



# Application of Fourier transform infrared spectroscopy and orthogonal projections to latent structures/partial least squares regression for estimation of procyanidins average degree of polymerisation

Cláudia P. Passos<sup>a</sup>, Susana M. Cardoso<sup>b,c</sup>, António S. Barros<sup>b</sup>, Carlos M. Silva<sup>a</sup>, Manuel A. Coimbra<sup>b,\*</sup>

<sup>a</sup> CICECO, Departamento de Química, Universidade de Aveiro, 3810-193 Aveiro, Portugal

<sup>b</sup> Departamento de Química, Universidade de Aveiro, 3810-193 Aveiro, Portugal

<sup>c</sup> CIMO/Escola Superior Agrária, Instituto Politécnico de Bragança, Campus de Santa Apolónia, 5301-854 Bragança, Portugal

## ARTICLE INFO

### Article history:

Received 2 June 2009

Received in revised form

17 December 2009

Accepted 18 December 2009

Available online 29 December 2009

### Keywords:

Fourier transform infrared spectroscopy  
Orthogonal projections to latent structures  
Partial least squares regression-1  
Grape seed  
Flavan-3-ols  
Graded procyanidin fractionation

## ABSTRACT

Fourier transform infrared (FTIR) spectroscopy has been emphasised as a widespread technique in the quick assess of food components. In this work, procyanidins were extracted with methanol and acetone/water from the seeds of white and red grape varieties. A fractionation by graded methanol/chloroform precipitations allowed to obtain 26 samples that were characterised using thiolysis as pre-treatment followed by HPLC-UV and MS detection. The average degree of polymerisation (*DP<sub>n</sub>*) of the procyanidins in the samples ranged from 2 to 11 flavan-3-ol residues. FTIR spectroscopy within the wavenumbers region of 1800–700 cm<sup>-1</sup> allowed to build a partial least squares (PLS1) regression model with 8 latent variables (LVs) for the estimation of the *DP<sub>n</sub>*, giving a RMSECV of 11.7%, with a *R*<sup>2</sup> of 0.91 and a RMSEP of 2.58. The application of orthogonal projection to latent structures (O-PLS1) clarifies the interpretation of the regression model vectors. Moreover, the O-PLS procedure has removed 88% of non-correlated variations with the *DP<sub>n</sub>*, allowing to relate the increase of the absorbance peaks at 1203 and 1099 cm<sup>-1</sup> with the increase of the *DP<sub>n</sub>* due to the higher proportion of substitutions in the aromatic ring of the polymerised procyanidin molecules.

© 2009 Elsevier B.V. All rights reserved.

## 1. Introduction

Grape seeds are a rich source of flavan-3-ols. These structures are present in mixtures of monomers together with oligomeric and polymeric procyanidins, mainly composed by residues of (–)-epicatechin, (+)-catechin, and (–)-epicatechin-*O*-gallate (Fig. 1) [1,2]. In grape seed tissues, although the quantification of procyanidins has only been accurately performed for oligomers up until tetramers, they have been shown to represent more than 70% of the total flavan-3-ols [3]. Procyanidins with different degrees of polymerisation have different properties, namely, they confer different organoleptic properties to the food [4,5] and they have different absorption behaviour at the gastrointestinal tract level [6]. The estimation of the average degree of polymerisation (*DP<sub>n</sub>*) of procyanidins is thus a useful parameter to evaluate the type of procyanidins present in a sample. Because the majority of procyanidin-rich samples contain molecules with different degrees of polymerisation [7], fractionation steps can be used to improve their homogeneity.

For the calculation of the *DP<sub>n</sub>*, the polymers need to be submitted to an acid-catalysed degradation in the presence of nucleophilic agents, such as thiolysis that promote the formation of distinct monomers corresponding to the terminal and extension units of the polymer [1,8]. These compounds are then usually characterised by reversed-phase HPLC with UV detection at 280 nm [9–12]. More recently, advanced techniques based on mass spectrometry analysis have become very effective in qualitative analysis. These highly sensitive methods are able to characterise procyanidins in complex matrices, allowing the direct identification of the molecules with different degrees of polymerisation [13–16]. Anyway, the analyses using HPLC-UV are both time consuming and expensive and the advanced techniques such as HPLC-MS require highly specialised equipments. Therefore, versatile, cheap, and rapid analyses are desirable and of great interest for its use in routine and generalised analyses.

Infrared spectroscopy has been a very useful tool for rapid evaluation of the procyanidin composition of a given sample [17,18]. Nevertheless, the signals obtained for potential information are complex. Chemometric methods allow the extraction of useful information from the large amount and complex data generated. Among the vast field of chemometrics, the multivariate regression methods have been widely used to provide a better insight into the

\* Corresponding author. Tel.: +351 234 370706; fax: +351 234 370084.  
E-mail address: [mac@ua.pt](mailto:mac@ua.pt) (M.A. Coimbra).

