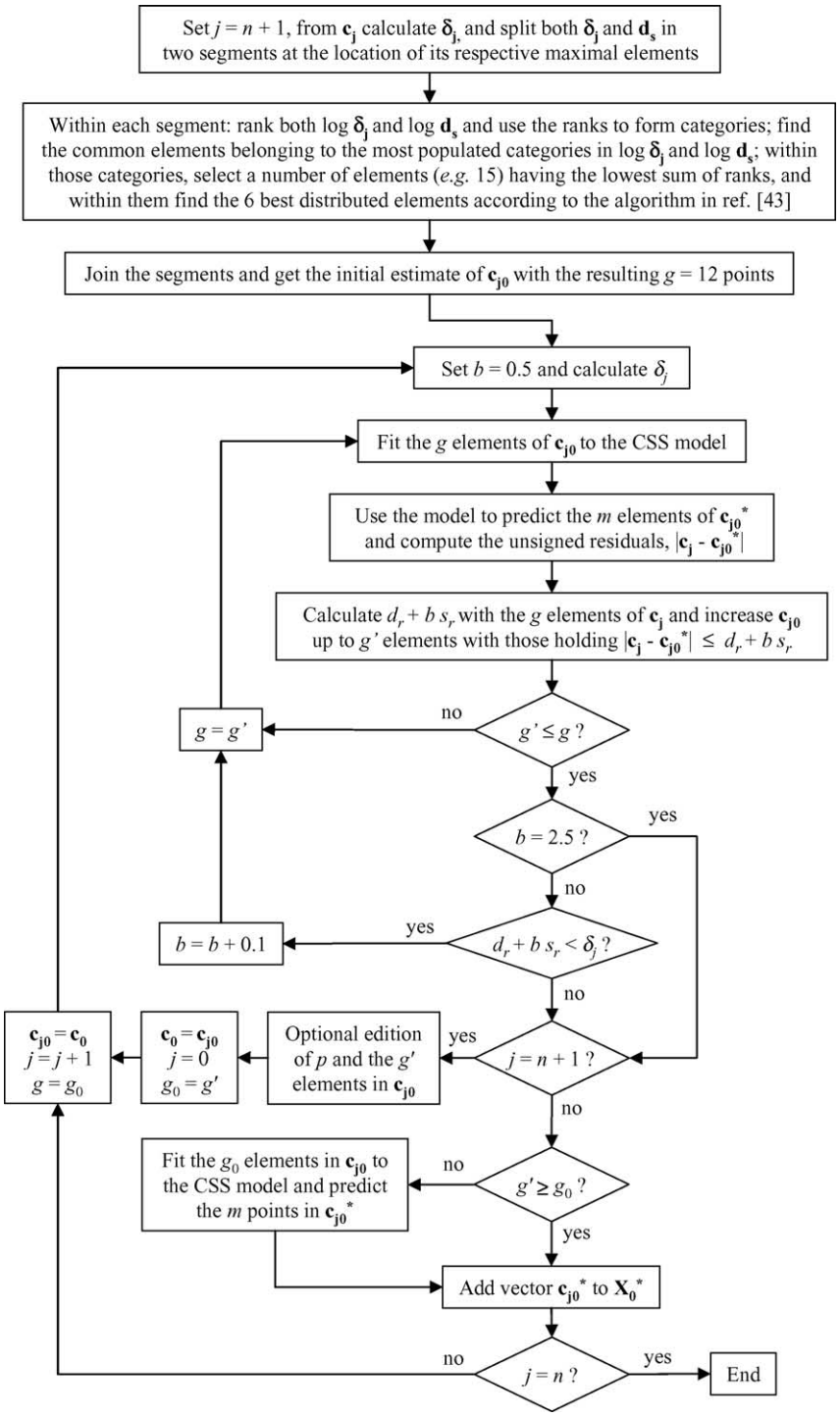

Fig. 5. Flow diagram describing the background selection algorithm. Symbols: $\boldsymbol{\delta_j}$, vector of the deviations of x_{ij} with respect to the mean of $x_{(i-1)j}$ and $x_{(i+1)j}$, being $x_{0j} = x_{1j}$ and $x_{(m+1)j} = x_{mj}$; δ_j, median of the unsigned elements of $\boldsymbol{\delta_j}$; d_r and s_r, mean and standard deviation of the unsigned residuals, $|\mathbf{c_j} - \mathbf{c_{j0}^*}|$; g_0, number of elements of the definitive mean electropherogram baseline, $\mathbf{c_0}$; subindex 0 denotes baseline and an asterisk means fitted to a model; other symbols are as in Fig. 2 or self-explained.

sorbance at least at one wavelength region exists. The working time domain was established at will from the first to the last of these regions, plus two small regions at the ends.

