Talanta 121 (2014) 273–280

Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Resolution of co-eluting compounds of *Cannabis Sativa* in comprehensive two-dimensional gas chromatography/mass spectrometry detection with Multivariate Curve Resolution-Alternating Least Squares

Jone Omar [a,*], Maitane Olivares [a], José Manuel Amigo [b], Nestor Etxebarria [a]

[a] Department of Analytical Chemistry, Faculty of Science and Technology, University of the Basque Country, UPV/EHU, PO Box 644, Bilbao 48080, Basque Country Spain

[b] Department of Food Science, Quality and Technology, Faculty of Sciences, University of Copenhagen, Rolighedsvej 30, DK-1958 Frederiksberg C, Denmark

ARTICLE INFO

Article history:
Received 1 August 2013
Received in revised form
17 December 2013
Accepted 22 December 2013
Available online 9 January 2014

Keywords:
Multivariate Curve Resolution-Alternating Least Squares
MCR-ALS
Comprehensive two-dimensional gas chromatography
Cannabis sativa
Cannabinoids
Sesquiterpenes

ABSTRACT

Comprehensive Two Dimensional Gas Chromatography – Mass Spectrometry (GC × GC/qMS) analysis of *Cannabis sativa* extracts shows a high complexity due to the large variety of terpenes and cannabinoids and to the fact that the complete resolution of the peaks is not straightforwardly achieved. In order to support the resolution of the co-eluted peaks in the sesquiterpene and the cannabinoid chromatographic region the combination of Multivariate Curve Resolution and Alternating Least Squares algorithms was satisfactorily applied. As a result, four co-eluting areas were totally resolved in the sesquiterpene region and one in the cannabinoid region in different samples of *Cannabis sativa*. The comparison of the mass spectral profiles obtained for each resolved peak with theoretical mass spectra allowed the identification of some of the co-eluted peaks. Finally, the classification of the studied samples was achieved based on the relative concentrations of the resolved peaks.

1. Introduction

Comprehensive two-dimensional gas chromatography (GC × GC) provides remarkable features in the understanding of the qualitative and quantitative composition of complex samples. According to the literature, GC × GC offers some advantages in comparison to the classical one-dimensional chromatography such as signal enhancement (higher signal to noise ratio), higher separation capacity and structure-retention dimensionality [1,2]. However, it requires specific approaches to analyze experimental data and to extract the chemical information [3].

The combination of high order instrumental data and efficient data treatment methodologies widens significantly the possibilities of chemical fingerprinting of the targeted and non-targeted analysis in sensitive fields such as bioactive compounds [4] or metabolomics [5]. It is precisely in the assessment of natural lead compounds in drug development where those two fields merge and the needs for sample clustering, classification and chemical fingerprinting are especially highlighted [6].

Among the plant extracts, the analysis of cannabinoids is gaining interest not only due to its extended recreational use but also because *Cannabis sativa* is a strong candidate for new drug source [7]. As a matter of fact, in these extracts more than 400 different compounds can be found, which contain more than 100 terpenoids and around 65 cannabinoids [8,9].

GC–MS based methods for the analysis of *Cannabis sativa* extracts are the most extended, either with derivatization (silylation or methylation of carboxylic groups) or directly without any derivatization step (i.e. chemical fingerprints) [9,10]. In this sense, the development of GC × GC methods would improve the separation and identification drawbacks of one dimension chromatographic techniques.

Many works on GC × GC/qMS emphasize that the technique can be used to analyze complex mixtures in unsurpassed detail [2] but it requires fast detectors that generate large data sets [11]. As a consequence, pre-processing steps such as baseline correction, noise reduction or retention time alignment are usually applied to the raw data in order to reduce the variations occurred during the data collection [12]. Once these artifacts have been corrected and/or

* Corresponding author. Tel.: +34 946015551.
E-mail address: jone.omar@ehu.es (J. Omar).

0039-9140/$ - see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.talanta.2013.12.044

reduced the resulting 2D chromatograms are treated to carry out identification and quantification tasks [1], and to do so, both proprietary licenses and open source programs are used following different data treatment approaches [3,13,14].

The use of chemometric techniques has been proved to be efficient to obtain the essential information from chromatographic data. As a consequence, several techniques such as Parallel Factor Analysis- PARAFAC-, PARAFAC2 or multivariate curve resolution – MCR have been successfully applied in diverse research fields due to the ability to handle multidimensional data [15,16]. In this sense, MCR and PARAFAC2 have demonstrated to provide an excellent resolution of the overlapped peaks even in the presence of retention time shifts or irregular baseline drifts [17].

Recently, the combined use of multivariate curve resolution with alternating least squares algorithm (MCR-ALS) has gained attention because it provides a bilinear description of the data under chemical constrains, as reviewed by Ruckebusch et al. [18–20]. These features provide a promising alternative for resolution of overlapping peaks in two dimensions avoiding previous peak alignment steps [21–23]. The MCR-ALS analysis of a single 2D chromatogram can be extended to all the samples gaining the robustness of a cumulative analysis [24].

The aim of this work is to analyze different co-elution regions of the GC × GC/qMS chromatograms from *Cannabis sativa* extracts by means of MCR-ALS. In this particular approach, instead of using a high resolution mass spectrometer (e.g. TOF), we wanted to combine the GC × GC-MS and multivariate approaches and test their capabilities to resolve complex mixtures. In addition to this aim, the identification of the resolved peaks was attempted by means of the recovered mass spectral profile (MCR-ALS) and their retention times. Moreover, the ability of the MCR-ALS approach to differentiate the studied *Cannabis* sativa species was compared to the one offered by commercially available software packages.

2. Materials and methods

2.1. Samples

The *Cannabis sativa* plants analyzed in this work were collected from local gardens. A total of 17 samples of five different kinds of plant buds (*AK*-47, *amnesia*, *somango*, 1024 and *critical*) were analyzed and whenever possible leaves were analyzed as well. The plants were cryogenically milled in a cryogenic grinder 6770 freezer/mill® (SPEX Sample Prep, Metuchen, New Jersey, USA). The pre-treatment took a previous cooling time of 5 minutes, a milling time of 4 min and 1 cycle. The obtained particle size was about few μm so the homogeneity of the analyzed samples was guaranteed. All milled samples were stored in amber vials at −20 °C until analysis.