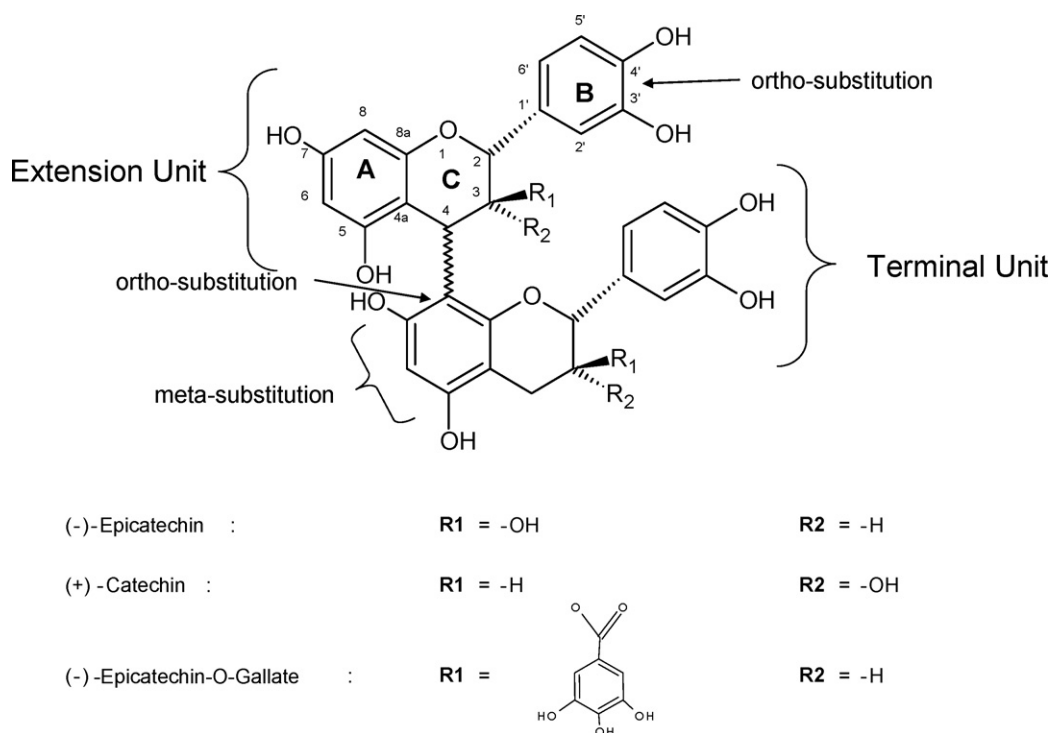


Fig. 1. Structure of procyanidin dimer containing a C4–C8 interflavanic linkage.

systems and to build calibration and prediction models. Partial least squares (PLS) regression is one of the most used models [19–22]. The PLS algorithm is based on a bilinear model, where the information contained in the **X** data matrix is projected over a small number of latent variables known as PLS components. The **Y** data matrix is actively used for estimating the latent variables and ensure that the first components of these are the most relevant for predicting the **Y** dependent variables. The interpretation of the relationships between the **X** and **Y** data is simplified to the relationships between the smaller number of PLS components [23,24].

The method of multivariate regression by PLS has been extensively used in chemometrics, where a wide field of applications has been found. This has been proposed and implemented in the routine analysis of a large number of parameters on wine industry [17,25,26]. Also, it has been used for the identification and characterisation of polysaccharides and proteins and for their quantification [27–29]. The estimation of the procyanidins *DP<sub>n</sub>* in the range of 2–6 residues by using PLS algorithm was recently attempted in dried red wine samples previously purified by C18-solid phase extraction [30].

In this work, regression models based on the FTIR spectral region between 1800 and 700  $\text{cm}^{-1}$  using PLS1 and O-PLS1 were assayed in grape seeds freeze-dried procyanidin extracts. Identification of procyanidins molecular features related to the absorbance characteristics in the mid infrared region were established and correlated with procyanidins *DP<sub>n</sub>*.

## 2. Experimental

### 2.1. Materials

Methanol, ethyl acetate, *n*-hexane, and acetone, from Sigma–Aldrich Co. (St. Louis, MO, USA), were of analytical grade. Water, from Merck (Darmstadt, Germany), acetic acid and acetonitrile, from Sigma–Aldrich Co. (St. Louis, MO, USA), were of chromatographic grade quality. (+)-Catechin, (–)-epicatechin, (–)-epicatechin-*O*-gallate, procyanidin B2, and benzyl mercaptan,

were purchased from Fluka Sigma–Aldrich Co. (St. Louis, MO, USA). Other reagents were of analytical grade or higher available purity.