A thoroughly detailed and highly reliable background model was conveniently and automatically constructed with the algorithm which is detailed in the flow diagram of Fig. 5. Rather conservative criteria are applied by the algorithm all along the point selection process, thus to maintain a high probability of including only true background points. First, a rough baseline common for all the diodes, including all the time regions exempted of solute peaks at any wavelength, $\mathbf{c_{j0}}$ (being $j = n + 1$, thus the points being selected from the mean electropherogram), is established. For this purpose, the minimal values in the following two vectors are attended: $\boldsymbol{\delta_j}$, which contains the deviations of a point in $\mathbf{c_{j0}}$ with respect to its two closest neighbours, and $\mathbf{d_s}$. These vectors are particularly sensitive to the presence of outliers, thus their minimal elements have a high probability of exclusively including true background points. Each one of the two vectors, $\boldsymbol{\delta_j}$ and $\mathbf{d_s}$, is next split in two segments at the location of their respective maximal elements. The points to be initially included in $\mathbf{X_0}$ are selected by independently processing each one of these four segments. In this way, a better distribution of the selected points all along the time domain is assured.

Then, the points within the segments are ranked by categories. Within each segment, this accumulates the points with the highest probability of belonging to the true background within the most populated category. Only a small set of points picked up from the lowest ranked elements common to the two vectors, $\boldsymbol{\delta_j}$ and $\mathbf{d_s}$, attending only to their respective most populated category, are included in $\mathbf{c_{j0}}$. Since $\boldsymbol{\delta_j}$ and $\mathbf{d_s}$ are defined by rather different algorithms, the probability of exclusively selecting true background points is greatly increased by this condition. This probability is further enhanced by reducing the number of selected points to a minimum. We decided to select only six common points at the left side of the maximal elements of $\boldsymbol{\delta_j}$ and $\mathbf{d_s}$, and six more points at the right side, thus coming to 12 points in the initial mean background sample, $\mathbf{c_{j0}}$ (indicated as $g = 12$ in Fig. 5). Further, these points are selected with the condition of being located as far apart as possible from each other, which assures the adequate sampling of the background all along the time domain, also including points close to the ends of the domain. The Kennard–Stone algorithm [43] was used to automatise the selection of these points within each side of $\mathbf{c_{j0}}$.

Next, a regression algorithm is applied to obtain a model, $\mathbf{c^*_{j0}}$ ($j = n + 1$), from the 12 points initially included in $\mathbf{c_{j0}}$. As discussed in Section 3.4, CSS yielded the best models. Once a model has been constructed, the trend set by the model itself can be exploited to increase the number of points in $\mathbf{c_{j0}}$ while maintaining a high probability of including only true background points. For this purpose, the unsigned residuals of the mean electropherogram with respect to the background model, $|\mathbf{c_j} - \mathbf{c^*_{j0}}|$ (also with $j = n + 1$), are examined. Only a few points are added within each iteration, thus to avoid to bias the CSS model. When optimisation has finished, both the points in the resulting common baseline, $\mathbf{c_{j0}}$, and the CSS parameter p (see Section 3.4) can be optionally edited. The spline is respectively strengthened or loosened by increasing or decreasing p.

Then, $\mathbf{c_{j0}}$ (which now contains $g' > g$ points) is fixed (by making $g_0 = g'$ and $\mathbf{c_0} = \mathbf{c_{j0}}$) to serve as the initial estimation of the background for each particular diode ($\mathbf{c_{j0}} = \mathbf{c_0}$). Index j is then varied from 1 to n, each time repeating the same refining cycle that was used before to construct $\mathbf{c_0}$. This results in additional background points for each diode. Since the analysis is independently performed spectrum by spectrum, time drifts do not influence the process. This process progressively converts the g_0 background points included in vector $\mathbf{c_0}$ in a matrix, $\mathbf{X_0}$, which contains the n g_0 points which are simultaneously observed at the n wavelengths, plus a variable number of additional background points at each wavelength. This point selection process automatically yielded a detailed and reliable description of the two-way background, $\mathbf{X_0}$. The set of the corresponding n definitive $\mathbf{c^*_{j0}}$ models constitute the two-way background model, $\mathbf{X^*_0}$.

