Research Article

Liquid chromatography with diode array detection combined with spectral deconvolution for the analysis of some diterpene esters in Arabica coffee brew

In this manuscript, the separation of kahweol and cafestol esters from Arabica coffee brews was investigated using liquid chromatography with a diode array detector. When detected in conjunction, cafestol, and kahweol esters were eluted together, but, after optimization, the kahweol esters could be selectively detected by setting the wavelength at 290 nm to allow their quantification. Such an approach was not possible for the cafestol esters, and spectral deconvolution was used to obtain deconvoluted chromatograms. In each of those chromatograms, the four esters were baseline separated allowing for the quantification of the eight targeted compounds. Because kahweol esters could be quantified either using the chromatogram obtained by setting the wavelength at 290 nm or using the deconvoluted chromatogram, those compounds were used to compare the analytical performances. Slightly better limits of detection were obtained using the deconvoluted chromatogram. Identical concentrations were found in a real sample with both approaches. The peak areas in the deconvoluted chromatograms were repeatable (intraday repeatability of 0.8%, interday repeatability of 1.0%). This work demonstrates the accuracy of spectral deconvolution when using liquid chromatography to mathematically separate coeluting compounds using the full spectra recorded by a diode array detector.

Keywords: Cafestol esters / Diode array detection / Kahweol esters / Spectral deconvolution

DOI 10.1002/jssc.201401095

Additional supporting information may be found in the online version of this article at the publisher’s web-site

1 Introduction

In separation science, a hyphenated technique is the combination of a separation (LC, GC, CE . . . ) and a spectrometric technique (diode array detection (DAD), MS, FTIR . . . ) that is used as the detector [1]. The recorded chromatographic data are often referred to as first-order data [2] and they are (or can be arranged into) a 2D table with the time axis in one dimension, the spectrum axis in the other dimension and the amplitude of the signal as responses. The simplest way to work with such a dataset is by generating mono-dimensional traces (chromatogram, electropherogram, MS spectra, UV spectra . . . ) that are the responses as a function of times at a particular spectral coordinate (chromatogram) or the responses as a function of the spectral coordinates at a given time (spectra). This allows obtaining spectra at different times for identification purposes as well as obtaining chromatograms in which the peak areas can be modulated by selecting the spectral coordinate [3]. Hyphenated MS techniques are probably the best combination for such applications [4–12], and it is often possible with those instruments to choose a spectral coordinate at which one or more compounds of interest will be present in the chromatogram but where potentially interfering species will be transparent. However, MS instruments remain costly to buy and run. DAD, due to its low price and good precision, is a commonly hyphenated detector but its low spectral selectivity does not often allow the resolution of coelution problems as easily as with a MS detector. The application of chemometric approaches in conjunction with the DAD dataset allows the mathematical separation of coeluted species [2, 13–19], where mathematical separation refers to the deconvolution of the dataset to a series of chromatograms by mathematical means.

Spectral deconvolution (SD) is a technique that can be applied to any spectrometric dataset [20, 21]. The key
assumption in this approach is that at any time, the spectrum that is recorded by the detector is a linear combination of the spectra of each compound that are present in the solution. This assumption is true as long as the detector is used within its linear range. This can be expressed mathematically as \[ Y(t) = X\beta(t) + \epsilon(t) \] (1)

where \( Y(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{pmatrix} \) is a vector of length \( n \) that is the measured UV absorption spectrum obtained at time \( t \), \( X = \begin{pmatrix} 1 & x_{21} & \cdots & x_{2p} \\ 1 & x_{31} & \cdots & x_{3p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{np} \end{pmatrix} \) is the \( n \times p + 1 \) design matrix where each \( p \) column corresponds to a spectrum of one of the compound that may be presented in the solution \( \beta(t) = \begin{pmatrix} \alpha(t) \\ \beta_1(t) \\ \vdots \\ \beta_p(t) \end{pmatrix} \) is the slope vector of length \( p + 1 \) and \( \epsilon(t) \) is the error vector. In this equation \( Y \) and \( X \) are known and the goal is to estimate \( \beta(t) \) as each slope parameters in this vector correspond to the relative contribution of the corresponding spectrum in \( X \) to the measured \( Y \) spectrum. The slope parameters are directly related to the concentration of compounds in the solution. This equation is a multilinear regression that can be solved using either matrix algebra or statistical tools such as principal component analysis to estimate the design matrix. Numerous reviews in this subject are available [2, 13, 15, 23–31].

