ELSEVIER

Carbohydrate Polymers 50 (2002) 85–94

Carbohydrate Polymers

www.elsevier.com/locate/carbpol

Determination of the degree of methylesterification of pectic polysaccharides by FT-IR using an outer product PLS1 regression

A.S. Barros[a], I. Mafra[a], D. Ferreira[a,b], S. Cardoso[a,c], A. Reis[a], J.A. Lopes da Silva[a], I. Delgadillo[a], D.N. Rutledge[d], M.A. Coimbra[a,*]

[a]*Departamento de Química, Universidade de Aveiro, 3810-193 Aveiro, Portugal*
[b]*Escola Superior Agrária de Viseu, Instituto Politécnico de Viseu, 3500 Viseu, Portugal*
[c]*Escola Superior Agrária de Bragança, Instituto Politécnico de Bragança, 5300 Bragança, Portugal*
[d]*Lab. Chimie Analytique, Institut National Agronomique, Paris-Grignon, 75005 Paris, France*

This paper is to honour Professor Joachim Thiem on the occasion of his 60th anniversary

Received 8 June 2001; revised 26 November 2001; accepted 14 December 2001

Abstract

The combination of the absorbance spectra in two infrared regions (1800–1500 and 1200–850 cm^{-1}) was used to build a calibration model for the determination of the degree of methylesterification (DM) of pectic polysaccharides. The wavenumbers in the region 1800–1500 cm^{-1} are related to the carbonyl esters and carboxylates and those in the region 1200–850 cm^{-1} are related to the sugars composition. The model was done by means of a matrix of the outer product of the two regions and PLS1 regression, using cell wall pectic polysaccharide extracts from olive and pear pulps. The application of the Durbin–Watson criterion to the regression vector of OP-PLS1 gave an 'optimal' model with nine latent variables, with a RMSEP of 7.7% and a coefficient of determination (R^2) of 0.95. The model explained 95% of the total variability present in the **y** vector. The results allowed the prediction of the DM of commercial standard pectic polysaccharides and of impure pectic polysaccharide extracts. A prediction error of 22% was found, which contrasted with the prediction error of 72% when only the region 1800–1500 cm^{-1} region was used.

Keywords: Degree of methylesterification; Pectic polysaccharides; FT-IR spectroscopy; Outer product analysis; PLS regression

1. Introduction

Pectic polysaccharides are involved in the complex fibrillar network of plant cell wall structure that defines the mechanical and functional properties of the cell wall (Cosgrove, 2001; Roberts, 2001). As structural components, pectic polysaccharides influence the texture of fruits on ripening (Jiménez et al., 2001; Mafra et al., 2001; Martin-Cabrejas, Waldron, Selvendran, Parker, & Moates, 1994; Paull, Gross, & Qiu, 1999; Vierhuis, Schols, Beldman, & Voragen, 2000), storage (Bartley & Knee, 1982) and processing (Femenia, Sanches, Simal, & Rosselló, 1998). Pectic polysaccharides are also of great importance in the food industry as gelling agents in jams and jellies, fruit preparations for dairies, stabilisers in fruit and milk beverages (Claus, Nielsen, & Glahn, 1998) and dietary fibres (Sun & Hughes, 1999; Sun, Fang, Goodwin, Lawther, & Bolton, 1998).

Pectic polysaccharides are constituted mainly of galacturonic acid residues that can be partially esterified. The degree of methylesterification (DM) is defined as the percentage of carboxyl groups esterified with methanol (Voragen, Pilnik, Thibault, Axelos, & Renard, 1995). The presence of methyl ester groups affects the cross-linking of pectinate molecules by Ca^{2+}, which plays an important role in the organisation of polysaccharides in plant cell walls (Brett & Waldron, 1996; Wellner, Kacuráková, Malovíková, Wilson, & Belton, 1998) and, consequently, may influence the texture properties of fruits during ripening and processing. The gelation mechanisms of pectins are

* Corresponding author. Fax: +351-234-370-084.
E-mail address: mac@dq.ua.pt (M.A. Coimbra).

Abbreviations: Ara, arabinose; CDTA, *trans*-1,2-cyclohexanediamine-*N*,*N*,*N*′,*N*′-tetraacetic acid sodium salt; CWM, cell wall material; DM, degree of methylesterification; CV, coefficient of variation; Gal, galactose; GalA, galacturonic acid; HexA, hexuronic acid; LV, latent variable; MLR, multiple linear regression; OPA, outer product analysis; OP-PLS1, outer product-partial least squares regression; PCR, principal component regression; PLS, partial least squares regression; RMSEP, root mean square error of prediction; SDS, sodium dodecyl sulphate

0144-8617/02/$ - see front matter
PII: S0144-8617(02)00017-6

also dependent on DM (Grant, Morris, Rees, Smith, & Thom, 1973; Lopes da Silva, Gonçalves, & Rao, 1995; Walkinshaw & Arnott, 1981).