Finally, the limits of the wavelength domain are established by using a restricted standard deviation spectrum, $\mathbf{d_0}$, in which only background regions common to all the time values are included. A vector containing the signal-to-noise ratios is obtained by dividing element by element vectors $\mathbf{s_M}$ (the spectrum of the maxima) and $\mathbf{d_0}$. The wavelength domain is selected at the sight of this vector plotted versus the wavelength. This step was programmed to be manually performed, which is advisable to avoid calculation artefacts and misleading results produced by the significant noise affecting the spectra at short wavelengths. After selection of the wavelength working domain, both $\mathbf{X}$ and $\mathbf{X^*_0}$ were definitely set, and the corrected matrix, $\mathbf{X_c}$, used in the subsequent deconvolution studies was calculated as $\mathbf{X_c} = \mathbf{X} - \mathbf{X^*_0}$. The efficiency of the algorithm for background selection and modelling can be appreciated in Fig. 4B, where a three-dimensional plot of a corrected matrix $\mathbf{X_c}$ can be compared with the plot of the corresponding uncorrected matrix, $\mathbf{X}$, which is given in the part A of the same figure.

3.4. The CSS algorithm applied to background modelling and correction of the CE signals

In order to construct reliable background models, a variety of algorithms and modelling strategies were tried, including linear fitting within small consecutive time regions, and quadratic, cubic and polynomial fitting using several functions. Although the quality of those fittings was apparently good, minor residual disturbances often led to biased deconvolutions, especially when large time windows were processed (which as indicated above was of utmost interest to expedite data treatment). Much better results, also showing universal applicability, were achieved by CSS. This technique was specifically developed to model the underlying low-frequency trends hidden under noisy data of very different nature [37]. The splines mimic complex functions by

using a set of r local polynomial approximations, $f_r(t)$, which are piecewise self-defined giving rise to a set of $(r+1)$ knots. The degree of the polynomials is selected by the user, the cubic degree being the most usual. In smoothing splines the following likelihood estimator is minimised:

$$r_j = p_j \sum_{k=1}^{g'} [x_{kj} - f_r(t_{kj})]^2 + (1 - p_j) \int_{t_{1j}}^{t_{g'j}} \left(\frac{\mathrm{d}^2 f_r}{\mathrm{d}t^2} \right)^2 \mathrm{d}t \quad (1)$$

Eq. (1) represents a balance between two stresses, i.e. that represented by the first term, which tends to fit a curve passing exactly through all the points (applying thus a pure cubic spline), and that expressed by the second one, which tends to perform a simple linear least-squares fitting. The relative weight of the two forces is governed by the smoothing parameter, p, which yields pure splines when its value is one and linear fittings when zero. Thus, the more irregular the background the higher the optimal p-value, excellent fittings to smooth curves being obtained with low p-values. For all the cases tried in this work, p-values ranging from 10^{-4} to 10^{-7} were adequate, variations of an order of magnitude or even larger giving rise to fittings of similar quality. It was also checked that the same p-value selected to optimally fitting the background of one of the diodes in a given two-way electropherogram, also yielded satisfactory fittings for all the other diodes, including the mean electropherogram, $\mathbf{c_{n+1}}$. In agreement with the smooth appearance of the background, the low values of p used in this work approximated the CSS to a low degree polynomial rather than to a loose function. An advantage is that the new points which were conservatively added during the background construction process tended to consolidate the CSS, which subsequently reduced the probability of introducing points contaminated with significant signals. Subindex k refers to the rows of $\mathbf{X_0}$ (the known background values at the g' times selected as baseline points for the considered diode, being $g' < m$), and index j indicates the diode (or the mean electropherogram, $\mathbf{c_{n+1}}$, when $j = n+1$) whose baseline is being fitted.

3.5. Multivariate deconvolution by OPA-ALS

Two steps are involved in two-way multivariate curve resolution, namely, a purity assay followed by deconvolution. In the purity assay, a set of the purest spectra that are present in $\mathbf{X_c}$ (ideally, one for each solute class) is established. During deconvolution, iterative least-squares are used to retrieve the pure contributions of the underlying compounds from the estimates provided by the purity assay. The efficiency and reliability of the deconvolution depends critically on: (a) the adequate correction of the background disturbances; (b) the identification of all the solute classes giving different spectra; (c) the differences among the spectra; and (d) the signal-to-noise ratios of the spectra used as initial estimates. Thus, all strategies focussed to improve the performance of the purity assay are highly beneficial.

When purity assays are used, the probability of finding selective regions with high signal-to-noise ratios, containing uncontaminated or scarcely contaminated spectra, increases with the size of the time domain being processed. Concerning to solute classes constituted by homologues having highly similar spectra (i.e. able to be represented by a single reference spectrum), processing greater time domains increases the probability of finding isolated spectra with good signal-to-noise ratios of each of the classes. This way, selective and relatively distant time regions can provide the spectral information required to deconvolve overlapped peaks of a given solute class located at other time regions. Further, the quality of the information retrieved at distant time regions improves after appropriate background subtraction. Thus, to process the data on the whole time domain, instead of splitting it, is recommended in order to improve the quality of the initial estimates of the spectra, and to reduce processing time. Actually in this work, full signal processing extended along the entire time domain worked properly with all the electropherograms. When small signals should be quantified, a splitting of $\mathbf{X_c}$ in two sections yielded more accurate results owing to both the reduction of local complexity and the attenuation of the influence of diode irregularities. Also, processing small time windows where complex peak clusters appeared was occasionally required to enhance resolution.