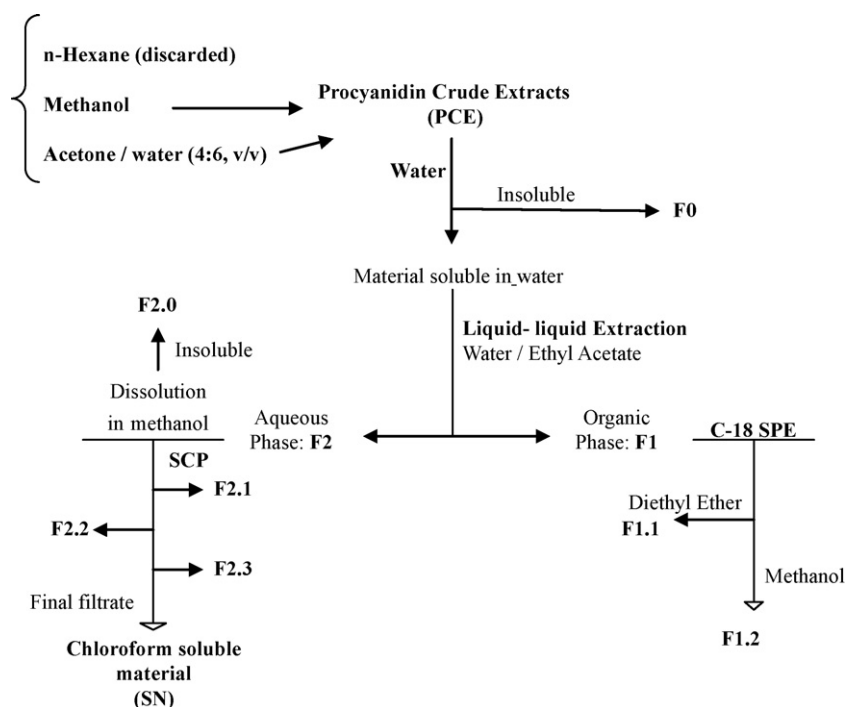
### 2.2. Plant material

Seeds were obtained from grapes (*Vitis vinifera* L.) of the white variety ‘Chardonnay’ at technological maturity, in Bairrada Appellation, from an experimental vineyard (Estação Vitivinícola da Bairrada, Anadia, Portugal), during transfer of the musts for fermentation. A mixture of red grape varieties ‘Touriga Nacional’, ‘Touriga Francesa’, and ‘Tinta Roriz’, were also obtained from the wine producers of Adega Cooperativa de Pinhel (Pinhel, Portugal). The remaining pulp and skins were separated from grape seeds by decantation and sieving (pore size <2.8 mm diameter). The seeds were then submitted to several wash cycles with water (200  $\text{g L}^{-1}$ ) under gentle stirring with a magnetic bar at 4 °C for a minimum of 3 days, with two water exchanges a day, until a minimum turbidity was constant, assuring that no remaining adherent tissues were present. The purified seeds were then washed with ethanol, air dried at room temperature, and stored at 4 °C until further analysis.

### 2.3. Data set preparation and characterisation

#### 2.3.1. Procyanidin crude extracts (PCE)

Seeds were immersed into liquid nitrogen, milled in a domestic coffee mill and sieved (pore size <0.75 mm diameter). The extraction methodology was adapted from Guyot et al. [12], as described by Cardoso et al. [31]. Seed powder was extracted three times with *n*-hexane to remove the lipids. It was then treated three times with methanol containing 5% acetic acid to extract the phenolic compounds. The methanol extracts were filtered through a G3 sintered glass filter, combined, concentrated under vacuum at 40 °C, with several additions of water to assure the complete removal of methanol and acetic acid, frozen, and freeze-dried, to give the methanol procyanidin crude extracts (PCE). The residue resultant from the methanolic treatment was extracted three times with acetone/water solution containing 5% acetic acid, whereas the acetone



**Scheme 1.** Extraction and fractionation of grape seed flavan-3-ols. SPE: solid phase extraction; SCP: sequential chloroform precipitations.

was eliminated as described for methanol. The aqueous solution was frozen and freeze-dried to obtain the acetone/water PCE.

### 2.3.2. Procyanidins fractionation

The methanol PCE and the acetone/water PCE from white and red grape seeds were fractionated according to the methanol/chloroform graded precipitations proposed by Saucier et al. [9] as summarised in Scheme 1. The PCE powder ( $10 \text{ g L}^{-1}$ ) was dissolved in water containing 5% acetic acid and the undissolved material ( $F_0$ ) was removed by centrifugation (Centrifuge 3K30, Sigma, St. Louis, MO, USA). The supernatant was then submitted three times to a liquid–liquid extraction with ethyl acetate, using a water/ethyl acetate ratio of 6:4 (v/v), resulting in an organic phase ( $F_1$ ) and an aqueous phase ( $F_2$ ). The  $F_1$  solution (PCE organic phase) was concentrated and loaded into a C18-solid phase extraction column (SPE-C18, SPE, Supelco-Discovery – 5 g) by eluting with diethyl ether followed by methanol, allowing to obtain fractions  $F_{1.1}$  and  $F_{1.2}$ , respectively. The  $F_2$  solution (PCE aqueous phase) was evaporated to dryness and redissolved in methanol ( $10 \text{ g L}^{-1}$ ). The undissolved material was removed by centrifugation and the supernatant was submitted to successive additions of chloroform until a new precipitate was formed. The precipitate was then collected by centrifugation, dissolved in water, rotary-evaporated with several additions of water to completely remove the organic solvents, frozen, and freeze-dried. The material soluble in chloroform after the last precipitation was named “SN” and was recovered as described for the precipitated fractions. The composition of the fractions containing polymeric procyanidins is given in Table 1.

### 2.3.3. Thiolytic and HPLC

Thiolytic was carried out according to the methodology described by Naczki and Shahidi [10] and HPLC analysis followed the conditions described by Peng et al. [32]. The HPLC apparatus used was from PerkinElmer (series 200), with UV–vis detector (785A UV-VIS Detector). Samples were loaded at  $30^\circ\text{C}$  into a C<sub>18</sub> column (LichroCart 250–4 Superspher 100 RP-18) equipped with a C<sub>18</sub> guard cartridge with the same packing material equilibrated

with 0.2% (v/v) formic acid (eluent A). Phenolic compounds were eluted by a gradient with 82% (v/v) acetonitrile and 0.04% (v/v) formic acid (eluent B) from 0% to 15% eluent B in the first 15 min; 15% to 16% from 15 to 40 min; 16% to 17% from 40 to 45 min; 17% to 43% from 45 to 48 min; 43% to 52% from 48 to 49 min, held isocratic at 52% from 49 to 56 min, reduced from 52% to 43% from 56 to 57 min, reduced from 43% to 17% from 57 to 58 min, and reduced from 17% to 0% from 58 to 60 min. Samples were loaded, at least, in duplicate. Peaks were detected at 280 nm and the monomers and procyanidin B2 dimer were identified by comparison of their retention times with standards. The epicatechin thioderivative was identified by comparison with the retention time of the products of procyanidin B2 dimer after thiolytic; the catechin and epicatechin-*O*-gallate thioderivatives were identified by their retention times and abundance and confirmed by analysis of their mass spectra using a LC–MS (Waters alliance 2690) as described by Passos et al. [14]. The average degree of polymerisation ( $DP_n$ ) was calculated as the ratio of all the areas of flavan-3-ols units (thioether adducts plus terminal units) to the sum of the areas of catechin, epicatechin, and epicatechin-*O*-gallate corresponding to terminal units. The  $DP_n$  estimated was in the range  $DP_n \pm 0.5$ .