Several analytical methods have been proposed for the determination of the DM of pectic polysaccharides. They include alkali hydrolysis of the methylester groups and subsequent determination of the DM by titration (Mizote, Odagiri, Toei, & Tanaka, 1975) in galacturonic acid rich samples. The independent quantification of the total amount of hexuronic acids colorimetrically and the methanol released after alkali hydrolysis by GLC (Knee, 1978; McFeeters & Armstrong, 1984; Waldron & Selvendran, 1990) and by HPLC (Voragen, Schols, & Pilnik, 1986), and by enzymatic oxidation (Klavons & Bennet, 1986) are used when the pectic polysaccharides contain also neutral sugars. Another approach is the reduction of pectin methyl ester groups of GalA to Gal and the determination of the DM either by the increase in the amount of Gal, or by the change in the amount of GalA, quantified by GLC and colorimetric analysis, respectively (Maness, Ryan, & Mort, 1990). Instrumental techniques such as ^{1}H NMR (Grasdalen, Bakoy, & Larsen, 1988; Renard & Jarvis, 1999), ^{13}C NMR (Pfeffer, Doner, Hoagland, & McDonald, 1981) and FT-IR (Chatjigakis et al., 1998) spectroscopies have also been proposed.

The use of infrared spectroscopy on pectic substances was previously applied to distinguish between high and low methoxyl contents (Reintjes, Musco, & Joseph, 1962), and was proven to be a useful tool to distinguish and evaluate the methoxyl content of different commercial pectins with high and low levels of esterification (Haas & Jager, 1986). FT-IR spectroscopy as proposed by Chatjigakis et al. (1998) is a simple, quick and non-destructive method of DM evaluation in cell wall material extracts. The estimation of DM is based on a calibration curve using samples of standard pectins with known degree of esterification and the spectral bands at 1749 and 1630 cm^{-1}, assigned, respectively, to the absorption of the esterified and non-esterified carboxyl groups of pectin molecules. However, this methodology is shown not to be suitable for the determination of the DM of the pectic polysaccharides when other carboxylates and carbonyl ester groups, such as those from cell wall phenolics, are present.

FT-IR spectroscopy associated with chemometric techniques when applied in the region between 1200 and 850 cm^{-1}, has been used as a reliable and fast method for the evaluation of polysaccharide composition of pectic and hemicellulosic samples derived from orange and olive tissues (Coimbra, Barros, Barros, Rutledge, & Delgadillo, 1998; Coimbra, Barros, Rutledge, & Delgadillo, 1999) and to follow the effect of processing in cell wall polysaccharide extracts of fresh and dried pears (Ferreira, Barros, Coimbra, & Delgadillo, 2001). Based on the FT-IR spectroscopy associated with multivariate analysis, this work proposes the application of a novel method for the determination of the DM of pectic polysaccharides present in raw cell wall extracts using the combination of the 1800–1500 cm^{-1} and 1200–850 cm^{-1} regions of the FT-IR spectra, by means of an outer product analysis (OPA).

2. Material and methods

2.1. Sample extracts

The extracts analysed were from two different origins: olive (pomace and pulp) and pear. The cell wall materials (CWM) of olive and pear pulps were prepared as described by Mafra et al. (2001) and Ferreira et al. (2001), respectively. The pectic polysaccharides were solubilised from the CWMs by sequential extraction with imidazole solutions for olive and CDTA for pear. The olive pomace pectic polysaccharides were obtained by extraction of the alcohol insoluble residue with 0.02 M HNO_3 at 80 °C for 2 h.

Some pectic extracts used were fractionated by graded precipitation with ethanol and anion-exchange chromatography (Coimbra, Delgadillo, Waldron, & Selvendran, 1996). Commercial citrus pectin with DM 35 was provided by HP Bulmer Ltd and those with DM 26 (potassium salt), 67 (potassium salt) and 93 were obtained from Sigma.

2.2. Determination of the DM by GLC

The determination of the degree of esterification of pectic polysaccharides was based on the estimate of methanol content released by saponification (Waldron & Selvendran, 1990), expressed as mol% of methanol in relation to HexA. The sample (5 mg) was dispersed in water (2 ml) and sonicated for 10 min; 1-propanol (0.4 ml of 0.5 g l^{-1} stock solution) was added as internal standard. The polysaccharides were saponified with 0.8 ml of 2 M NaOH for 1 h at 20 °C. The sample suspension was neutralised by the addition of 0.8 ml of 2 M HCl and the resulting solution was filtered through a glass fibre filter (Whatman GF/C) and a nylon membrane filter NL16 0.2 μm (Schleicher & Schuell). The filtrate was injected (0.2 μl) in a gas chromatograph HP 5890 series II with a purged packed injector and an FID detector. The injector and detector temperatures were 180 and 220 °C, respectively. A DB-Wax column (30 m length, 0.53 mm i.d., and 1.0 μm film thickness) (J & W) was used with N_2 as carrier gas and a temperature program between 50 (1 min) and 140 °C, at a rate of 15 °C min^{-1}.

2.3. Carbohydrate analysis

Neutral sugars were released by Saeman hydrolysis (Selvendran, March, & Ring, 1979) and analysed as their alditol acetates by GLC (Blakeney, Harris, Henry, & Stone, 1983; Harris, Blakeney, Henry, & Stone, 1988) using a Carlo Erba 6000 with a split injector (split ratio 1:60) and a FID detector. A 30 m column DB-225 (J & W) with i.d. 0.25 mm and 0.15 μm film thickness was used. The injector and detector temperatures were 220 and 230 °C, respectively.