Owing to its excellent characteristics, the hyphenated OPA-ALS technique was selected. OPA extracts from $\mathbf{X_c}$ their purest spectra, which are subsequently adjoined to form an auxiliary matrix. This matrix, which is iteratively completed, contains a collection of reference spectra, all of them being as much mutually orthogonal as possible. The property monitored to assess which spectrum in $\mathbf{X_c}$ is the least compatible with regard to those in the collection of reference spectra is called the dissimilarity. This assessment is the determinant of product $\mathbf{Y}^{\mathrm{T}} \cdot \mathbf{Y}$, where $\mathbf{Y}$ is an auxiliary matrix containing the normalised reference spectra plus the target spectrum being analysed. The process is repeated up to compare all the spectra contained in $\mathbf{X_c}$ with those in the collection. The most incompatible spectrum in $\mathbf{X_c}$ with regard to those in the collection is indicated by the largest element in the vector of the dissimilarities. This spectrum is then added to the collection, and the process is repeated until only noise remains in the dissimilarity vector.

This process does not guarantee the true purity of the selected "purest" spectra; therefore, further refinement is required in most actual cases. Refinement is achieved by rotating the solution up to a number of physicochemical restrictions are met. This is frequently performed using ALS, which decomposes the bilinear matrix in their two data orders (time and wavelength) by alternating two least-squares regressions, i.e. the spectral profiles on $\mathbf{X_c}$ to obtain the concentration profiles, $\mathbf{c_j^*}$, and the concentration profiles on $\mathbf{X_c}$ to obtain the spectral profiles, $\mathbf{s_i^*}$, repeating this process alternatively:

$$\mathbf{c_j^*} = \mathbf{X_c} \cdot (\mathbf{s_i^{*T}} \cdot \mathbf{s_i^*})^{-1} \cdot \mathbf{s_i^{*T}} \quad (2)$$

$$\mathbf{s_i^*} = (\mathbf{c_j^{*T}} \cdot \mathbf{c_j^*})^{-1} \cdot \mathbf{c_j^{*T}} \cdot \mathbf{X_c} \quad (3)$$

Constraints imposed to the intermediate solutions, such as non-negativity to both data orders, usually assure convergence in meaningful results. Further details about the OPA and ALS algorithms can be found elsewhere [27,29,32]. An advantage of full signal processing along the time domain is the easy achievement of convergence without imposing too compelling constraints, e.g. selectivity on the time order is usually not necessary. Unimodality, which is the other most common restriction, cannot be applied to samples constituted by two or more peaks sharing the same spectrum. Note that a straightforward projection of the two-way electropherograms of standards is unadvisable. Owing to the large noise variations from one injection to another, and to the occasional presence of unexpected compounds with partially collinear spectra, this can lead to strongly biased solutions. In the mixtures processed in this work, such a strategy was tried and discarded.

3.6. Identification and quantification of mixtures upon deconvolution

The concentration profiles obtained by OPA-ALS, $\mathbf{c_j^*}$, are adequate to estimate the homologue distribution within each solute class; however, to build up calibration graphs, true reconstructed electropherograms should be calculated. This is achieved by multiplying each concentration profile by the respective normalised absorbances within a wavelength window which is selected from the corresponding spectral profile. The results obtained at different wavelengths within the window are finally averaged. This gives rise to electropherograms which can be directly compared with those obtained in routine work. The reconstructed electropherograms for the solute classes found in the electropherogram of Fig. 1B, but using wavelength windows selected attending to the respective optimal signal-to-noise ratios (204–210, 204–218 and 212–232 nm for CAPB, CDEA and ABS, respectively), are shown in Fig. 6. All the homologues of the three classes and the unknowns were correctly described by the reconstructed

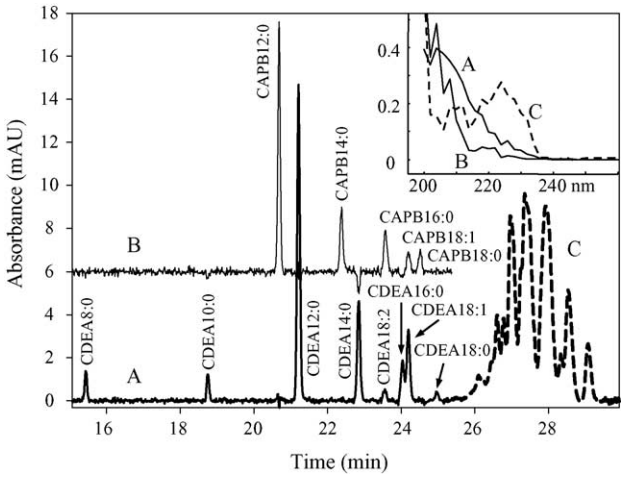

Fig. 6. Reconstructed electropherograms retrieved by OPA-ALS from the $\mathbf{X_c}$ matrix (mixture of Fig. 1B): CDEA (A, continuous line), CAPB (B, shifted upwards for clarity) and ABS (C, dashed line). The inserted plot shows the normalised spectra of A–C after deconvolution by OPA-ALS.