The calibration curves for estimation of phenolic compounds were obtained using (+)-catechin, (–)-epicatechin, (–)-epicatechin-*O*-gallate, and procyanidin B2 dimer, in the range of concentration of  $0.005$ – $0.5 \text{ g L}^{-1}$ . The quantification of phenolics in the fractions was made by comparison of the chromatographic area after thiolytic degradation of the samples with the respective calibration curve. As thioderivative standards were not available, they were quantified by using the respective monomer calibration curves based on the similar response factors to the correspondent monomeric units [13].

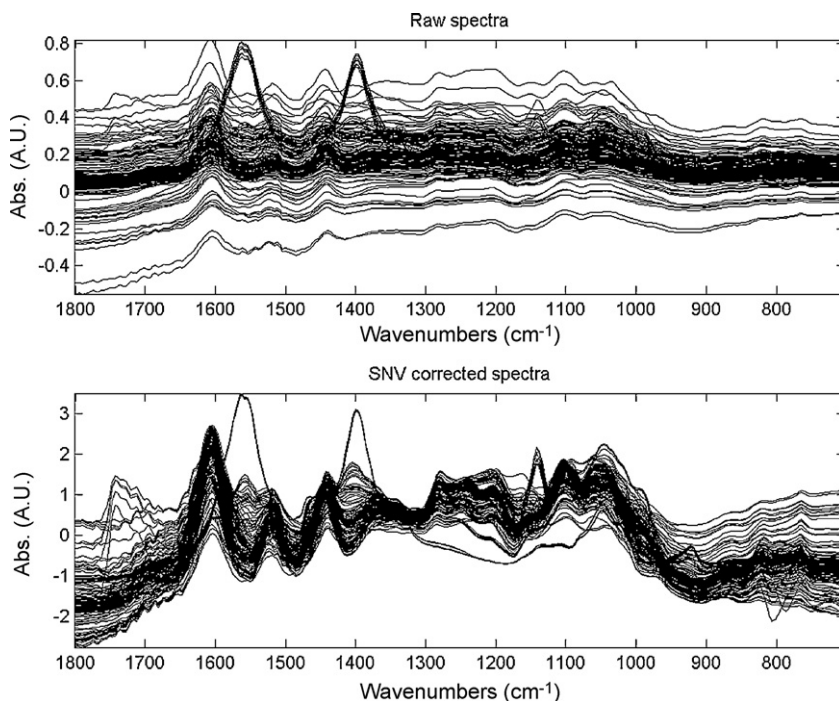
## 2.4. FTIR spectroscopy and multivariate analyses

The FTIR spectra of each fraction presented in Table 1 were obtained using a Golden Gate single reflection diamond ATR system in a Bruker IFS-55 spectrometer with a Deuterated Triglycine Sul-

**Table 1**  
Procyanidin composition of grape seed extract obtained according to the fractionation in Scheme 1.

Fraction	CHCl <sub>3</sub> (%)	Yield (%)	Total procyanidin (% w/w)	Flavan-3-ol composition (%)						DP <sub>n</sub>
				(+)-Catechin		(–)-Epicatechin		(–)-Epicatechin-O-gallate		
				Terminal	Extension	Terminal	Extension	Terminal	Extension	
WM-PCE	–	21.4	55.2	8.7	9.4	2.9	63.9	1.9	13.1	7.4 (7.5)
WM-F0	–	11.3	8.6	6.4	5.8	3.3	65.7	1.6	17.2	8.8
WM-F1	–	8.7	84.1	34.5	17.5	17.4	22.5	3.7	4.4	1.8
WM-F2.0	–	1.4	22.1	5.3	4.7	3.2	65.4	5.5	16.1	7.2
WM-F2.1	29	5.6	35.2	5.9	4.1	3.6	68.8	1.2	16.4	9.4
WM-F2.2	52	10.5	45.1	7.1	4.6	3.7	66.4	1.7	16.5	8.0
WM-F2.3	66	8.5	43.6	8.3	5.1	4.9	65.6	2.4	13.8	6.4
WM-F2.4	73	6.6	68.5	7.5	3.7	4.5	65.1	2.7	16.5	6.8
WM-F2.5	79	4.2	49.2	9.8	6.5	5.5	65.8	2.2	10.1	5.7
WM-F2.6	84	3.7	43.6	11.1	7.8	6.6	62.6	2.4	9.7	5.0
WA-PCE	–	4.5	53.7	11.4	5.6	9.9	58.4	2.5	12.2	4.2 (3.6)
WA-F0	–	21.4	33.2	7.7	4.7	4.6	64.1	1.8	17.1	7.1
WA-F2.1	42	8.2	29.4	7.6	4.9	3.6	61.9	1.8	20.3	7.7
WA-F2.2	58	7.2	31.7	8.6	3.2	3.6	64.2	2.5	17.9	6.8
WA-F2.3	71	4.0	25.5	11.1	7.0	5.8	60.5	2.4	13.2	5.2
WA-F2.4	84	2.8	13.6	13.6	6.1	10.0	60.5	2.0	7.7	3.9
RM-PCE	–	7.6	8.8	11.9	15.1	2.6	59.8	1.6	9.0	6.2 (7.0)
RM-F0	–	11.1	11.8	8.9	8.5	3.6	62.9	0.9	15.3	7.5
RM-F1.2	–	0.7	28.2	25.5	18.1	7.8	35.6	2.4	10.6	2.8
RM-F2.0	–	0.7	46.6	6.6	7.9	2.8	65.6	0.2	16.9	10.4
RM-F2.1	43	3.1	27.3	6.9	7.4	2.5	67.0	0.9	15.2	9.7
RM-F2.2	60	3.0	29.4	10.3	11.1	4.3	69.8	0.3	4.2	6.7
RM-F2.3	82	2.2	27.0	15.7	11.7	6.9	59.0	1.1	5.5	4.2
RA-PCE	–	3.0	20.2	9.1	9.0	3.3	58.1	1.1	19.4	7.4 (7.5)
RA-F0	–	14.8	41.4	6.8	7.5	2.7	55.2	1.3	26.5	9.3
RA-F1.2	–	2.3	44.7	10.7	10.2	3.9	42.8	4.7	27.8	5.2
RA-F2.0	–	6.9	31.0	5.7	6.5	2.6	56.1	1.0	28.4	10.8
RA-F2.1	43	8.6	55.2	8.4	7.1	3.5	57.8	1.2	21.8	7.6
RA-F2.2	60	3.4	53.0	11.3	7.4	4.7	59.2	1.2	16.1	5.8
RA-F2.3	79	5.2	15.5	21.4	9.3	9.4	40.8	2.5	16.5	3.0