The goal of this work was the quantification of diterpene esters in Arabica coffee brew by HPLC–DAD. Arabica coffee is regarded as an important source of diterpenes mainly cafestol and kahweol. These compounds are of interest due to their potential beneficial effect on human health, for example, some literature point out the anti-carcinogenic, anti-inflammatory as well as anti-angiogenic properties of diterpenes [32, 33]. In coffee, diterpenes are rarely present in free form but mostly exist as esters of fatty acids mainly: palmitic, linoleic, oleic, and stearic acids [34]. For their quantification by HPLC, cafestol and kahweol esters are normally converted to their corresponding diterpene alcohols by a saponification reaction before being analyzed [35]. However, to study the health effect of those compounds, the quantification of the esterified form may provide a better understanding. Kurzrock and Speer used GPC with LC–DAD–MS to identify different kahweol and cafestol fatty acid esters in Arabica coffee [36], but some compounds that were not baseline separated could only be analyzed by MS. The authors also demonstrated that the kahweol esters could be selectively detected with the wavelength set at 290 nm, but not the cafestol esters.

In this work, in absence of a MS detector, a mathematical separation using spectral deconvolution will be used in conjunction with LC–DAD (LC–DAD/SD) to obtain deconvoluted chromatograms that are specific to the cafestol and kahweol esters. Because kahweol esters can be quantified using the chromatogram obtained at 290 nm or using deconvoluted chromatograms, those compounds will be used to compare the analytical performances of both methods.

2 Materials and methods

2.1 Chemicals and samples

Cafestol linoleate, oleate, and stearate along with kahweol linoleate, oleate, and stearate, were obtained from LKT lab (MN, USA). Individual standards of cafestol palmitate and kahweol palmitate were acquired from Sigma–Aldrich (MO, USA). Acetonitrile, methanol (HPLC gradient grade) and diethyl ether were obtained from VWR (Belgium). Sodium chloride was supplied by Panreac Quimica (Spain). All stock solutions were prepared in acetonitrile with the following

\[
SSR = \sqrt{\sum_t (y_t - \hat{y}_t)^2}
\] (2)

where \( y_t \) and \( \hat{y}_t \) are the true and the estimated values, respectively. However, for the solutions to be unique and accurate the following conditions should be verified:

(i) the spectrum of every compound that will contribute to the recorded spectrum, \( Y \), should be taken into account in \( X \).

(ii) Each spectrum should be independent (e.g. no multicollinearity where one spectrum in \( X \) can be expressed as a linear combination of other spectra).

(iii) The number of column \( p \) (number of spectrum used in \( X \)) in the design matrix should be less than the amount \( n \) of observation to avoid overfitting.

Although SD is sometimes used with UV-vis spectroscopy [20], to the best of our knowledge it has never been applied to chromatographic data where the aim is the quantification of few compounds. However, this should be advantageous because at every time few components should contribute to the recorded spectrum. Moreover, the
concentration: cafestol and kahweol palmitate (300 mg/L), cafestol and kahweol linoleate, oleate, stearate (200 mg/L). To protect standards from degradation under sunlight or heat, all of them were wrapped in aluminum foil and stored at –22°C.