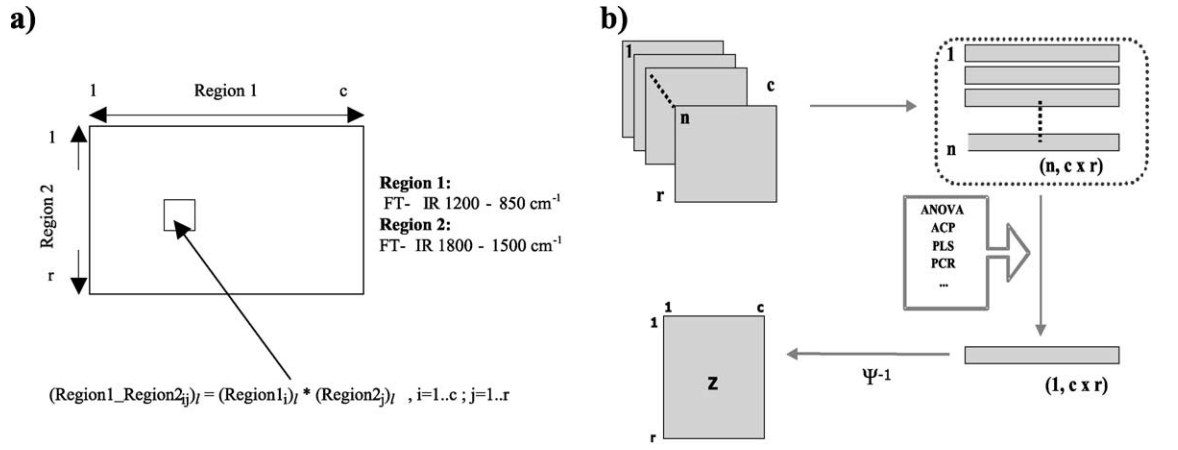

Fig. 1. (a) Calculation of the outer product between the two FT-IR regions; (b) Unfolding of the OPMs, concatenation of the vectors, statistical analysis of the resulting matrix and refolding of the vectors of calculated values.

The oven temperature program used was: 220 °C for 4 min, followed by 230 °C for 6.5 min with a rate of 25 °C min^{-1}. The flow rate of the carrier gas (H_2) was set at 1 ml min^{-1} at 220 °C. Hexuronic acids were determined colorimetrically as described by Coimbra et al. (1996).

2.4. FT-IR spectra

FT-IR spectra of pectic polysaccharide extracts were acquired with a Golden-gate single reflection ATR in a Bruker IFS-55 at a resolution of 8 cm^{-1} and 128 co-added scans. Spectra were recorded, at least, in triplicate, for each sample, in the absorbance mode from 4000 to 400 cm^{-1}. The FT-IR spectral regions used were 1800–1500 cm^{-1} and 1200–850 cm^{-1}. A total of 169 FT-IR spectra were acquired and 82 of them were used in the calibration set, according to a PCA analysis. The samples considered for the calibration set had a sugar content higher than 40%.

2.5. Data pre-treatment

The spectra were transferred in the JCAMP.DX format (Rutledge & McIntyre, 1992) into the data analysis software package developed by Barros and Rutledge (Barros, 1999). In order to minimise the effect of baseline shifts and other factors that may interfere with the multivariate analysis, the spectra were autoscaled. For the case of OPA (described below), the spectra were normalised between 0 and 1.

2.6. Partial least squares regression

Partial least squares regression (PLS) is a procedure used to model the relationship between a set of predictor variables **X** (n objects $\times$ k variables) and a set of response variables **Y** (n objects $\times$ m responses). In this study there is only one response (DM), thus **Y** has n (objects) $\times$ 1 (response). The PLS regression procedure has the advantage of accepting more variables than objects in the data and of avoiding the problem of collinearity among variables. The classical PLS regression method is a special case of non-linear iterative partial least squares (NIPALS) method, in which the information explained by each dimension is subtracted from the **X** matrix in an iterative process, until all the important variance is extracted (Geladi & Kowalski, 1986)

$$\mathbf{E}_a = \mathbf{E}_{a-1} - \mathbf{t}_a\mathbf{p}_a^{\mathrm{T}}$$

and optionally from the **Y** matrix

$$\mathbf{F}_a = \mathbf{F}_{a-1} - \mathbf{t}_a\mathbf{c}_a^{\mathrm{T}}$$

a is the component to be extracted from **X** and **Y**; **p**, the loadings vector (contribution of each variable to component a) of the **X** matrix; **c**, the loadings vector of the **Y** matrix; **t**, the scores vector (projection of objects in a new coordinate space); and **E** and **F** are the remaining variance after factor extraction.

The PLS regression procedure may be written as:

$$\mathbf{Y} = \mathbf{XB} + \mathbf{F}$$

The regression model was generated by calculating the **B** coefficients matrix which minimised the **F** error matrix. The main difference between this and the other multivariate regression procedures, such as MLR and PCR, concerns the way the **B** matrix is calculated. In the PLS procedure the information present in the **Y** matrix is used in the calculation of the **B** coefficients.