W: white grape seeds; R: red grape seeds; M: methanol extracts; A: acetone/water extracts; CHCl<sub>3</sub>: percentage of chloroform used; yield: percentage of mass recovered from grape seed into the corresponding PCE, and from the correspondent PCE to all the remaining fractions; DP<sub>n</sub>: average degree of polymerisation. In parenthesis are the values obtained by FTIR/O-PLS1 regression.



**Fig. 2.** FTIR spectra in the 1800–700 cm<sup>−1</sup> region. Top: raw spectra and bottom: SNV corrected spectra.

fate (DTGS) detector. The spectra were recorded at the absorbance mode from 4000 to 400  $\text{cm}^{-1}$  (mid infrared region) at the resolution of 8  $\text{cm}^{-1}$ . Five analytical replicate spectra (128 co-added scans) were collected for each sample. The measured spectra were transferred via a JCAMP.DX format into the data analysis software developed in the Institut National Agronomique Paris-Grignon in collaboration with the University of Aveiro [33]. The multivariate calibration was applied in the 1800–700  $\text{cm}^{-1}$  region and due to amplification spectrum effects they were pre-processed using SNV (Standard Normal Variate). Fig. 2 shows the used spectral region before and after the SNV correction.

### 2.5. Calibration model framework

In order to build the calibration model for the quantification of the *DPn*, a Monte Carlo cross-validation [34] framework was used. The dataset was split into a calibration (learning set) and a validation (external) set to assess the predictive power of the *DPn* model. Table 1 includes the 26 samples used to calibrate and the four PCE samples (which are the crude extracts, containing a heterogeneous material) for prediction. The splitting process consisted in sorting the *DPn* values and then, randomly, selected 40% of the samples, where replicates are considered a sample, were used as validation set. The remaining 60% of the samples, with replacement, were used as calibration set. This procedure was repeated several times (iterations), 200 regression models were built and for each one the “optimal” model dimensionality based on the RMSECV value and LV was recovered. This allowed one to see how many times of a given LV/RMSECV “optimal” pair (distribution profile) was used to build a predictive model. The selection of model complexity was based on the most frequent pair of LV/RMSECV. Then, the selected model dimensionality was used to predict the parameters of interest from the external set, expressed as root mean square error of prediction (RMSEP).

Since this approach is very computational demanding when using PLS1 [20], the Principal Component Transform PLS1 (PCT-PLS1) [23] was used instead to build the calibration models in order to accelerate the Monte Carlo cross-validation process.

## 3. Results and discussion

### 3.1. Characterisation of grape seed procyanidin fractions by HPLC

The procyanidin dataset used in this study is composed by 26 samples purified from the methanol and acetone/water procyanidin crude extracts (PCE) from seeds of white and red grapes. The flavan-3-ols content vary from 8.6% (WM-F<sub>0</sub>) to 84.1% (WM-F<sub>1</sub>), having an average of 36.1% and standard deviation of 17.8%, thus, covering a large range and dispersion of procyanidin concentrations in samples (Table 1). These flavan-3-ols are composed by flavan-3-ol monomers of catechin, epicatechin and epicatechin-O-gallate, and by procyanidins (oligomers and polymers) with these same three constituting units. In particular, the procyanidin crude extracts (Table 1, PCE) have a *DPn* range between 4.2 and 7.4. By fractionation in chloroform/methanol solutions, the data set was enlarged, allowing to obtain fractions containing procyanidins with *DPn* ranging from 1 (SN material) to 10.8. According to the data in Table 1, an increase in the percentage of chloroform allows to obtain precipitated procyanidin fractions with lower *DPn* (Table 1, F<sub>2</sub> extracts).

For the majority of the samples, catechin is the main terminal unit, whereas epicatechin occurs as the main extension unit. The epicatechin-O-gallate unit accounts for approximately 4.5–32.5% of the total procyanidin residues.

### 3.2. Characterisation of procyanidins by FTIR

The characteristic wavenumbers related to the phenolic compounds are associated to the presence of an OH band between 3600 and 3200  $\text{cm}^{-1}$ . Also, they show aromatic, ester, alcohol, and ether bands in the region between 1800 and 700  $\text{cm}^{-1}$ . All spectra have a similar profile (Fig. 2) and, according to the bibliography [35,36], changes in the procyanidins aromatic ring bands are expected to occur in this spectral region.

### 3.3. Calibration models for estimation of procyanidin *DPn*

#### 3.3.1. Calibration model for estimation of procyanidin *DPn* using PCT-PLS1

Using the 1800–700  $\text{cm}^{-1}$  region for all the FTIR spectra of the procyanidin fractionated samples, excluding the crude extracts, a PLS1 regression procedure was applied for estimation of their *DPn*. To accomplish that, a calibration model with 8 latent variables (LVs), using an internal-cross-validation (leave-5-out) procedure, was seen to have predictive power. The relative root mean square error of cross-validation (rRMSECV) obtained was 11.7%, with a coefficient of determination ( $R^2$ ) of 0.91 and a root mean square error of prediction (RMSEP) of 2.58.