Table 2
Relative peak areas obtained without and with OPA-ALS deconvolution

Solute	Without OPA-ALS at 200 nm[a,b]	With OPA-ALS at optimal windows[a,c]	
		Long end injection	Short end injection
CDEA8:0	(0.11)	0.044 (0.09)	0.038 (0.06)
CDEA10:0	(0.08)	0.039 (0.08)	0.028 (0.04)
CDEA12:0	(1)	0.503 (1)	0.637 (1)
CDEA14:0	(0.29)	0.172 (0.34)	0.145 (0.23)
CDEA16:0	–	0.068 (0.14)	–
CDEA18:0	(0.03)	0.016 (0.03)	0.008 (0.01)
CDEA18:1	–	0.138 (0.27)	–
CDEA18:2	–	0.020 (0.04)	–
CDEA18:2 + 16:0 + 18:1	–	0.174 (0.35)	0.144 (0.23)
CDEA total area[d]	–	4.48 (25%)	6.22 (26%)
CAPB12:0	(1)	0.562 (1)	0.419 (1)
CAPB14:0	(0.28)	0.190 (0.34)	0.214 (0.51)
CAPB16:0	–	0.129 (0.23)	0.123 (0.29)
CAPB18:0	(0.11)	0.060 (0.11)	–
CAPB18:1	–	0.059 (0.10)	–
CAPB18:1 + 18:0	–	0.119 (0.21)	0.246 (0.59)
CAPB total area[d]	–	1.58 (9%)	1.12 (5%)
ABS total area[d]	–	11.6 (66%)	16.9 (70%)

[a] Without parenthesis, relative areas referred to the total area of the class, and between parenthesis, relative areas referred to the 12:0 homologue of the class. Relative standard deviations of the areas ranged from 10 to 16% ($n = 4$ electropherograms).
[b] Only relative areas of the resolved peaks referred to the 12:0 homologue of the class could be measured.
[c] Optimal spectral windows for quantification: 204–218, 204–210 and 212–232 nm for CDEA, CAPB and ABS, respectively.
[d] Total area of the class and percentage with respect to the sum of the areas of the three classes.

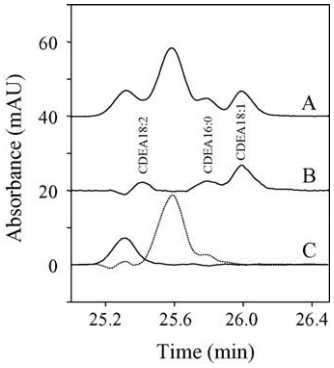

Fig. 7. Enlarged 25–26 min region of an electropherogram of an aged CDEA sample (A) and the corresponding deconvolved concentration profiles of CDEA (B) and two impurities (C) with a different spectral profile.

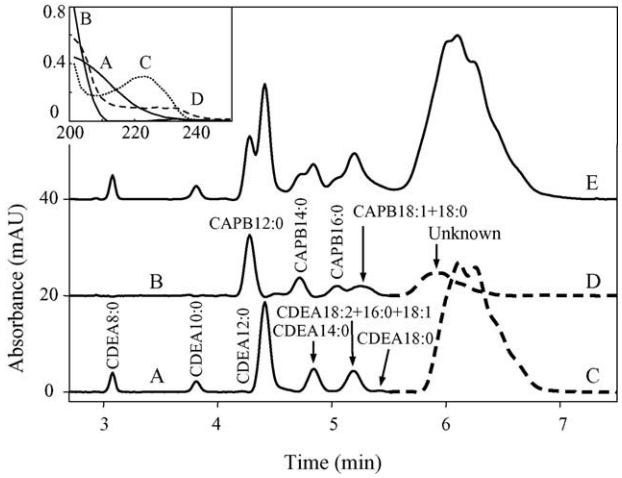

Fig. 8. Electropherogram of a mixture of CDEA, CAPB and ABS (1000 μg ml^{-1} each) injected by the short capillary end (E), and reconstructed electropherograms retrieved by OPA-ALS from the corresponding $\mathbf{X_c}$ matrix: CDEA (A), CAPB (B), ABS (C) and unknown solutes (D). The inserted plot shows the normalised spectra of A–D after deconvolution by OPA-ALS (the axis of the insert are normalised absorbance vs. wavelength in nm).

electropherograms. As also observed in the inserted plot at the upper right corner of Fig. 6, the normalised reconstructed spectra of the classes were similar to the spectra obtained with the spectrophotometer (dotted lines in Fig. 3). The relative peak areas obtained from the electropherograms both without and with application of the OPA-ALS deconvolution are compared in Table 2. Similar relationships between the peak areas of the resolved homologues within each class can be observed.