2.2 Separation and detection

Boiled coffee was prepared by boiling ground coffee (11.25 g of pure Arabica) with 150 mL of distilled water for 10 min followed by 2 min of settling time followed by decanting the liquid. Three cups (150 mL) were prepared and kept at –22°C. Before the extraction, frozen samples were defrosted and mixed, heated (55–60°C) and stirred well to reach a homogeneous mixture. Then, 2.5 mL of coffee brew was transferred to polyethylene tubes and left to cool for some minutes and subjected to LLE. Extraction of diterpene esters was performed in duplicate according to the developed and validated methodology described previously by Moeenfard et al. [37]. Briefly, coffee brews were extracted directly using 5 mL of diethyl ether. The mixture was vortexed for 2 min and after centrifugation (Rotofox 32A, Germany) at 4000 rpm during 10 min, the upper phase was transferred to a clean test tube. The aqueous solution was re-extracted using diethyl ether (5 mL) then the combined ether phase was washed with 5 mL of 2 M NaCl solution to remove interfering compounds followed by centrifugation (4000 rpm, 10 min). In each step of extraction and cleaning, 0.5 mL of methanol was added to break the emulsion and create a neat interface between the aqueous and organic phases. The clean ether phase was transferred to an amber glass vial and dried under a N₂ stream. Samples were kept at –22°C until analysis using LC–DAD.

2.3 Chromatographic analysis

Analysis was carried out in a Merck Hitachi Elite LaChrom (Tokyo, Japan) system equipped with a quaternary pump (L-2130), an L-2200 autosampler. Separation was achieved using a Purospher STAR LichroCART RP 18 end-capped (250 × 4 mm, 5 μm) column attached to a guard column (4 × 4 mm, 5 μm) of the same kind. Before injection, dried extract was dissolved in 2.5 mL of acetonitrile and filtered through 0.45 μm filter membrane (PTFE, VWR, USA). The chromatographic conditions for analyzing diterpene esters were adapted from Ref. [36] with slight modifications. Twenty microliter of sample was injected and the separation was achieved using isocratic conditions during 35 min with the mobile phase made of acetonitrile/isopropanol (70:30, v/v) and pumped at 0.4 mL/min. The detection was made using an L-2455 (Merck Hitachi) UV/vis spectrophotometry diode array detector in the range of 200–400 nm. Two detection wavelengths were also set: 225 and 290 nm for cafestol esters and kahweol esters, respectively. EZChrom Elite 3.1.6 software was used for data acquisition and analysis.

2.4 Deconvolution

After each run, data were exported as comma-separated values format by the acquisition software (EZChrom Elite 3.1.6). Those files were open using Matlab R2013b and the chromatographic data loaded into a 2D array. The different Matlab functions were programmed using Matlab 2013b and run on a personal computer equipped with 4 GB of RAM. The code and procedures used are further described in the Supporting Information.

3 Results and discussion

3.1 Separation and selection of the model

Different mobile phase compositions and flow rates were assayed using HPLC–DAD to separate the target compounds, namely: acetonitrile/isopropanol (70:30, v/v) at 0.2, 0.4, and 0.8 mL/min; acetonitrile/isopropanol (90:10, v/v) at 0.4, 0.6, and 0.8 mL/min; acetonitrile/isopropanol (50:50, v/v) at 0.4 and 0.6 mL/min; acetonitrile/methanol (70:30, v/v) at 0.8 mL/min; acetonitrile/methanol (90:10, v/v) at 0.8 mL/min. Results were not better than those previously reported by Kurzrock [34]. Within our experimental trials, the best separation was obtained using acetonitrile/isopropanol (70:30, v/v) with a flow rate of 0.4 mL/min. Chromatograms of a sample of coffee, extracted, and separated as described in Section 2 are presented in Fig. 1, with the wavelength set at (A) 205 ± 4 nm, (B) 225 ± 4 nm, and (C) 290 ± 4 nm. The peaks corresponding to the eight diterpene esters of interest are indicated by numbers (see Table 1). While all kahweol esters are baseline separated and can be quantified setting the wavelength to 290 nm (Fig. 1C), this is not the case for the cafestol esters. For those compounds there is no wavelength that is specific enough to resolve the coelution between the cafestol and kahweol esters. As the separation of those compounds is particularly difficult in absence of a MS detector, a mathematical separation may be the easiest approach. Because the kahweol esters can be quantified using a classical method, they were used to compare the analytical performances of both approaches.