2.7. Durbin–Watson criterion

The Durbin–Watson statistic (DW) is classically used as a criterion to measure the randomness of residuals after a regression (Durbin & Watson, 1950). In this work, it was used as a measure of the structure, the non-randomness or the information content of the loadings, the **B** coefficients and other vectors derived from signals. In this way, DW was used to characterise the 'signal/noise' ratio of the vector

Table 1
Origin and sugar composition of cell wall extracts used in the PCA to select those suitable for the OP-PLS1 calibration curve

Sample number	Origin	HexA (mol%)	Ara (mol%)	Gal (mol%)	Total sugars (mg/g)	DM[a] (mol%)
Samples selected						
1	Olive pulp	64	26	1	525	5
2	Olive pulp	79	18	0	700	8
3	Olive pulp	52	39	2	570	29
4	Olive pulp	68	17	4	729	34
5	Olive pulp	73	23	1	605	36
6	Olive pulp	54	36	4	853	38
7	Olive pulp	74	17	1	752	42
8	Olive pomace	81	12	4	818	43
9	Olive pomace	73	17	6	749	50
10	Pear	57	33	3	735	52
11	Olive pulp	61	23	5	787	52
12	Olive pomace	74	4	13	602	53
13	Pear	71	23	2	733	54
14	Olive pomace	66	22	7	480	55
15	Olive pomace	68	21	6	816	59
16	Pear	84	7	1	720	59
17	Olive pulp	85	10	2	792	59
18	Olive pomace	72	18	6	772	64
19	Pear	91	4	0	737	64
20	Pear	76	17	1	478	72
21	Pear	90	4	0	669	91
Samples not selected						
22	Olive pulp	19	29	2	467	0
23	Olive pulp	20	30	2	506	0
24	Olive pulp	14	29	3	533	0
25	Olive pulp	21	28	3	533	0
26	Olive pulp	26	32	2	555	0
27	Olive pulp	18	30	2	503	0
28	Olive pulp	61	24	1	548	17
29	Citrus	97	0	2	890	35
30	Pear	53	37	3	510	35
31	Olive pomace	74	17	4	508	38
32	Olive pomace	41	40	12	868	48
33	Olive pomace	41	39	12	754	50
34	Olive pulp	17	42	17	469	53
35	Olive pomace	86	4	6	610	53
36	Olive pulp	17	29	2	461	55
37	Olive pulp	21	27	2	446	56
38	Olive pomace	49	34	8	623	57
39	Olive pomace	50	35	9	703	57
40	Olive pulp	18	26	6	571	66

[a] Degree of methylesterification of pectic polysaccharides determined by GLC.

which can be useful, for example, in selecting the optimal number of latent variables (LV) to use in a regression model.

The DW criterion is given by

$$\mathrm{DW} = \frac{\sum_{i=2}^{n} (\delta x_i - \delta x_{i-1})^2}{\sum_{i=1}^{n} (\delta x_i \delta x_i)}$$

where δx_i and δx_{i-1} are the values for the successive points in a series.

If there is a strong correlation between the successive values, DW tends towards zero. If there is a weak correlation (i.e. a 'noisy' series), DW tends towards 2.0 (Durbin & Watson, 1950).[1]

2.8. Outer product analysis

One sometimes wishes to determine the relations that exist between two types of signal such as near IR and mid IR, mid IR and Raman, UV–Visible and NMR, etc. To do this it may be useful to acquire two sets of signals for the same samples and analyse how they vary simultaneously as

[1] URL: http://www.csus.edu/indiv/j/jensena/mgmt105/durbin.htm.

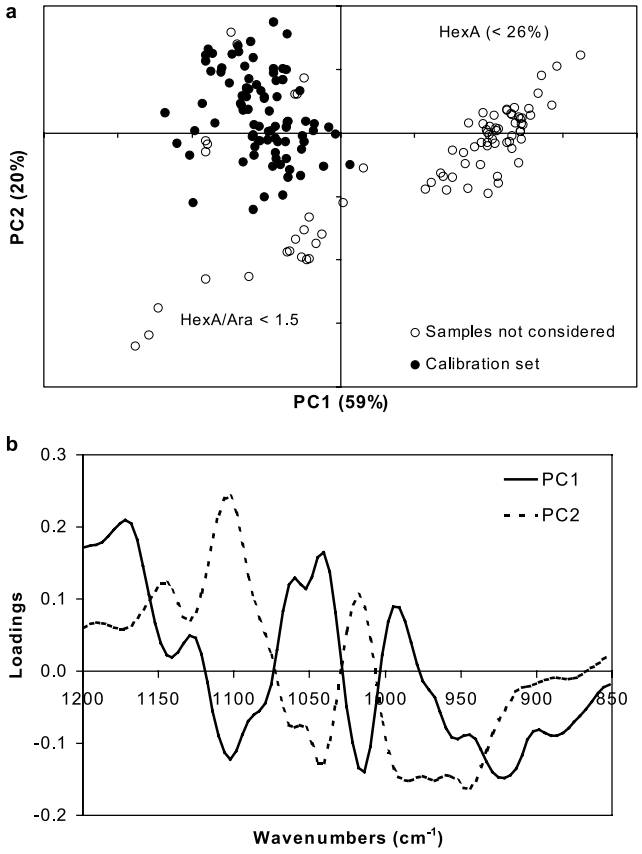

Fig. 2. PCA for identification of the calibration set design. (a) Scores scatter plot (PC1 vs. PC2); (b) Loadings profiles (PC1 and PC2) for characterisation of the calibration set design.

a function of some property, such as concentration. One possibility is to apply statistical techniques to the n outer product matrices (OPMs) calculated, for each of the n samples, by outer product multiplication of their two signal vectors.