2.2. Supercritical fluid extraction (SFE) of the plants

An amount of 50 mg of milled plant was homogeneously mixed with 150 mg of diatomaceous earth and accurately placed in a high pressure extraction vessel of 1 mL (EV-1 Jasco). The extraction conditions were previously optimized and are thoroughly described elsewhere [25]. Briefly, the extraction was performed using SC–CO_2 (Carburos Metálicos 99.9995%, Barcelona, Spain) at 100 bar, 35 °C and 1 mL min^{-1} without co-solvent for the extraction of monoterpenes and 20% of EtOH (purity, Lab-Scan, Spain) as co-solvent for the extraction of sesquiterpenes and cannabinoids. Once the extraction was over, the samples were stored in vials at 4 °C until analysis.

2.3. FID/MS analysis

All the extracts obtained by SFE were analyzed by means of GC × GC-Flame Ionization Detector-Mass Spectrometer (GC × GC-FID/MS). A GC7890A gas chromatograph (Agilent Technologies, PA, USA) equipped with a FID and 5975C MS detector and an Agilent G-3486 A capillary flow plate modulator was employed. The control of the second pressure source was handled with a pressure control module. A three-way solenoid, Fluid Automation System Valve, was used for flow switching. The column set for GC × GC-FID/MS analysis consisted of two columns connected by a valve modulator. The first dimension consisted of HP-5 MS capillary column (Agilent Technologies, 30 m × 250 μm i.d. × 0.25 μm film thickness) and the second dimension consisted of DB-17 MS (Agilent Technologies, 5 m × 250 μm i.d. × 0.25 μm film thickness). Two deactivated but not coated fused silica tubes (restrictor) were used in order to split the flow to the detectors: a 0.70 m 0.32 mm i.d. restrictor connected to the FID and a 0.45 m 0.10 mm i.d. connected to the MS.

The subsequent temperature programmed conditions were set as follows: from 60 °C to 102 °C at 4 °C · min^{-1}, from 102 °C to 165 °C at 12 °C · min^{-1} and from 165 °C to 300 °C at 6 °C · min^{-1} (hold 5 min). The GC was equipped with a split/splitless injector (290 °C). The injections were performed in the splitless mode injecting 2 μL of each sample into the GC using a 7683 Agilent autosampler. Modulation periods of 1.42 s, first column flow of 1.23 mL min^{-1} and second column flow of 17.55 mL min^{-1} were used. Hydrogen (> 99.9995%, AD-1020 Hydrogen Generator, Cinel Strumenti Scientifici, Padova, Italy) was employed as carrier gas taking into account all the safety issues necessary in the laboratory. The FID was operated at a data collection frequency of 100 Hz at 300 °C. The MS detector worked in full scan mode from m/z 50 to 350, at the faster electronic mode that assured a sampling rate of 12,500 amu/s, and temperatures of quadrupole and source were 150 °C and 230 °C respectively.

2.4. MCR-ALS analysis

The GCxGC/qMS chromatograms were exported from GC Image (v. 2.0, Zoex Corporation, Houston, USA) as * cdf files into MATLAB 7.0 (MathWorks, Natick, MA, USA, 2010R2). In order to reduce the amount of information and to facilitate the calculations, only the *regions of interest* (RoI) in the chromatograms were chosen and saved as new chromatograms.

The theory behind MCR-ALS has been discussed in previous works [18,26–28]. Herein we will briefly highlight the main features of MCR-ALS. MCR is a model-free method (this means, MCR does not require the pre-assumption of an empirical model for the peaks to be modeled) that focuses on describing the evolution of the experimental multicomponent measurements through their pure component contributions [26]. Each RoI **D** (*I*,*J*, *K*), composed by *I* elution times in the first chromatographic dimension, *J* elution times in the second chromatographic dimension and *K* *m*/*z* intensities in the third dimension, must be unfolded to adapt the three-way structure to a bidimensional matrix **D** (I^*J,*K*) as indicated in Fig. 1.

MCR-ALS looks for a bilinear data decomposition of the experimental matrix **D** (I^*J,*K*) for *N* components (*N* stands for the number of chemical components present in the RoI) using an iterative algorithm based on constrained linear least-squares steps based on Eq. (1).

$$D = CS^T + E \quad (1)$$

where **C** (I^*J, *N*) is the matrix of pure elution profiles, $\mathbf{S^T}$ (*N*, *K*) is the matrix of pure mass spectra and **E** (I^*J, *K*) is the residual matrix reflecting the experimental error unexplained by the resolution