However, the relative high number of LVs in the previous model introduces difficulties in the interpretation of the **pp** (predictive loadings) vector profiles. As a matter of fact the **y** variance explained by this PLS1 regression model shows some irregularity concerning LV2 (Table 2), suggesting that some systematic variations present in the spectra are not related to the **y** variability. In these cases, one approach is to remove from the model all spectra variations orthogonal to the factor of interest (*DPn*). One such method is the O-PLS (Orthogonal Projections to Latent Structures) [37]. Therefore, this procedure was applied with the aim of improving the interpretation of the PLS1 regression model by removing orthogonal artefacts not related to the *DPn* profile. Nonetheless, it should be noted that O-PLS method does not improve the robustness of a calibration model.

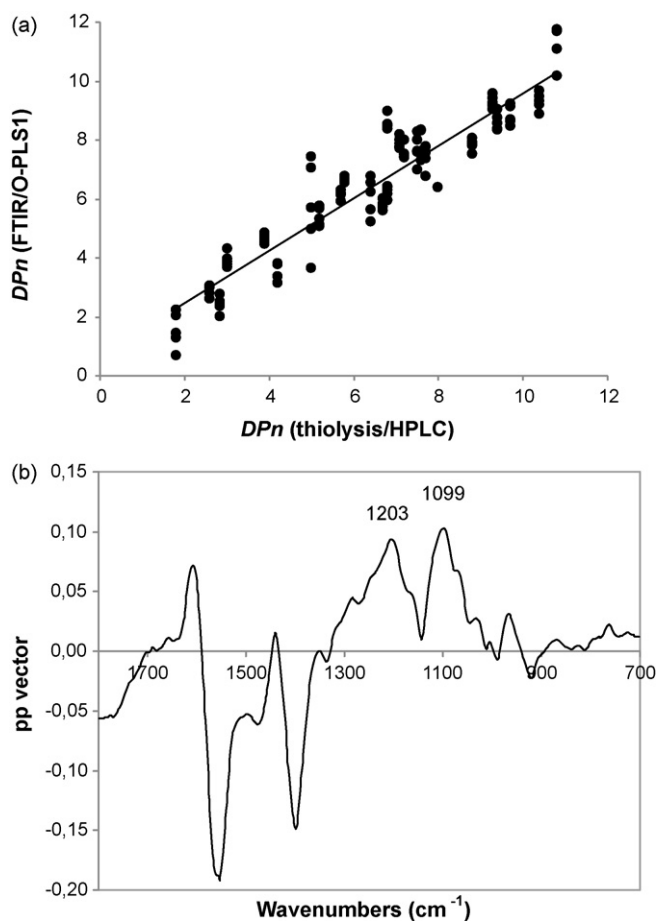
#### 3.3.2. Calibration model for estimation of procyanidin *DPn* using O-PLS

The O-PLS was used in the Monte Carlo cross-validation procedure (similar to the one described in Section 2.5) to remove orthogonal systematic variations from the spectra with respect to the *DPn* values. The Monte Carlo cross-validation indicated seven O-PLS components to be removed from the regression model resulting into a 1 LV model for calibration/interpretation purposes. This procedure has removed 88% of non-correlated variations present in FTIR spectra. The regression model obtained presented a relative rRMSECV of 8.6%, with a  $R^2$  of 0.95 and a RMSEP of 2.58 (Fig. 3a), which shows that it is not significantly different from the PLS1 model with 8 LVs.

**Table 2**  
PLS1 explained variance values for each LV for the estimation of procyanidin *DPn*.

LV	$R^2\mathbf{X}$ (%)	Accumulated $R^2\mathbf{X}$ (%)	$R^2\mathbf{y}$ (%)	Accumulated $R^2\mathbf{y}$ (%)
1	41.8	41.8	23.1	23.1
2	33.3	75.1	9.8	32.9
3	10.7	85.8	27.3	60.2
4	5.2	91.0	8.8	69.0
5	2.4	93.4	8.3	77.3
6	2.1	95.4	5.7	83.0
7	2.0	97.5	4.5	87.5
8	0.6	98.0	3.0	90.5

LV: latent variables;  $R^2\mathbf{X}$ :  $\mathbf{X}$  variance;  $R^2\mathbf{y}$ :  $\mathbf{y}$  variance.



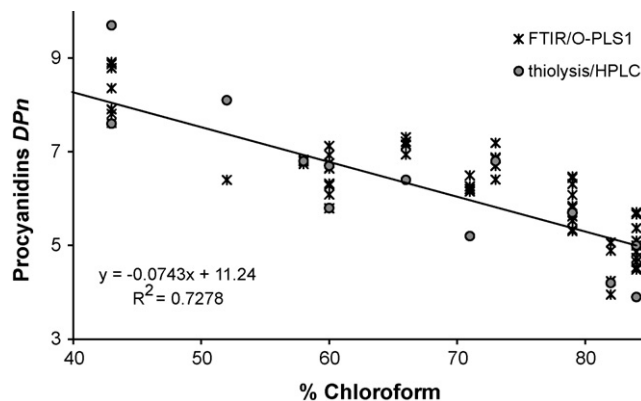
**Fig. 3.** Regression model for estimation of procyanidin  $DP_n$ . (a) Relationship between  $DP_n$  values estimated by FTIR/O-PLS1 and thiolysis/HPLC. (b) O-PLS1 **pp** vector profile with 1 LV.