The procedure starts by calculating the products of the intensities in the two signal domains for each sample. All the intensities of one domain were multiplied by all intensities in the other domain, resulting in a data matrix containing all possible combinations of the intensities in the two domains (Fig. 1(a)). The outer product of two signal-vectors of lengths r and c for the n samples gave n (r rows by c columns) matrices which were then unfolded to give $n(r \times c)$-long row-vectors (Fig. 1(b)). This procedure corresponds to a mutual weighting of each signal by the other (Barros, 1999; Barros et al., 1997; Rutledge, Barros, & Giangiacomo, 2001):

- (i) if the intensities are high simultaneously in the two domains, the product is higher;
- (ii) if the intensities are low simultaneously in the two domains, the product is lower;
- (iii) if one is high and the other low, the resulting product tends to an intermediate value.

After analysis of the set of $n(r \times c)$-long row-vectors, each vector of calculated statistical parameters was folded back to give a (r rows by c columns) matrix, which may be easily examined to detect the relations between the two domains.

In this work, the two considered domains belonged to mid infrared region: the first one to the region 1800–1500 cm^{-1} (79 variables) and the second one, to the region 1200–850 cm^{-1} (91 variables). The outer product of these two regions gave an **X** matrix with ($79 \times 91 = 7189$ variables). This matrix was then used in PLS1 regression to model the DM. The obtained **b** vector, which established the relationship between the **X** variables and the **y** vector, was therefore composed of 7189 values. In order to facilitate the interpretation of this vector, it was folded back to give a matrix **B** (79×91) which highlighted the links between the variables (wavenumbers) interactions in the two regions.

3. Results and discussion

3.1. Selection of samples

Prior to the construction of a calibration model for the DM of pectic polysaccharide samples, it was necessary to characterise and select, among a wide set of cell wall samples, those that were rich in pectic polysaccharides. Therefore, the design of the calibration set was based on the application of a PCA to the FT-IR spectra of 40 samples (region associated to the variability of pectic polysaccharide samples), in the region between 1200 and 850 cm^{-1}. The sugar composition of the samples used is shown in Table 1 and the scores scatter plot of the entire set of samples is shown in Fig. 2(a). The analysis of this plot, along with the loadings profiles (Fig. 2(b)) allows a better insight on the samples characteristics. Therefore, along PC1 positive axis, one can see a cluster that was characterised by pectic samples where the amount of HexA was lower than 26 mol%. The correspondent loadings profile (PC1—Fig. 2(b)) showed that these samples had indeed a lower HexA content, as observed by the negative bands located at 1100 and 1014 cm^{-1}, and a higher content of neutral sugars, as shown by the positive bands located at 1060 and 1041 cm^{-1} (Coimbra et al., 1998, 1999; Ferreira et al., 2001). The observation of the PC2 loadings showed that this profile was very similar, although with opposite sign, to the PC1 loadings between 1150 and 1000 cm^{-1}. However, for PC2, the elimination of samples was not clear as was with PC1. Hence, a careful examination of those samples was taken and it was found that all of them with a HexA to Ara ratio lower than 1.5 were eliminated from the calibration set (PC2 more negative). Other three samples, although in PC2 positive, were found to be out of the calibration model. Those were samples 29 (commercial sample, from citrus), 31 and 35 (from olive pomace) as, in these samples, the non-esterified carboxyl groups were in the acid form and not in the salt form. This was shown to be required for the

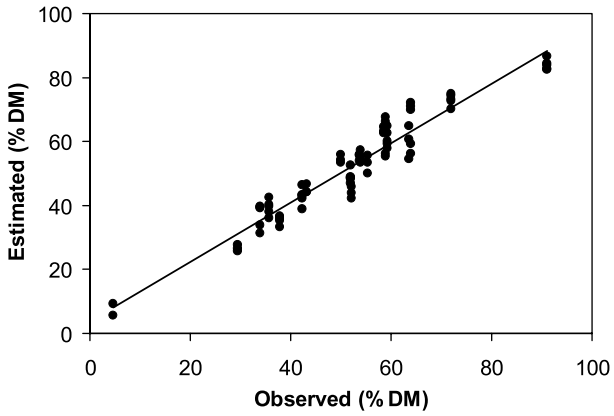

Fig. 3. OP-PLS1 calibration curve plot for determination of the DM of pectic polysaccharides.

determination of the DM by FT-IR. The final calibration model was built using the 21 samples shown in Table 1 and Fig. 2(a), i.e. samples rich in pectic polysaccharides with a HexA amount greater than 52 mol%. They were obtained from olive pulp and pear matrices after extraction using different aqueous solutions followed by dialysis at neutral pH. The extracts used were those from water, diluted acid, CDTA and imidazole. The characterisation of the extracts by sugar analysis indicated the presence of HexA (52–91 mol%), Ara (4–39 mol%) and Gal (1–13 mol%), with a total sugar content within 48–85%, and a DM, estimated by the GLC quantification, that ranged from 5 to 91% (Table 1).