The first step in the spectral deconvolution is to build the design matrix, X (Eq. (1)), that will contain all the spectra of the compounds of interest which in this case include: the spectrum of eight diterpene esters, a background and a constant spectrum as well as the spectra of potential interfering compounds. To obtain reliable spectra a series of samples, each of them containing one diterpene ester at concentration of 100 mg/L, were injected and separated. The spectrum for each diterpene ester was measured using the peak of the main compound. Subsequently, the spectrum was corrected for background adsorption (see Supporting Information for more details). The resulting spectra are presented in Fig. 2, with (A) the four spectra related to the kahweol esters, (B) the four spectra related to the cafestol esters, (C) a background spectrum and (D) the spectra of potential interfering species.
Figure 1. Separation of diterpene esters in coffee by LC–DAD with detection at (A) 205 ± 4 nm, (B) 225 ± 4 nm and (C) 290 ± 4 nm. The coffee sample (boiled coffee produced with 100% Arabica coffee) has been extracted with diethyl ether and separated using a C<sub>18</sub> column with a mobile phase constituted of acetonitrile/isopropanol (70:30, v/v) at a flow rate of 0.4 mL/min. The peaks corresponding to the diterpene esters are indicated as (1) (KL: kahweol linoleate), (2) (KO: kahweol oleate), (3) (KP: kahweol palmitate), (4) (KS: kahweol stearate), (5) (CL: cafestol linoleate), (6) (CO: cafestol oleate), (7) (CP: cafestol palmitate), and (8) (CS: cafestol stearate).

Table 1. Names and structures of the diterpenes esters of interest

<table>
<thead>
<tr>
<th></th>
<th>Cafestol esters</th>
<th>Kahweol esters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>with side chain R as:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C(CH&lt;sub&gt;2&lt;/sub&gt;)&lt;sub&gt;2&lt;/sub&gt;CH&lt;sub&gt;3&lt;/sub&gt;</td>
<td>C(CH&lt;sub&gt;2&lt;/sub&gt;)&lt;sub&gt;2&lt;/sub&gt;CH&lt;sub&gt;2&lt;/sub&gt;</td>
</tr>
<tr>
<td></td>
<td>(CH&lt;sub&gt;3&lt;/sub&gt;)&lt;sub&gt;3&lt;/sub&gt;CH&lt;sub&gt;3&lt;/sub&gt;</td>
<td>(CH&lt;sub&gt;3&lt;/sub&gt;)&lt;sub&gt;2&lt;/sub&gt;CH&lt;sub&gt;3&lt;/sub&gt;</td>
</tr>
<tr>
<td></td>
<td>Cafestol – linoleate (CL)</td>
<td>Kahweol – linoleate (KL)</td>
</tr>
<tr>
<td></td>
<td>Cafestol – oleate (CO)</td>
<td>Kahweol – oleate (KO)</td>
</tr>
<tr>
<td></td>
<td>Cafestol – palmitate (CP)</td>
<td>Kahweol – palmitate (KP)</td>
</tr>
<tr>
<td></td>
<td>Cafestol – stearate (CS)</td>
<td>Kahweol – stearate (KS)</td>
</tr>
</tbody>
</table>

All spectra have been normalized by their highest absorbance value to facilitate their visual comparison. As it can be clearly seen, the spectra from the same diterpene only differ by their absorbance values below 220 nm. The matrix of correlation between the different spectra was calculated using Matlab.

All the kahweol esters were found to be correlated with a score higher than 0.93 and all the cafestol esters with a score higher than 0.91. Because the background also absorbs below 220 nm, the data below 220 nm and over 320 nm were removed (1) to avoid problem of colinearity with the background.
Figure 2. UV-vis absorption spectra of (A) cafestol esters, (B) kahweol esters, (C) background, and (D) main impurities.