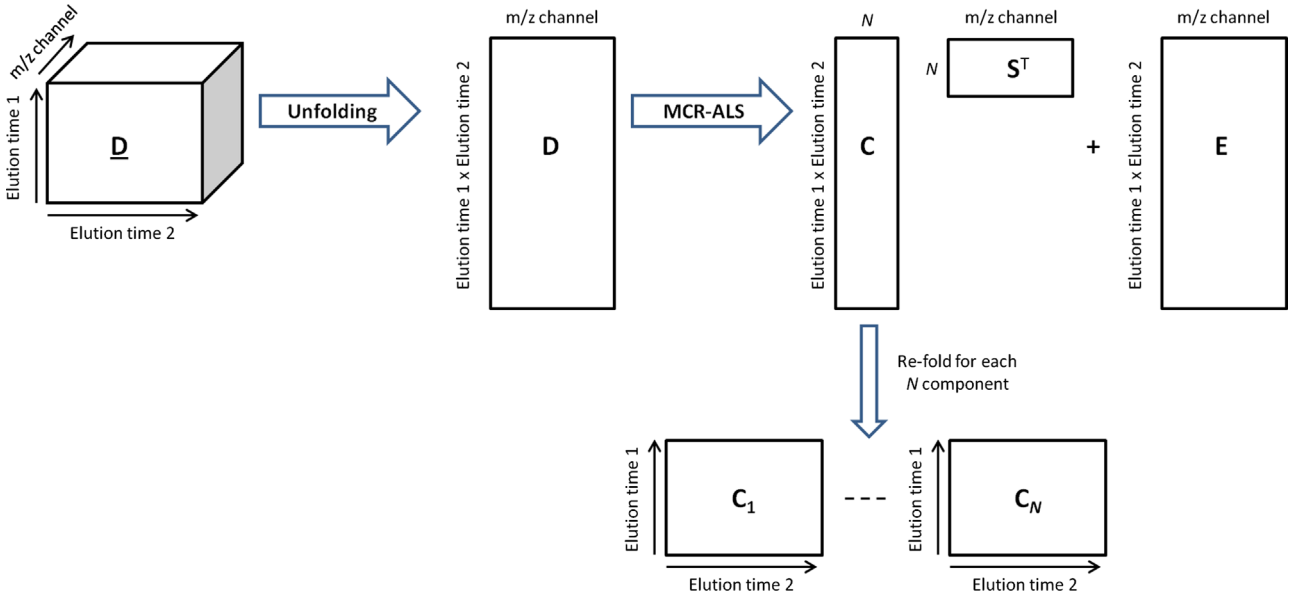

Fig. 1. Graphical representation of the MCR-ALS decomposition for the two dimensional chromatograms of one RoI.

model. After MCR-ALS optimization, the concentration vectors obtained in the pure component matrix are refolded into two-dimensional GC × GC/qMS chromatograms as indicated in Fig. 1 [29].

One of the most popular algorithms for solving this least squares problem is ALS. This algorithm has important benefits in the calculation, being extremely flexible and only considering the bilinear relationship between the concentration profiles and the pure spectra of all the components in the interval. This strength of ALS in the calculations is, at the same time, its weakness, being extremely sensitive to ambiguities. The most common ones are the rotational and intensity ambiguities that can lead the algorithm to obtain a perfectly plausible mathematical solution but totally wrong chemical interpretation. Rotation ambiguities refers to the existence of different linear combinations fitting equally well the original data; whereas intensity ambiguities are originated by the occurrence of scale uncertainty, which describe the original data with the same fit.

In order to decrease the effect of the ambiguities, constraints can be set, limiting the number of possible solutions. In this way, it is possible to tune the contribution of each pure profile, but taking into account that, the application of constraints must be directly linked to the physicochemical nature of every system [30]. A constraint can be defined as any mathematical or chemical property systematically fulfilled by the whole system or by some of its pure contributions, forcing the iterative optimization process to model the profiles respecting the conditions desired, being the most common ones non-negativity, unimodality, equality and closure [30]. In this case, unimodality was employed in the unfolded retention time; whereas nonnegativity constraints were applied in the retention time and mass spectral modes. The application of some of these constraints decreases dramatically the ambiguity of the related profiles and provides fitted parameters of physicochemical and analytical interest [31].

The convergence is achieved when the relative difference in lack of fit (LOF) between two consecutive iterations goes below a threshold value (often 0.1%). The percentage of LOF, estimated as shown in eq. (2), gives a measure of the quality of the model fitting.

$$\%LOF = 100x\sqrt{\frac{\sum_i^J\sum_j^J\sum_k^K e_{ijk}^2}{\sum_i^J\sum_j^J\sum_k^K d_{ijk}^2}} \quad (2)$$

where e_{ijk} is the ijk^{th} element of the unfolded residual matrix **E**, and d_{ijk} is the ijk^{th} element of the unfolded sample **D**.

Another way of decreasing ambiguities is the possibility of applying MCR-ALS to a series of samples that belong to the same elution interval simultaneously. When more than one sample comes into play, the model in eq. 1 can be extended as showed in eq. 3:

$$\begin{bmatrix} D_1 \\ D_2 \\ \cdot \\ \cdot \\ \cdot \\ D_m \end{bmatrix} = \begin{bmatrix} C_1 \\ C_2 \\ \cdot \\ \cdot \\ \cdot \\ C_m \end{bmatrix} S^T + \begin{bmatrix} \begin{bmatrix} E_1 \\ E_2 \\ \cdot \\ \cdot \\ \cdot \\ E_m \end{bmatrix} S^T \end{bmatrix} \quad (3)$$

where ***D*** and ***C*** are column-wise augmented data matrices, with the unfolded submatrices of the individual experiments, D_m and C_m one on top of each other, with a total amount of *m* experiments. The main condition for applying this column-wise augmentation of samples is that the mass-spectra profiles ($\boldsymbol{S}^T$) must be common to all the samples. This condition is, at the same time, a great advantage, since a large variability in the intensities in one sample will help to model that component in samples with small variability of intensities. Therefore, the shape of the concentration profile of that particular component can change in a completely free manner from experiment to experiment [24,32].

In this work, all the MCR-ALS analysis were performed by means of the MCR-ALS Toolbox ("http://www.mcrals.info"; last visited 15th of December 2013) [32] run under MATLAB environment. Further PCA analysis was performed by using the PLS-Toolbox (v. 6.5 Eigenvector Research Inc, WA, USA).

3. Results and discussion

3.1. GC × GC/qMS analysis of Cannabis sativa extracts

As explained in the previous work [25], the extraction of terpenoid and cannabinoid compounds of *Cannabis sativa* was performed in two steps. Firstly, the terpenes (basically monoterpenes) were extracted employing only SC-CO_2, denoted as first SFE fraction henceforth. Once the most volatile fraction was obtained, a mobile phase with 20% EtOH as co-solvent was employed to obtain the second fraction containing the more polar compounds including the cannabinoids (denoted as second SFE fraction from now on).