3.2. Estimation of the DM using the 1800–1500 cm^{-1} region

The classical multivariate approach for the determination of the degree of methylesterification in the region 1800–1500 cm^{-1}, using the bands located at 1750 and 1630 cm^{-1}, did not give a regression model for olive pulp and pear pectic polysaccharide extracts with acceptable predictive power. In fact, when a PLS1 procedure was applied to this region, using samples from 1 to 21 of Table 1, one obtained a model with 9 LV with a RMSEP of 14.7% and a R^2 of 0.79. This fact might be related to the presence of esters and carboxylate groups from phenolics in the samples, as shown by the total sugar content (less than 85%), and by the UV absorbing materials present.

In order to relate more precisely the esters and carboxylate groups with the amount of GalA present, the absorbance in the regions 1800–1500 and 1200–850 cm^{-1} regions was combined by means of an OPM.

3.3. Outer product-partial least squares regression

The obtained OPM (82×7189) was used to build a calibration model by means of PLS1 in order to determine the DM (**y** vector) with 82 DM values. Since the size of the **X** matrix was very large, a methodology based on the signal structure was used for the determination of the model dimensionality (Barros, 1999; Barros & Rutledge, 1998. The application of the DW criterion to the regression vector of OP-PLS1 gave an optimal model with 9 LV, with a RMSEP of 7.7% and a coefficient of determination (R^2) of

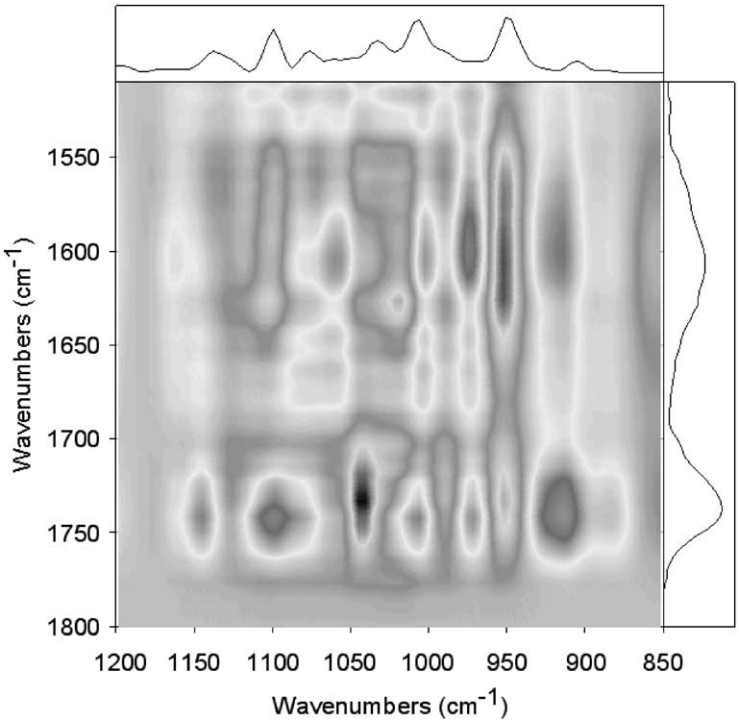

Fig. 4. OP-PLS1 2D **b** vector map for determination of the DM of pectic polysaccharides; the variance profiles of each region are shown in the left and top sides.

Table 2
b vector variables—main relationships as a function of the DM (s, strong; m, medium)

1800–1500 cm^{-1} region	1200–850 cm^{-1} region
Positive relationship	
1746 (ester carbonyl)	913 (s), 971 (m), 1006 (m), 1100 (s), 1145 (m)
1603 (anti-symmetric carboxylate stretching)	913 (m), 975 (s), 1002 (m), 1056 (m)
Negative relationship	
1746 (ester carbonyl)	952 (m), 1044 (s)
1603 (anti-symmetrical carboxylate stretching)	952 (s)
1626	1018 (m), 1100 (m)

0.95. These first 9 LV explained 95% of the total variability presented in the **y** vector. Also, the DW value for the calibration curve residuals was about 0.9, an indication of the randomness residuals (Durbin & Watson, 1950).

From the obtained calibration curve, plotted in Fig. 3, it is possible to observe a linear relationship between the estimated (by OP-PLS-1) and observed (by GLC) DM values. The folded **b** vector is shown in Fig. 4 as a 2D grey level map. This 2D map allowed establishing relationships between the wavenumbers of the two FT-IR regions. The most important links between variables in the two domains, represented by darker spots, are shown in Table 2. The positive relationships are related to those variables links that are correlated to the DM; conversely, the negative ones are anti-correlated to the DM. According to Fig. 4 and Table 2, the most important wavenumbers in the 1800–1500 cm^{-1} region are 1746 cm^{-1} (assigned to carbonyl esters), and 1626 and 1603 cm^{-1} (assigned to anti-symmetric stretching mode of carboxylates). The **b** vector profiles of these wavenumbers in the region 1200–850 cm^{-1} are shown in Fig. 5. The major differences occurred at 1100 and 1018 cm^{-1}, which correlated positively to the ester vibration and negatively to the carboxylate vibrations. These wavenumbers have been assigned to GalA (Coimbra et al., 1998; Ferreira et al., 2001; Kacuráková, Capek, Sasinková, Wellner & Ebringerová, 2000). The ester band was also positively related to the band at 1145 cm^{-1} and negatively related to the band at 1041 cm^{-1}. The carboxylate bands were negatively correlated to 1100 and 1018 cm^{-1} and positively correlated to 1060 cm^{-1}. These observations are in accordance with the relative absorbance of GalA and neutral sugars (Coimbra et al., 1998; Ferreira et al., 2001). The results showed that the absorbance values of the pectic polysaccharides in the region 1200–850 cm^{-1} were positively correlated with the absorbance of the ester groups and anti-correlated with the absorbance of the carboxylate groups in the estimation of their DM.