An example of resolution enhancement by local data treatment is shown in Fig. 7. In trace A, the enlarged 25–26 min region of an electropherogram of an aged CDEA sample was plotted. Trace B of the same figure shows the deconvolved concentration profile of CDEA, and trace C those of two impurities with a different spectral profile. Particularly, the CDEA18:2 peak, which was entirely hidden by the more intense signals of the impurities, was satisfactorily retrieved.

The reduction of the analysis time to a minimum is of interest for the industrial exploitation of analytical procedures. As shown in Fig. 8, trace A, analysis time was reduced from 30 to 7 min when a mixture with the same composition of that used in Fig. 1B, was injected by the short end of the capillary. As expected, resolution along the shorter time domain decreased largely, but the signal-to-noise ratio improved, leading to better spectral profiles (see the inserted plot at the upper left corner of Fig. 8). This helped in obtaining the deconvoluted concentration profiles. The reconstructed electropherograms are shown in Fig. 8, traces B–E, the latter corresponding to that of an unknown impurity. Thus, short end injection was fully valid for the incidental detection of interferences, and as shown next, was also useful for class quantitation. The relative peak areas are given in Table 2. In comparison to the results obtained with long end injection, the relationships between the peak areas of the homologues within the classes varied significantly. However, the ratio between the total area of the class and the sum of the areas of all the classes was almost the same for CDEA, and increase slightly for ABS at the cost of a significant reduction for CAPB. Thus, short end injection can be also useful for global class evaluation, although logically the errors can be large for classes giving small total areas.

In other experiments, mixtures containing 900 μg ml^{-1} of both ABS and one of the two alkyl amide surfactant classes, together with decreasing concentrations of the other class (from 900 to 300 μg ml^{-1}) were injected by the long end of the capillary and processed. To reduce the CAPB concentration while maintaining high ABS and CDEA concentrations represents the most difficult deconvolution problem, since CAPB homologues absorb within a spectral region with a higher noise. In all cases, the reconstructed electropherograms were correctly retrieved from the data matrix of the mixtures. Calibration straight-lines for total CDEA and CAPB were constructed by summing the corrected areas of the peaks of the respective homologues. The straight-lines showed a low dispersion and good linearity ($r = 0.990$, $n = 5$). Calibration straight-lines of similar quality ($r = 0.991$, $n = 9$) were also obtained by summing the corrected areas of the resolved CDEA and CAPB peaks on the unprocessed electropherograms, thus neglecting the overlapped peaks. However, this later approach increases the risk of mistakes in peak assignments to the classes, leading to biased predictions when the analysed samples had impurities or homologue distributions different than those of the standards.

The limits of detection (LODs) for the CDEA and CAPB homologues were calculated as the concentrations giving a peak height equal to three times the standard deviation of the background, s. This value was estimated from the background width at the monitored wavelength, which was taken as 5s. For individual homologues, 3s was divided by $h_i/[M_i]$, where h_i and $[M_i]$ are the peak height and the molar concentration of the i-homologue. Both the CAPB and CDEA peaks gave LODs of ca. 1.3 mM. The electropherograms of two samples

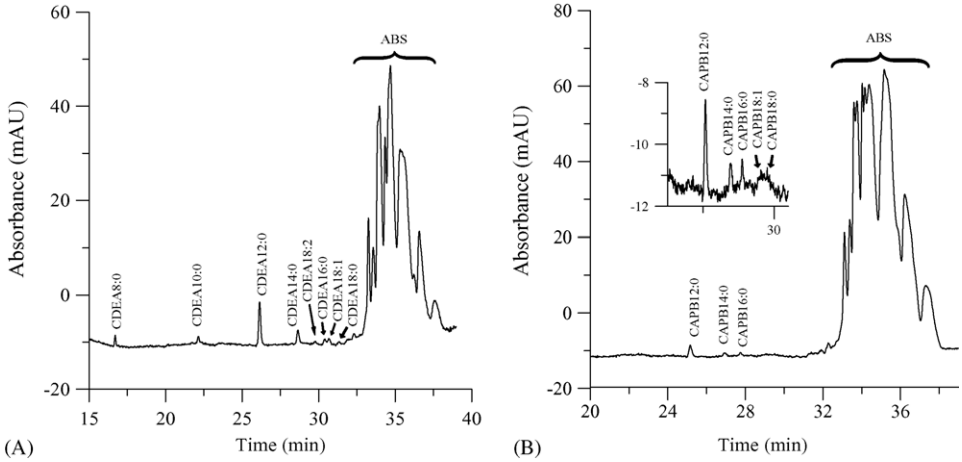

Fig. 9. Electropherograms of a cleaner for delicate clothes (A, with 9% ABS and 1.5% CDEA) and a dishwasher (B, with 16% ABS and 0.8% CAPB).