Figure 3. Deconvoluted chromatograms obtained using LC–DAD/SD with of a standard mixture of diterpene esters. Panels with the suffix I display chromatograms obtained with a four-spectra model and Panels with the suffix II display chromatograms obtained with a seven-spectra model. Panels with the suffix A, B, and C display the deconvoluted chromatogram related to kahweol, the deconvoluted chromatogram related to cafestol, and the sum squared residuals, respectively. Other conditions are as stated in Fig. 1.

spectra and (2) to remove the range where the diterpene esters do not absorbed. Within this new wavelength range, both spectra of esters from the same diterpene correlated with a score higher than 0.998, thus the same spectra were used for the four cafestol esters and the four kahweol esters. Two deconvolution models will be used subsequently as follows: (I) the four-spectra model (one spectrum for all the cafestol esters, one spectrum for all kahweol esters, one background spectrum, and one constant in the design matrix) and (II) the seven-spectra model that include the four spectra of model (I) plus another three spectra from impurities that may coelute with the diterpene esters and had been detected in the standard solutions. Those three additional spectra are shown in Fig. 2D.

3.2 Spectral deconvolution

Using Matlab, SD was applied to every spectrum acquired during the separation (see Supporting Information). Deconvoluted chromatograms were obtained as the variation as a function of time of one of the slope parameter obtained...
### 3.3 Analytical performances

The analytical performances obtained when quantifying the kahweol esters with a classical approach with detection at 290 nm or after SD with a four-spectra model and a seven-spectra model were compared to validate the HPLC–DAD/SD methodology. In the deconvolution model, the same spectra were always used. The exception is the background spectra which were measured in each experiment, individually.

The calibration curves were build using standard solutions, each of them containing the eight diterpene esters at known concentration ranging from 2 to 600 mg/L for cafestol palmitate (CP) and kahweol palmitate (KP) (2, 5, 10, 20, 50, 150, 300, and 600 mg/L) and between 2 and 600 mg/L (2, 5, 10, 20, 50, 100, 200, and 600 mg/L) for cafestol oleate (CO), cafestol linoleate (CL), cafestol stearate (CS), kahweol oleate (KO), kahweol linoleate (KL), and KS. In the classical approach the peak areas were measured using the acquisition software. With the deconvoluted chromatograms peak areas were measured using a small routine programmed in Matlab (see Supporting Information). Each standard solution was run in duplicate. Results obtained are presented in Table 2.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Spectral deconvolution</th>
<th>Calibration curve</th>
<th>LOD (mg/L)</th>
<th>r²</th>
<th>Intercept</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Classical</td>
<td>Four-spectra model</td>
<td>Seven-spectra model</td>
<td>290 ± 4 nm</td>
<td>KO</td>
</tr>
<tr>
<td></td>
<td>0.87 ± 0.09</td>
<td>0.94 ± 0.01</td>
<td>0.99 ± 0.01</td>
<td>207 ± 30</td>
<td>227 ± 19</td>
</tr>
<tr>
<td></td>
<td>± 0.08</td>
<td>± 0.01</td>
<td>± 0.02</td>
<td>± 30</td>
<td>± 19</td>
</tr>
</tbody>
</table>

a) Values calculated by excel using the LINEST function.

b) The LODs calculated as the concentration at which the amplitude of the response is equal to the intercept plus three times σy/x.

c) The value in bracket are the SDs.
four-spectra model, clearly demonstrating that the seven-spectra model gives better results.

The intra-day repeatability was measured using the seven-spectra model with six successive runs using a mixture of the eight diterpene esters, each at a concentration of 75 mg/L. The average intraday peak area RSD was 0.8% (min CL, 0.3%; max KO, 1.7%). The interday repeatability was measured on three successive days, each sample run in triplicate. The average interday peak area RSD was measured to be 1.0% (min KL and KS, 0.5%; max CS, 1.4%).