### 3.3.3. Validation of O-PLS model for estimation of procyanidin $DP_n$

In order to validate the regression model obtained for the estimation of the procyanidins  $DP_n$  using the O-PLS1 model (Fig. 3), the procyanidin crude extracts (PCE) were used. The results show that similar values were obtained using the FTIR/O-PLS1 calibration curve and the thiolysis/HPLC procedures. Such evidence is even more important considering that the PCE extracts represent a more complex matrix than the fractionated extracts, allowing to infer that this approach can be a useful tool for the estimation of  $DP_n$  in non-purified extracts.

### 3.4. Molecular features relating the IR absorbance characteristics and the procyanidins $DP_n$

Fig. 3b shows the **pp** vector profile of the O-PLS1 regression model where it is possible to identify the most important bands related to the  $DP_n$  value. Since this model was built using 1 LV, its interpretation is easier than the previous 8 LV PLS1 model. The observation of the positive peaks at 1203 and 1099  $\text{cm}^{-1}$  suggests a possibility to assess the  $DP_n$  of procyanidins by monitoring the aromatic substitutions. The 1,2-disubstitutions on the aromatic rings have been related with the existence of three peaks (two medium and one strong) in the range of 1200–900  $\text{cm}^{-1}$  and 1,3-disubstitutions on the aromatic rings have been related with the existence of four peaks (two medium and two strong) in 1100–700  $\text{cm}^{-1}$  range [35]. The  $DP_n$  is obtained by the balance between one terminal and several extension units. The molecular



**Fig. 4.** Relation of procyanidin  $DP_n$  estimated by FTIR/O-PLS1 and thiolysis/HPLC vs. solubility of the fractions in methanol/chloroform solutions.

structure of a dimer such as the one represented in Fig. 1, comprises two flavan-3-ol residues (one terminal and one extension unit) bonded by an interflavanic linkage. Each flavan-3-ol residue is composed by two aromatic rings (A and B) and a non-aromatic ring (C). The hydroxyl groups on ring A occur in C5 and C7, resulting in a 1,3-disubstitution (meta-substitution), whereas on B ring, the hydroxyl groups occur in C3' and C4', forming a 1,2-disubstitution of the aromatic ring (ortho-substitution). The interflavanic linkage between the two composing units of the dimer originates an extra C7–C8 ortho-substitution of the aromatic ring A of the terminal unit of the procyanidin. In total, a dimeric procyanidin with a  $DP=2$ , formed by two flavan-3-ol units, has one meta-substitution and one ortho-substitution per unit plus one meta-substitution due to the bond formed between the two composing units. In general, a procyanidin formed by  $n$  flavan-3-ol units ( $DP=n$ ), has  $n+(n-1)$  1,2-disubstitutions and  $n$  1,3-disubstitutions. As a consequence, the ratio 1,2-disubstitutions/1,3-disubstitutions can be expressed by the following formula:  $(2n-1)/n$ .

### 3.5. Relation of procyanidins $DP_n$ and solubility in methanol/chloroform solutions

Fig. 4 shows the relationship of the solubility of procyanidins in methanol/chloroform solutions and their average degree of polymerisation. The material that precipitate in solutions of approximately 40% chloroform tend to have a  $DP_n$  of 8–10, the material that precipitate in solutions near 60% chloroform tend to have a  $DP_n$  of 6–7, and the samples that precipitate in solutions in the order of 80% chloroform tend to have a  $DP_n$  of 3–5. The samples that do not precipitate in these solutions contain mainly flavan-3-ol monomers or, at least, procyanidin dimers (data not shown). The graded precipitation in methanol/chloroform solutions is a simple and quick method to evaluate the  $DP_n$  of procyanidins in a fraction, as well as a rough but suitable methodology to recover enriched fractions of procyanidins with a defined degree of polymerisation that can be used, for example, in industry. In this work, the precipitation methodology was an undoubtedly important step in order to obtain samples for performing a suitable calibration/validation data set.

## 4. Concluding remarks

The need for testing fast and cheap but reliable methodologies for the estimation of the  $DP_n$  of procyanidins has been the basis of the present work. The difference in the solubility of the procyanidins in methanol/chloroform solutions was shown to be useful for this purpose. This methodology of fractionation also allowed to obtain the number and representative samples required to build a

FTIR model for the determination of the *DPn* of procyanidins. Using the resulting calibration model it is possible to estimate the *DPn* of grape seed procyanidins and to assign the major positive changes in the **pp** vector at 1203 and 1099  $\text{cm}^{-1}$  to the increase of aromatic substitutions in polymerised molecules.

Although the methodology has not yet been tested for raw materials other than grape seeds, it is expected that applications to other procyanidin sources can be successfully obtained using these methodologies.

### Acknowledgements

The authors gratefully acknowledge Pedro Vasco (Universidade de Aveiro) and Eng. A. Dias Cardoso, Eng. José Carvalheira, and all staff from Estação Vitivinícola da Bairrada, Portugal, for the technical support and raw material provided. Thanks are also due to Adegas Cooperativa de Pinhel for providing the seed from red grapes. This work was supported by the Research Unit 62/94 QOPNA and project POCTI/EQU/47533/2002. Cláudia Passos was supported by a PhD grant by FCT, SFRH/BD/19072/2004.