of unknown composition are shown in Fig. 9A and B. The cleaner for delicate clothes turned out to contain 9% ABS and 1.5% CDEA, and 16% ABS and 0.8% CAPB were found in the dishwasher.

4. Conclusions

The practical application of OPA-ALS multivariate deconvolution of two-way electropherograms to the separation of non-ionic, zwitterionic and anionic surfactant classes with resolution of most homologues in 30 min has been demonstrated. Data processing, including deconvolution, is seriously compromised by fluctuations and drift in the background signal of the individual diodes of the detector array. In this work, an algorithm capable of constructing a background model to be subtracted from the two-way raw signal with scarce supervision was developed. For this purpose, a small initial set of background points was selected by using two independent criteria. This set was iteratively increased for the background regions common to all the diodes with the support of CSS models. The common background obtained in this way, was improved by adding new points for all the individual diodes, also with support of CSS models. All along the whole process, many safeguards and rather conservative criteria were always applied, thus to maintain a high probability of excluding points associated to significant signals. Thanks to the parsimonious character of this process, a detailed and highly reliable description of the background, and an associated CSS quality model, were achieved. Other tried background filtering procedures led to artefacts which made unfeasible the subsequent deconvolution stages; thus, the good results obtained in this work by OPA-ALS deconvolution of the corrected two-way signals also guaranteed the reproducibility and quality of the background selection and modelling procedure here described. Further, the corrected signals showed a high sensitivity towards small solute peaks which were retrieved unbiased.

Advantages of using CSS were simplicity, reliability and robustness in the presence of disturbances of a wide variety of magnitudes and shapes. This is important, since the presence of irregular and unpredictable background disturbances of the same order of magnitude as the solute peaks has probably hindered until now the development of two-way analytical procedures in capillary electrophoresis. In comparison to the straightforward use of unprocessed data, several advantages are also achieved by applying multivariate deconvolution. First, the signal-to-noise ratio is improved by using the significant information generated at all the wavelengths. Second, owing to the use of quality spectra, the reliability in peak identification also improves, and additional information about unexpected absorbent species is further gained. Third, the risk of systematic error produced by both neglecting overlapped peaks and peak misidentification is highly reduced. Forth, analysis time can be dramatically reduced to 7 min by injecting through the short capillary end. In this case, although some resolution between the homologues of a given class was lost, the classes were properly characterised and quantified. Finally, with the home-made software, the whole process, from data file importation to MATLAB up to quantitation of the peak corrected areas on the reconstructed electropherograms, can be completed in less than 2 min.

Acknowledgements

We thank financial support from the ‘Ministerio de Educación y Ciencia’ (MEC) of Spain and the ‘Fondo Europeo de Desarrollo Regional’ of the European Union (Project CTQ2004-06302/ BQU), and from ‘Químicas Oro, S.A.’; VBZ thanks the MEC for an FPU grant; JRTL thanks the MEC and the University of Valencia for a ‘Ramon y Cajal’ Position; grant TS/03/05 of the ‘Generalitat Valenciana’ is also acknowledged.

References

[1] A. Nakae, K. Kunihiro, J. Chromatogr. 152 (1978) 137.
[2] A. Marcomini, S. Capri, W. Giger, J. Chromatogr. 403 (1987) 243.
[3] Y. Yokoyama, H. Sato, J. Chromatogr. 555 (1991) 155.