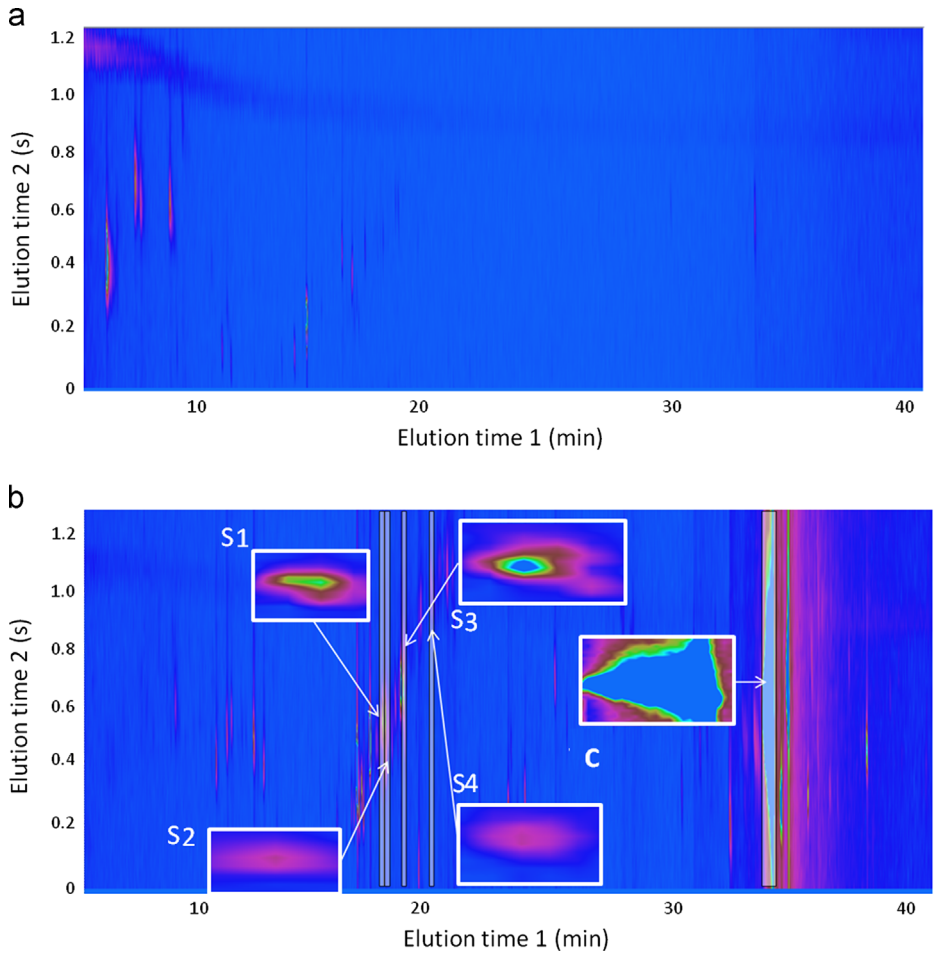

Fig. 2. Total Ion GC × GC/qMS chromatograms of: (a) the first SFE fraction containing monoterpenes and (b) the second SFE fraction containing both sesquiterpenes and cannabinoids. Four co-eluted areas (RoIs) were appreciated for sesquiterpenes (i.e., S1, S2, S3, and S4) and one co-eluted area for cannabinoids (C).

Fig. 2a and b shows the GC × GC/qMS chromatograms for one of the samples of the first and second SFE fractions (only CO_2 and CO_2-EtOH mixture, respectively), showing the extreme difficulty in the separation of complex mixtures. Since a good modulation should render symmetric peaks in both dimensions, the fact of obtaining non-symmetrical peaks suggests co-elution events. In those cases, a careful exam of the blob through the comparison of the mass spectra can confirm this co-elution event.

From the 2D chromatograms obtained when analyzing the second fraction of theSFE extracts, four RoIs were selected among the sesquiterpenes region and one in the cannabinoids region to be carefully examined. Based on the asymmetry of the blob and the different mass spectra present in these regions, they were examined. The analyzed co-eluting areas expressed in the first dimension were the following: sesquiterpene 1 (S1), 19.634–19.745 min; sesquiterpene 2 (S2), 19.765–19.823 min; sesquiterpene 3 (S3), 20.344–20.468 min; sesquiterpene 4 (S4), 22.072–22.214 min and cannabinoid (C), 34.820–35.396 min. The second dimension which is the same as the modulation period (1.42 s) was considered as a whole. Regarding the last RoI, we have to assume a highly co-eluted pattern due to the saturated signal of the tetrahydrocannabinol (Δ^9-THC) as it is the most abundant cannabinoid and all the cannabinoids are eluting close to each other. Moreover, a wraparound effect has to been mentioned since such a high concentration tetrahydrocannabinol (Δ^9-THC) is the most abundant cannabinoid and saturates distorts the output. However, it must be assumed as the less concentrated target compounds (i.e., sesquiterpenes) must be analyzed in the same run.

3.2. MCR-ALS results

The MCR-ALS procedure was applied to the five RoIs in order to reveal the number of compounds co-eluting in each RoI. First of all, the chemical rank (a.k.a. number of sources of variation) in the studied area was individually determined by the standard method of principal component analysis (PCA). Due to the low Signal-to-Noise (S/N) ratio and the high compositional variability of the samples, different number of components were determined for each RoI. Thus, 3 components were found for S1, S3 and C; whereas only 2 components were modeled in S2 and S4 regions.

Once the number of components for each RoI was assessed, it was also necessary to provide for MCR-ALS the initial estimations for starting the iterations. Due to the demonstrated low S/N, the spectral profiles obtained by the standard method SIMPLISMA [33] were used for each RoI. The main aim of SIMPLISMA applied on the spectral dimension is to locate the most dissimilar spectra considering the number of components previously assessed which minimizes the risk of converging to a local optimum [23]. MCR-ALS algorithm was then initialized applying non-negativity constraints to both concentration and mass spectral profiles [34].