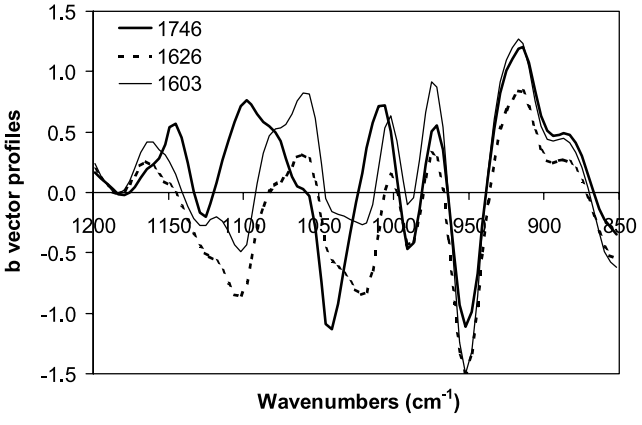

Fig. 5. OP-PLS1 **b** vector profiles at 1746, 1626 and 1603 cm^{-1}.

Table 3
Prediction of DM values for commercial pectins using OP-PLS1 (9 LV)

Sample number	HexA (mol%)	DM (mol%)		CV (%)
		Observed[a]	Predicted	
41	81	93	94	0.3
42	74	67	72	1.7
43	66	26	23	2.3

[a] Degree of methylesterification of pectic polysaccharides determined by GLC.

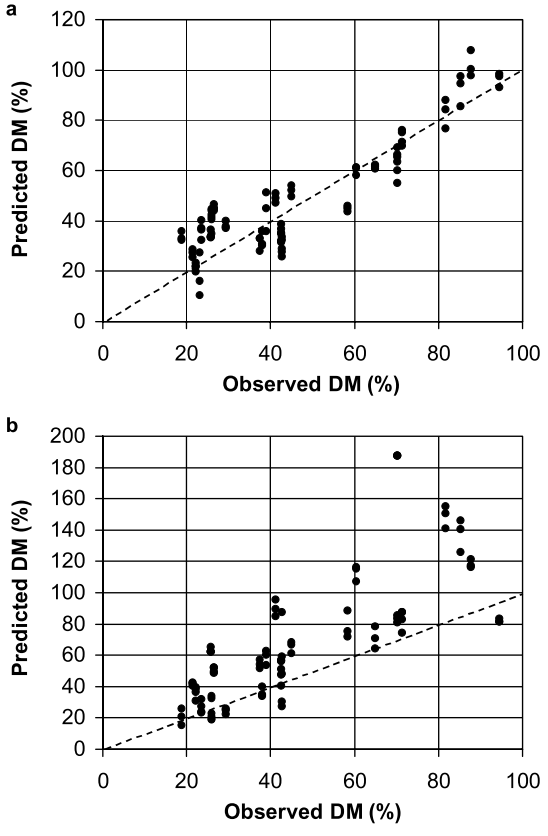

Fig. 6. Prediction of the DM of pectic polysaccharides for an external data set of olive samples (dotted line represents the optimal prediction). (a) Using the proposed OP-PLS1 regression model; (b) using PLS1 in the classical 1800–1500 cm^{-1} region.

Table 4
Prediction of DM values of pectic polysaccharides in non-purified extracts using OP-PLS1 (9 LV)

Sample number	Origin	HexA (mol%)	Ara	Gal	Total sugars (mg/g)	DM (mol%)		CV (%)
						Observed[a]	Predicted	
44	Olive pomace	41	50	4	996	19	34	5
45	Olive pulp	64	24	3	816	22	22	7
46	Olive pomace	61	27	7	768	22	27	6
47	Olive pomace	64	28	4	820	24	37	9
48	Olive pomace	44	41	4	742	26	34	0
49	Olive pomace	61	27	7	768	26	44	3
50	Olive pomace	45	43	4	743	26	39	9
51	Olive pomace	52	22	5	448	27	45	2
52	Olive pomace	61	26	7	580	29	38	4
53	Olive pulp	51	38	5	852	37	31	9
54	Olive pomace	31	50	8	718	38	32	10
55	Olive pomace	53	28	6	311	39	44	18
56	Olive pulp	60	16	12	530	41	49	4
57	Olive pulp	64	21	3	802	43	35	8
58	Olive pulp	51	23	6	695	43	30	11
59	Olive pomace	48	15	5	474	45	52	4
60	Olive pulp	54	18	5	697	58	45	2
61	Olive pulp	51	21	10	496	60	60	3
62	Olive pulp	59	25	7	448	65	61	1
63	Olive pulp	40	28	13	577	70	60	9
64	Olive pulp	34	21	9	287	70	66	5
65	Olive pulp	21	23	3	515	71	73	4
66	Olive pulp	18	25	4	435	82	83	7
67	Olive pulp	21	25	4	416	85	92	7
68	Olive pulp	19	25	3	376	88	102	5
69	Olive pulp	44	16	17	226	95	96	3

[a] Degree of methylesterification of pectic polysaccharides determined by GLC.