[4] Y. Yokoyama, M. Kondo, H. Sato, J. Chromatogr. 643 (1993) 169.
[5] R. Alzaga, A. Peña, L. Ortiz, J.M. Bayona, J. Chromatogr. A 999 (2003) 51.
[6] J. Zweigenbaum, Chromatographia 11 (1990) 9.
[7] P.L. Desbene, C.M. Rony, B. Desmazieres, J.C. Jacquier, J. Chromatogr. 608 (1992) 375.
[8] S. Chen, D.J. Pietrzyk, Anal. Chem. 65 (1993) 2770.
[9] V. Bernabé-Zafón, S. Ortega-Gadea, E.F. Simó-Alfonso, G. Ramis-Ramos, Electrophoresis 24 (2003) 2805.
[10] K. Heinig, C. Vogt, G. Werner, Analyst 123 (1998) 349.
[11] W.H. Ding, C.H. Liu, J. Chromatogr. A 929 (2001) 143.
[12] J.M. Herrero-Martínez, E.F. Simó-Alfonso, G. Ramis-Ramos, Electrophoresis 24 (2003) 681.
[13] R. Loos, J. Riu, M.C. Alonso, D. Barcelo, J. Mass Spectrom. 35 (2000) 1197.
[14] H. Salimi-Moosavi, R.M. Cassidy, Anal. Chem. 68 (1996) 293.
[15] J.M. Herrero-Martínez, E.F. Simó-Alfonso, G. Ramis-Ramos, Electrophoresis 22 (2001) 2017.
[16] L. Lunar, S. Rubio, D. Pérez-Bendito, J. Chromatogr. A 1031 (2004) 17.
[17] A. Kruse, N. Buschmann, K. Cammann, J. Planar Chromatogr. Mod.-TLC 7 (1994) 22.
[18] H. Grunewald, C. Kurowski, D. Timm, U. Grummisch, U. Meyhack, J. Near Infrared Spectrosc. 6 (1998) 215.
[19] K. Nakamura, Y. Morikawa, J. Am. Oil Chem. Soc. 61 (1984) 1130.
[20] Y. Miyamae, K. Yoshizawa, J. Tsuchiya, Bunseki-Kagaku 50 (2001) 61.
[21] M. Castillo, J. Riu, F. Ventura, R. Boleda, R. Scheding, H.F. Schroeder, C. Niston, J. Emneus, P. Eichhorn, T.P. Knepper, C.C.A. Jonkers, P. de Voogt, E. Gonzalez-Mazo, D. Barcelo, J. Chromatogr. A 889 (2000) 195.
[22] M. Petrovic, D. Barcelo, Fresenius J. Anal. Chem. 368 (2000) 676.
[23] L.H. Levine, J.L. Garland, J.V. Johnson, Anal. Chem. 74 (2002) 2064.
[24] O.P. Haefliger, Anal. Chem. 75 (2003) 371.
[25] A. Nakae, K. Kunihiro, J. Chromatogr. 156 (1978) 167.
[26] A. Tegeler, W. Ruess, E. Gmahl, J. Chromatogr. A 715 (1995) 195.
[27] F. Cuesta-Sánchez, B. van den Bogaert, S.C. Rutan, D.L. Massart, Chemom. Intell. Lab. Syst. 34 (1996) 139.
[28] F. Cuesta-Sánchez, J. Toft, B. van den Bogaert, D.L. Massart, Anal. Chem. 68 (1996) 79.
[29] R. Tauler, D. Barceló, Trends Anal. Chem. 12 (1993) 319.
[30] S. Sentellas, J. Saurina, J. Sep. Sci. 26 (2003) 1395.
[31] H. Li, F. Zhang, J. Havel, Electrophoresis 24 (2003) 3107.
[32] V. Bernabé-Zafón, J.R. Torres-Lapasió, S. Ortega-Gadea, E.F. Simó-Alfonso, G. Ramis-Ramos, J. Chromatogr. A 1036 (2004) 205.
[33] T.L. Cecil, S.C. Rutan, J. Chromatogr. 556 (1991) 495.
[34] H. Shen, J. Wang, Y.Z. Liang, K. Petterson, M. Josefson, J. Gottfries, F. Lee, Chemom. Intell. Lab. Syst. 37 (1997) 261.
[35] Y.Z. Liang, O.M. Kvalheim, A. Rahmani, R.G. Brereton, Chemom. Intell. Lab. Syst. 18 (1993) 265.
[36] P.J. Gemperline, J.H. Cho, B. Archer, J. Chemom. 13 (1999) 153.
[37] C. De Boor, Practical Guide to Splines, revised ed., Applied Mathematical Sciences Series, Vol. 27, Springer, New York, USA, 2001, pp. 207–220.
[38] O. Glatter, H. Greschoning, Mikrochim. Acta II (1986) 389.
[39] S.V. Romanenko, E.S. Romanenko, N.A. Kalpakova, J. Anal. Chem. 56 (2001) 51.
[40] T. Rotunno, F. Palmisano, G. Tiravanti, P.G. Zambonin, Chromatographia 29 (1990) 269.
[41] F.B. Erim, X. Xu, J.C. Kraak, J. Chromatogr. A 694 (1995) 471.
[42] E. Primo-Yúfera, Química de los alimentos, Síntesis, Madrid, 1998, p. 169 (chapter 5).
[43] R.W. Kennard, L.A. Stone, Technometrics 11 (1969) 137.