Figs. 3 and 4 show some of the results obtained in the S1 and in the *C* regions, respectively. The relative concentrations obtained for each of the components in all the samples are shown; moreover, the reconstructed chromatograms as well as the mass spectral profiles for one of the samples are shown. First of all, it can be observed that the total concentrations found in leaves were ten times smaller than those found in buds, which supposed the

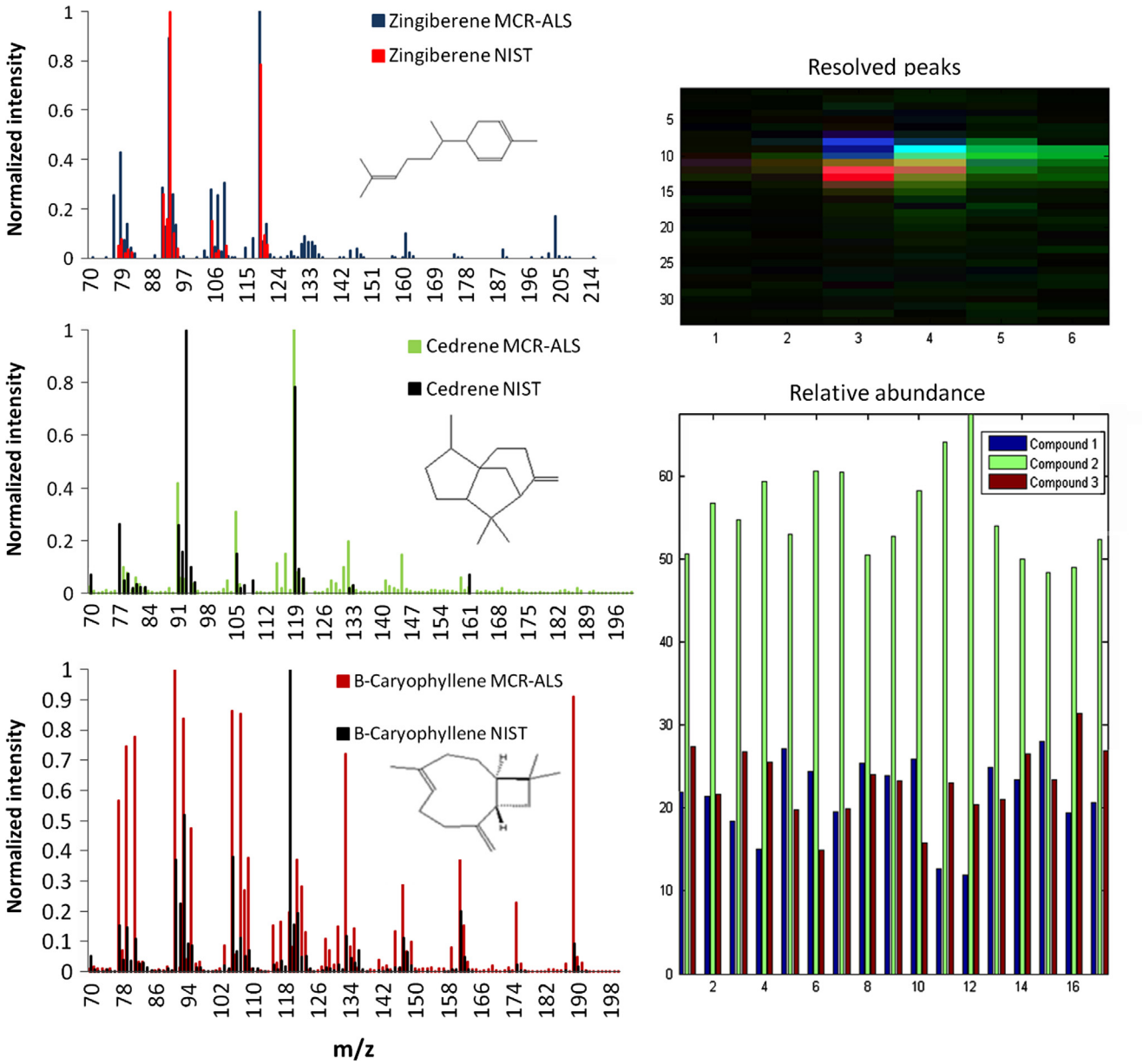

Fig. 3. MCR-ALS results for the co-eluting sesquiterpenes in the first detected region (S1). Top-right denotes a false color image of the resolved elution profiles for one of the samples. Each compound, has been resolved (red, green and blue colors) and both the mass spectra (left column) and the relative abundance of each compound in the RoI for all the samples is shown (right-bottom). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

absence of some compounds in a few leaf extracts and made the resolved profiles very different.

The results of MCR-ALS for all studied RoI as well as the goodness of the performed models are summarized in Table 1: relative abundance of each compound in its RoI, lack of fit (%LOF) of the resolved peaks and information about the identification of the pure compounds.

Once the number of components in each RoI was found out, with a LOF lower than 5% for all the regions, their identification was also considered. To this aim, several plausible target compounds (20 sesquiterpenes and 14 cannabinoids) were selected based on the literature [7]. Their mass spectra (GC-EI-MS) were obtained from the NIST 07 database (NIST'07 Mass Spectral Library, Version 2007, Scientific Instrument Services Inc., 1027, Old York Rd. Ringoes, NJ, USA). The mass spectra of the target compounds were loaded in MATLAB and the MCR-ALS resolved spectral profiles for each compound were compared to the loaded spectra in terms of correlation coefficients (Pearson's correlation coefficient, R) (see Table 1). As an example of this comparison, it can be highlighted the matching found for dronabinol (Δ^9-THC) in the *C* RoI, being 0.95 for the most abundant MCR-ALS spectral profile. The other two compounds were not fully identified since the MCR-ALS mass spectral profiles were highly similar to many cannabinoids. However, it is noteworthy that the highest correlation coefficient was achieved for dronabinol and the smallest lack of fit was also obtained in that region, meaning that the fitted model was in agreement with the raw data collected in *C* RoI. In the S1 RoI, the three compounds were identified as zingiberene, cedrene and β-caryophyllene with matching coefficients higher than 0.80.