### 3.4 Application to a real coffee sample

After extraction as detailed in Section 2, a coffee brew (boiled coffee prepared using 100% Arabica coffee) was separated, and the concentration of the eight diterpene esters was measured. Three extractions were performed, and each extraction was subjected to duplicate injections. A deconvoluted chromatogram can be seen in Fig. 4 with (A) the deconvoluted chromatogram related to cafestol, (B) the deconvoluted chromatogram related to kahweol and (C) the plot of the relative SSR. Results obtained using the chromatogram at 290 nm (only for KL, KO, KP, KS) or using the deconvoluted chromatograms are summarized in Table 3. In this table, the t-test (two-tail) between the two series was calculated using Excel. A value of $p$ higher than 0.05 indicates that the values are not significantly different [38] and, for all calculated concentration of kahweol esters perfect agreement between the two approaches was observed. Cafestol could only be calculated after spectral deconvolution. It should be noted that the SSR values shown in Fig. 4C are, at least, one order of magnitude higher than when working with the mixture of standard (Fig. 3IC and IIC). This is particularly noticeable for KO, KP, CO, and CP. This is probably due to the presence of coeluting species whose spectra are not in the model. This is not surprising when working with complex matrix and real samples. A better precision could probably be obtained if the spectra of those impurities were taken into account in the design matrix. However, in this particular example the error is small and should not significantly contribute to the measured concentration in Table 3. This is validated by the excellent agreement obtained using both approaches as demonstrated in Table 3. It should be emphasized that such problems also occur with a classical chromatogram. However it is generally unnoticed. Here the SSR allows us to verify the goodness of the model and to detect the presence of coeluting species even with perfect coelution, as long as the spectra of the impurity is significantly different from the one of the

<table>
<thead>
<tr>
<th></th>
<th>Cafestol</th>
<th>Kawheol</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CL</td>
<td>CO</td>
</tr>
<tr>
<td>Classical</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>(290 ±4 nm)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deconvolution</td>
<td>57.9 (±1.1)</td>
<td>37.2 (±1.9)</td>
</tr>
<tr>
<td>(seven-spectra model)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$P(T ≤ t)^{b)}$</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

a) t-test for two samples calculated by Excel. $P(T ≤ t) > 0.05$ indicates that the values are not significantly different.
b) Values in bracket are the SDs.
main components. It should also be noted in Fig. 4C the very high SSR of the peaks marked by an asterisk. Those peaks are due to unknown components that are not taken into account in the model. It is evident that those peaks will not impact the quantification of the target compounds and can be ignored.

4 Concluding remarks

LC–DAD/SD has been successfully applied to deconvolute, from the raw data, the contribution from cafestol and kahweol esters. Regarding kahweol esters, concentration could be quantified using the deconvoluted chromatogram or through the chromatogram obtained at 290 nm. Nevertheless slightly better analytical performances were obtained using the deconvoluted chromatogram. LC–DAD/SD demonstrated its full potential with the analysis of cafestol esters that could not have been achieved otherwise. Moreover, this approach was demonstrated to be accurate and cheap.

However, LC–DAD/SD is not a universal solution. While mathematical deconvolution could be an integrated part of the analytical method when using DAD, it should also be rigorously designed and validated. In particular, it is important to obtain high quality spectra from standards and to verify for linearity. The amount of spectra used at any time should also be kept minimal. While, theoretically, only the spectra of the target compounds and background spectra are needed, this is with the assumption that no impurities are coeluting. If this is not the case and if the concentrations of the impurities are high enough to interfere, the spectra of the impurities should be added in the design matrix. A careful examination of the plot of the SSR allows us visually to assess the performance of the deconvolution model and to optimize it if necessary.

This work was funded by FEDER funds through the Operational Program for Competitiveness Factors—COMPETE, ON.2—O Novo Norte—North Portugal Regional Operational Program and National Funds through FCT-Foundation for Science and Technology under the projects: PEst-C/EBB/UI0578/2013, NORTE-07-0124-FEDER-000255-RL2Environment & Health, and SFRH/BD/7931/2011.

This work was funded by FEDER funds through the Operational Program for Human Potential and by National Funds through FCT-Foundation for Science and Technology under the project IF/00528/2013.

The authors have declared no conflict of interest.

5 References