### References

- [1] T. Escribano-Bailón, Y. Gutiérrez-Fernández, J.C. Rivas-Gonzalo, C. Santos-Buelga, *J. Agric. Food Chem.* 40 (1992) 1794.
- [2] T. Fuleki, J.M. Ricardo da Silva, *J. Agric. Food Chem.* 45 (1997) 1156.
- [3] M. Monagas, C. Gómez-Cordovés, B. Bartolomé, O. Laureano, J.M. Ricardo da Silva, *J. Agric. Food Chem.* 51 (2003) 6475.
- [4] A.G.H. Lea, G.M. Arnold, *J. Sci. Food Agric.* 29 (1978) 478.
- [5] S. Vidal, L. Francis, S. Guyot, N. Marnet, M. Kwiatkowski, R. Gawel, V. Cheynier, E.J. Waters, *J. Sci. Food Agric.* 83 (2003) 564.
- [6] C. Santos-Buelga, A. Scalbert, *J. Sci. Food Agric.* 80 (2000) 1094.
- [7] C. Le Bourvellec, M. Picot, C.M.G.C. Renard, *Anal. Chim. Acta* 563 (2006) 33.
- [8] S. Guyot, N. Marnet, D. Laraba, P. Sanoner, J.F. Drilleau, *J. Agric. Food Chem.* 46 (1998) 1698.
- [9] C. Saucier, M. Mirabel, F. Daviaud, A. Longieras, Y. Glories, *J. Agric. Food Chem.* 49 (2001) 5732.
- [10] M. Naczki, F. Shahidi, *J. Chromatogr. A* 1054 (2004) 95.
- [11] D. Ferreira, S. Guyot, N. Marnet, I. Delgado, C.M.G.C. Renard, M.A. Coimbra, *J. Agric. Food Chem.* 50 (2002) 4537.
- [12] S. Guyot, N. Marnet, J.F. Drilleau, *J. Agric. Food Chem.* 49 (2001) 14.
- [13] N. Vivas, M.F. Nonier, N.V. de Gaulejac, C. Absalon, A. Bertrand, M. Mirabel, *Anal. Chim. Acta* 513 (2004) 247.
- [14] C.P. Passos, S.M. Cardoso, M.R.M. Domingues, P. Domingues, C.M. Silva, M.A. Coimbra, *Food Chem.* 105 (2007) 1457.
- [15] Y. Hayasaka, E.J. Waters, V. Cheynier, M.J. Herderich, S. Vidal, *Rapid Commun. Mass Spectrom.* 17 (2003) 9.
- [16] Y. Yang, M. Chien, *J. Agric. Food Chem.* 48 (2000) 3990.
- [17] J.S. Jensen, M. Egebo, A.S. Meyer, *J. Agric. Food Chem.* 56 (2008) 3493.
- [18] C.S. Robb, S.E. Geldart, J.A. Seelenbinder, Phyllis R. Brown, *J. Liq. Chromatogr. Relat. Technol.* 25 (2002) 787.
- [19] H. Martens, T. Naes, *Multivariate Calibration*, Wiley, Chichester, 1994.
- [20] S. Wold, H. Martens, H. Wold, in: A. Ruhe, B. Kagstrom (Eds.), *Lecture Notes in Mathematics* 1983. Proceedings of the Conference Matrix Pencils, March 1982, Springer Verlag, Heidelberg, 1983, p. 286.
- [21] S. Wold, M. Sjostrom, L. Eriksson, *Chemometr. Intell. Lab. Syst.* 58 (2001) 109.
- [22] S. Wold, J. Trygg, A. Berglund, H. Antti, *Chemometr. Intell. Lab. Syst.* 58 (2001) 131.
- [23] A.S. Barros, D.N. Rutledge, *Chemometr. Intell. Lab. Syst.* 73 (2004) 245.
- [24] D.A. Guillen, M. Palma, R. Natera, R. Romero, C.G. Barroso, *J. Agric. Food Chem.* 53 (2005) 2412.
- [25] A. Edelmann, J. Diewok, K.C. Schuster, B. Lendl, *J. Agric. Food Chem.* 49 (2001) 1139.
- [26] M. Dubernet, V. Dubernet, S. Coulomb, M. Lerch, I. Traineau, XXV Congrès Mondial de la Vigne et du Vin, vol. II, Paris, 2000, p. 215.
- [27] J.C. Boulet, T. Doco, J.M. Roger, *Chemometr. Intell. Lab. Syst.* 87 (2007) 295.
- [28] M.A. Coimbra, A.S. Barros, E. Coelho, F. Gonçalves, S.M. Rocha, I. Delgado, *Carbohydr. Polym.* 61 (2005) 434.
- [29] M. Cocchi, C. Durante, M. Grandi, P. Lambertini, D. Manzini, A. Marchetti, *Talanta* 69 (2006) 1166.
- [30] K. Fernández, E. Agosin, *J. Agric. Food Chem.* 55 (2007) 7294.
- [31] S.M. Cardoso, S. Guyot, N. Marnet, J.A. Lopes-da-Silva, C.M.G.C. Renard, M.A. Coimbra, *J. Sci. Food Agric.* 81 (2005) 21.
- [32] Z. Peng, Y. Hayasaka, P.G. Iland, M. Sefton, P. Hoj, E.J. Waters, *J. Agric. Food Chem.* 49 (2001) 26.
- [33] A.S. Barros, Contribution à la sélection et la comparaison de variables caractéristiques, PhD thesis, Institut National Agronomique Paris-Grignon, France, 1999, Chapter 1.
- [34] Q.-S. Xu, Y.-Z. Liang, *Chemometr. Intell. Lab. Syst.* 56 (2001) 1.
- [35] M.M. Ramos-Tejada, J.D.G. Durán, A. Ontiveros-Ortega, M. Espinosa-Jimenez, R. Perea-Carpio, E. Chibowski, *Colloid Surf. B: Biointerfaces* 24 (2002) 297.
- [36] M. Contreras-Domínguez, S. Guyot, N. Marnet, J. Le Petit, I. Perraud-Gaime, S. Roussos, *C. Augur, Biochimie* 88 (2006) 1899.
- [37] J. Trygg, S. Wold, *J. Chemometr.* 16 (2002) 119.