3.3. Classification of the samples

Since the analysis of the RoIs revealed a large variability, undoubtedly, attributed to the source variability of the samples, principal component analysis (PCA) was performed by using the area of the resolved chromatographic peaks for all the samples and

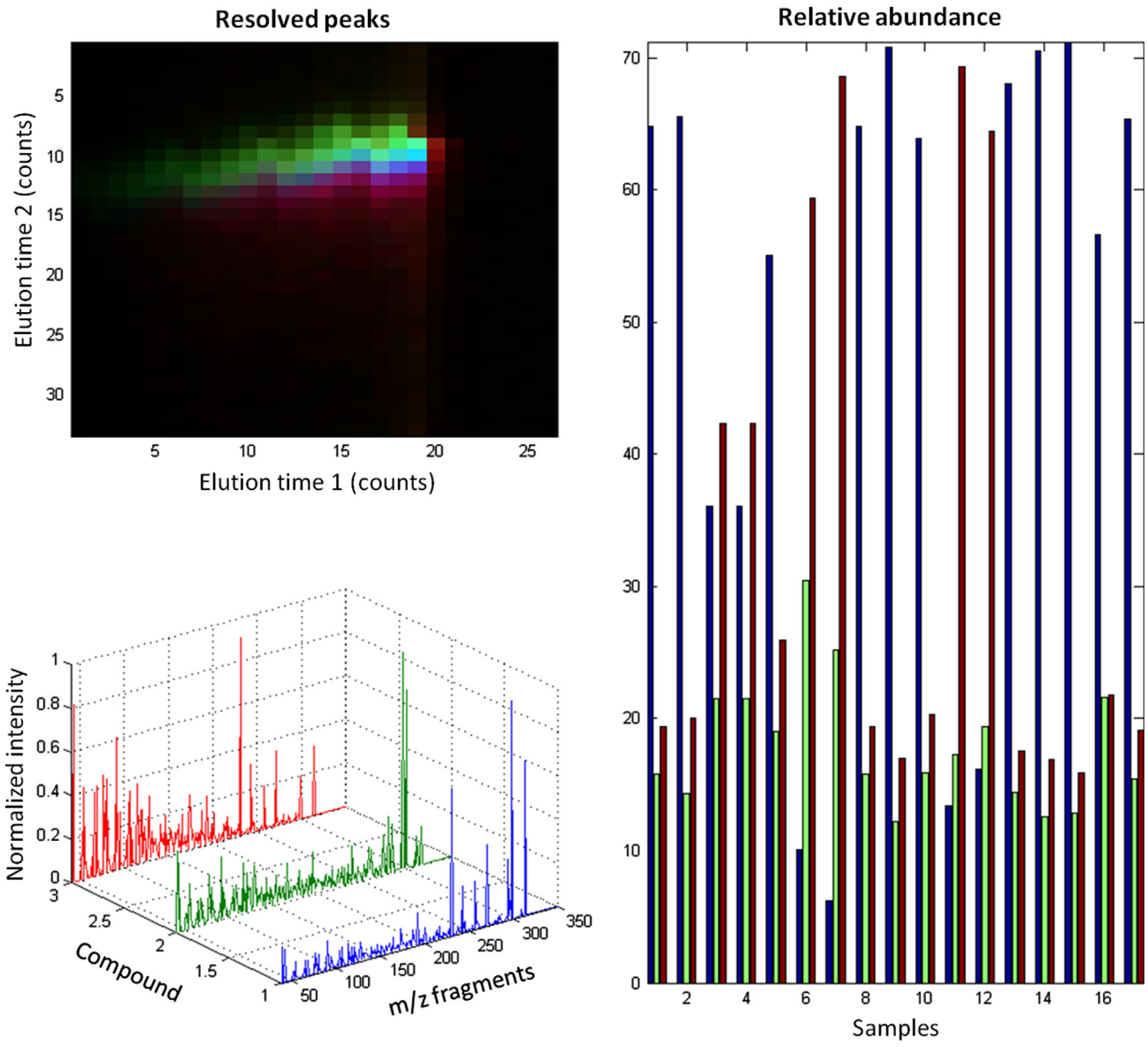

Fig. 4. MCR-ALS results for the co-eluting cannabinoids (C). Top-left denotes a false color image of the resolved elution profiles for one of the samples. Each compound has been resolved (red, green and blue colors) and the corresponding mass spectrum (profile) is shown in the bottom. In the right hand side, the relative abundance of each compound in the RoI for all the samples has been depicted. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

all the RoIs. Therefore, a new data matrix X (13,17) was created containing the relative concentration of the 13 resolved compounds in the 17 samples.

It has to be considered as well that many different marijuana types have been developed mixing both *Cannabis sativa* and *indica* seeds. Therefore, it is not so straightforward to get deep into the plant details. The PCA model of the autoscaled data matrix required 3 principal components to explain up to 99.9% of the total variance (PC1 50%, PC2 32%, PC3 17.9%).

The results of this PCA model are shown in Fig. 5. The first principal component (PC1) is highly related with the total concentration of sesquiterpenes and cannabinoids, it showed the separation of the analyzed plant parts (buds and leaves, results not shown). Nevertheless, the PC2 vs. PC3 scatter plots (Fig. 5a) shows a clear separation not only between buds and leaves but also between the different plants analyzed. As it can be seen in Fig. 5, all the samples (except one) denoted as *critical* are grouped in the positive part of PC3; whereas the rest of the species are placed in the negative part of PC3. The explanation of this grouping can be found in the PC2 vs. PC3 loadings scatter plot, where the samples denoted as *critical* are highly related to the presence of valencene; whereas the rest of species are related to the presence of farnesene and agarospirol, among other non-identified but significant compounds.

The analysis of 2D chromatograms can be carried out by proprietary software such as GC Image and Investigator softwares (v. 2.0. and v. 2.2., Zoex Corporation, Houston, USA). This software also allowed the performance of a peak corresponding matching which enabled direct comparison of analyte peak responses across samples. This kind of template matching can be used to identify both targeted and non-targeted analytes in two-dimensional chromatograms in a direct way [35]. Moreover, template matching involves two peak sets: a peak template and a target peak set. A peak template is a manual set of peaks with metadata that identifies and characterizes the peak (retention time, MS fragmentation pattern, chemical structure, etc.). A target peak set is a set of peaks whose metadata is to be determined. A target peak has only computed statistical features, such as peak location, area,

Table 1
Summary of the number of compounds found in each RoI and their relative abundance obtained by means of MCR-ALS method as well as the LOF for each model. The name, formula, *m/z* values of the identified compounds are also shown as well as the correlation coefficient of the theoretical mass spectra of each compound to that provided by the NIST'07 database.

Coeluting area	*RoI*	*MCR-ALS*			*Pure Compound*			
		Number of Compounds	*Relative Abundance (%)*	*LOF (%)*	*Compound*	*Formula*	*r*	*m/z*
Sesquiterpene	S1	1	21	3.4	Zingiberene	$C_{15}H_{24}$	0.88	119, 69
		2	55		Cedrene	$C_{15}H_{24}$	0.83	119, 91
		3	24		β-Caryophylene	$C_{15}H_{24}$	0.81	91, 105
	S2	1	77	5.2	Aromandrene	$C_{15}H_{24}$	0.87	105, 107
		2	23		Non identified	–	–	91, 78
	S3	1	23	2.2	Farnesene	$C_{15}H_{24}$	0.86	69, 93
		2	32		Valencene	$C_{15}H_{24}$	0.92	161, 107
		3	45		Non identified	–	–	91, 105
	S4	1	95	4.5	Non identified	–	–	91, 93
		2	5		Agarospirol	$C_{15}H_{26}O$	0.82	105, 93
Cannabinoid	C	1	51	0.96	Dronabinol	$C_{21}H_{30}O_2$	0.95	299, 314
		2	17		Non identified	–	–	295, 299
		3	31		Non identified	–	–	231, 91

*R denotes for Pearson's correlation coefficient.

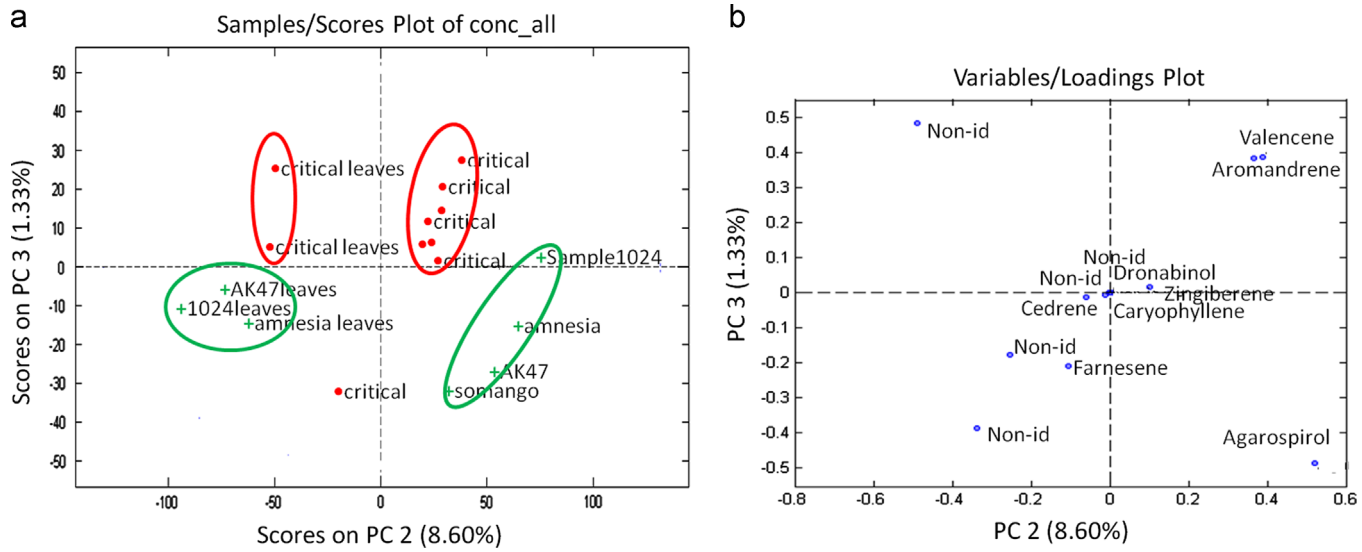

Fig. 5. PC2 vs. PC3 scatter plot for (a) the scores and (b) the loadings.

volume, etc. [13,36]. Once the main information of the peaks of interest is recorded in a template, the chromatograms must be compared one by one with the template; the matching algorithm determines the geometric algorithm that best fits the target peaks with the template pattern [37,38].

In this case, the analysis of the 2 D chromatograms included not only the second SFE fraction but also the first one (pure CO_2) and, therefore, the elution of the volatile compounds (monoterpenes). For this purpose, an initial template including all target analytes, i.e. monoterpenes (alpha-pinene, myrcene, beta-pinene and limonene) and cannabinoids (Cannabigerol (CBG), Cannabivarin (THV), Dronabinol (Δ^9-THC), Cannabinol (CBN) and Cannabichromene (CBC)), was generated and applied to all 2D chromatograms. This template matching across the samples allows alignment of chromatograms for a further comparison [13,37].

As it was performed previously, the results of this analysis were used to study the classification of the samples based on the peak volume of each analyte and analyzed using PCA. The sample data set consisted on 9 target compounds as variables (4 monoterpenes and 5 cannabinoid compounds). In the first approach, a PCA was carried out with the set of 17 samples of the first SFE extracts (rich in monoterpenes) in order to see if there was any difference of the Cannabis sativa samples. Afterwards, the second PCA was carried out with a set of 17 Cannabis sativa extracts obtained with the second SFE extract (i.e. rich in cannabinoids). When analyzing both monoterpenes and cannabinoids, sample grouping within species could be observed. In order to classify the plants considering only the monoterpenes, two big groups were appreciated. All the *critical* species plants rich in myrcene and alpha-pinene were found in one group, and the rest of the species were found in another group rich in limonene. Furthermore, the same two groups were distinguished if only cannabinoids were considered for PCA: one group was constituted by *critical* samples whereas the other group was formed by somango, amnesia, AK-47 and 1024 species richer in CBG.

4. Conclusions

The analysis of the co-eluted peaks in complex samples by GC × GC/qMS was satisfactorily carried out by MCR-ALS, making it feasible for the common goals of GC × GC analysis, i.e. chemical fingerprinting and classification. MCR-ALS of selected RoIs allowed the resolution of the overlapping peaks in terms of relative elution

profile and mass spectra of pure compounds, even though the mass spectra was collected with a *low resolution* detector.

The unsupervised classification patterns found for the studied samples were in good agreement, which reassures the strength of the commercial software and the results of the MCR-ALS approach.

Even if the information that is shared in both approaches (MCR-ALS and GC-Image+Investigator) is quite limited (i.e. THC and the most abundant cannabinoids), only MCR-ALS can provide the resolution and identification of otherwise unknown components.

Acknowledgments

This work has been partially financed by the project DIPE 09/09 418 of the Council of Biscay. J. Omar is grateful to the Basque Government for her PhD. Fellowship. IDOKI SCF is also acknowledged for the use of the SFE facilities. Søren Furbo from the University of Copenhagen is also acknowledged for his help with the data extraction.

References

[1] J.A. Murray, J. Chromatogr. A 1261 (2012) 58–68.
[2] S. Reichenbach, X. Tian, C. Cordero, Q. Tao, J. Chromatogr. A 1226 (2012) 140–148.
[3] K. Pierce, B. Kehimkar, L.C. Marney, J.C. Hoggard, R.E. Synovec, J. Chromatogr. A 1255 (2012) 3–11.
[4] K. Duarte, T. Rocha-Santos, A. Freitas, A.C. Duarte, TrAC, Trends Anal. Chem. 34 (2012) 97–110.
[5] M.M. Koek, R.H. Jellema, J. van der Greef, A.C. Tas, T. Hankemeier, Metabolomics 7 (2011) 307–328.
[6] N.D. Yuliana, A. Khatib, Y.H. Choi, R. Verpoorte, Phytother. Res. 25 (2011) 157–169.
[7] J.T. Fischedick, A. Hazekamp, T. Erkelens, Y.H. Choi, R. Verpoorte, Phytochemistry 71 (2010) 2058–2073.
[8] T.J. Raharjo, R. Verpoorte, Phytochem. Anal. 15 (2004) 79–94.
[9] M.A. ElSohly, W. Gul, M. Salem, Forensic Science, Handbook of Analytical Separations, Elsevier, 2008.
[10] A. Hazekamp, Ph.D. thesis, Proefschrift Universiteit Leiden, 2007.
[11] S. Reichenbach, X. Tian, Q. Tao, E.B. Ledford, Z. Wu, O. Fiehn, Talanta 83 (2011) 1279–1288.
[12] J.T.V. Matos, R.M.B.O. Duarte, A.C. Duarte, J. Chrom. B 910 (2012) 31–45.
[13] S. Reichenbach, M. Ni, V. Kottapalli, A. Visvanathan, Chem. Intell. Lab. Syst. 71 (2004) 107–120.
[14] H.G. Schmarr, J. Bernhardt, J. Chromatogr. A 1217 (2010) 565–574.
[15] J.M. Amigo, T. Skov, R. Bro, Chem. Rev. 110 (2010) 4582–4605.
[16] B. Khakimov, J.M. Amigo, S. Bak, S.B. Engelsen, J. Chromatogr. A 1266 (2012) 84–94.
[17] R. Manne, Chem. Intell. Lab. Syst. 27 (1995) 89.
[18] C. Ruckebusch, L. Blanchet, Anal. Chim. Acta (2014) (in press), (corrected proof).
[19] R. Tauler, D. Barcelo, TrAC, Trends Anal. Chem. 12 (1993) 319–327.
[20] A. de Juan, R. Tauler, Crit. Rev. Anal. Chem. 36 (2006) 163–176.
[21] L.W. Hantao, H.G. Aleme, M.P. Pedroso, G.P. Sabin, R.J. Poppi, F. Augusto, Anal. Chim. Acta 731 (2012) 11–23.
[22] M. Jalali-Heravi, H. Parastar, Talanta 85 (2011) 835–849.
[23] H. Parastar, J.R. Radović, J.M. Bayona, R. Tauler, Anal. Bioanal. 405 (2013) 6235–6249.
[24] J.M. Amigo, A. de Juan, J. Coello, S. Maspoch, Anal. Chim. Acta 567 (2006) 236–244.
[25] J. Omar, M. Olivares, M. Alzaga, N. Etxebarria, J. Sep. Sci. 36 (2013) 1397–1404.
[26] A. de Juan, R. Tauler, Crit. Rev. Anal. Chem. 36 (2006) 163–176.
[27] A. de Juan, R. Tauler, J. Chem. 15 (2001) 749–772.
[28] I. van Stokkum, K. Mullen, V. Mihaleva, Chemom. Intell. Lab. Syst. 95 (2009) 150–163.
[29] L.A. Fonseca de Godoy, L.W. Hantao, M.P. Pedroso, R.J. Poppi, F. Augusto, Anal. Chim. Acta 699 (2011) 120–125.
[30] I. Montoliu, R. Tauler, M. Padilla, A. Pardo, S. Marco, Sens. Actuators B 145 (2010) 464–473.
[31] A. de Juan, R. Tauler, Anal. Chim. Acta 500 (2003) 195–210.
[32] J.M. Amigo, A. de Juan, J. Coello, S. Maspoch, Anal. Chim. Acta 567 (2006) 245–254.
[33] N. Dupuy, Y. Batonneau, Anal. Chim. Acta 495 (2003) 205–215.
[34] H.P. Bailey, S.C. Rutan, Chemom. Intell. Lab. Syst. 106 (2011) 131–141.
[35] J. Omar, I. Alonso, M. Olivares, A. Vallejo, N. Etxebarria, Talanta 88 (2012) 145–151.
[36] S. Reichenbach, P. Carr, D. Stoll, Q. Tao, J. Chromatogr. A 1216 (2009) 3458–3466.
[37] C. Cordero, E. Liberto, C. Bicchi, P. Rubiolo, S. Reichenbach, X. Tian, Q. Tao, J. Chromatogr. Sci. 48 (2010) 251–261.
[38] B. Bäckström, M.D. Cole, M.J. Carrott, D.C. Jones, G. Davidson, K. Coleman, Sci. Justice 37 (1997) 91–97.