3.4. Model validation

The proposed OP-PLS1 calibration model was applied to predict the DM of three commercial citrus pectins of defined DM (93, 67 and 26%). The predicted values of the OP-PLS1 model were very similar to those present in the samples, with a RMSEP of 5.6% and $R^2 = 0.99$. The coefficients of variation (CV) for the spectra repetitions were also very low (Table 3). These results showed that this model could be used to predict the DM of pectic samples due to the low prediction error and to the high linearity observed over a wide range of DM in pure pectic samples.

For the case of complex samples, Fig. 6 shows the predicted values for the DM of different set of olive pectic polysaccharide extracts (Table 4). The observed DM values were obtained by the independent quantification of the total amount of hexuronic acids colorimetrically and the methoxyl groups released as methanol and analysed by GLC. The prediction error obtained (RMSEP) was 21.8% with a coefficient of determination (R^2) of 0.82 (Fig. 6(a)). Despite the observed variability, the OP-PLS1 could predict with some reliability the DM of pectic polysaccharides of samples with low purity and from different sources. It is important to note that using the classical 1800–1500 cm^{-1} region for the prediction of the DM, the obtained prediction error is much higher (RMSEP = 72.5%) with a very low R^2 (0.63). The prediction curve plot for this region (Fig. 6(b)) showed clearly that using solely this region, it is not possible to quantify the DM in complex pectic samples.

4. Concluding remarks

These results showed that the use of the combination of FT-IR spectra in the region 1200–850 cm^{-1} with the ester and carboxylate regions (1800–1500 cm^{-1}) allowed the prediction of the DM of pectic polysaccharides in extracts of olive and pear pulps in the presence of phenolic compounds. The obtained OP-PLS model seems to indicate that one can use this procedure to quantify the amount of the DM of pectic polysaccharides samples.

Acknowledgements

The authors acknowledge the European program FAIR CT97-3053, the Research Unit 62/94 of Química Orgânica, Produtos Naturais e Agro-Alimentares of University of Aveiro. António Barros was supported by a Post-Doctoral grant of FCT BPD/18824/98. Isabel Mafra was supported by a PhD grant of FCT SFRH/BD/1245/2000. Dulcineia Ferreira was supported by a PhD grant PRODEP II

C3/204.015/98. Susana Cardoso was supported by a PhD grant PRODEP III 5.3/N/199.006/00.

References

Barros, A. S. (1999). *Contribution à la sélection et la comparaison de variables caractéristiques*. PhD thesis, Institut National Agronomique, Paris-Grignon, chapter 1.

Barros, A. S., & Rutledge, D. N. (1998). Genetic algorithm applied to the selection of principal components. *Chemometric and Intelligent Laboratory Systems*, *40*, 65–81.

Barros, A. S., Safar, M., Devaux, M. F., Robert, P., Bertrand, D., & Rutledge, D. N. (1997). Relations between mid-infrared and near infrared spectra detected by analysis of variance of an intervariable data matrix. *Applied Spectroscopy*, *51*, 1384–1393.

Bartley, I. M., & Knee, M. (1982). The chemistry of textural changes in fruit during storage. *Food Chemistry*, *9*, 47–58.

Blakeney, A. B., Harris, P. J., Henry, R. J., & Stone, B. A. (1983). A simple and rapid preparation of alditol acetates for monosaccharide analysis. *Carbohydrate Research*, *113*, 291–299.

Brett, C. T., & Waldron, K. W. (1996). *Physiology and biochemistry of plant cell walls*, (2nd ed). London: Chapman & Hall.

Chatjigakis, A. K., Pappas, C., Proxenia, N., Kalantzi, O., Rodis, P., & Polissiou, M. (1998). FT-IR spectroscopic determination of the degree of esterification of cell wall pectins from stored peaches and correlation to textural changes. *Carbohydrate Polymers*, *37*, 395–408.

Claus, R., Nielsen, B. U., & Glahn, P. -E. (1998). Pectin. In S. Dumitriu, *Polysaccharides. Structural diversity and functional versatility* (pp. 37–432). New York: Marcel Dekker.

Coimbra, M. A., Delgadillo, I., Waldron, K. W., & Selvendran, R. R. (1996). Isolation and analysis of cell wall polymers from olive pulp. *Modern methods of plant analysis*, Vol. 17. Berlin: Springer pp. 19–44.

Coimbra, M. A., Barros, A., Barros, M., Rutledge, D. N., & Delgadillo, I. (1998). Multivariate analysis of uronic acid and neutral sugars in whole pectic samples by FT-IR spectroscopy. *Carbohydrate Polymers*, *37*, 241–248.

Coimbra, M. A., Barros, A., Rutledge, D. N., & Delgadillo, I. (1999). FTIR spectroscopy as a tool for the analysis of olive pulp cell-wall polysaccharide extracts. *Carbohydrate Research*, *317*, 145–154.

Cosgrove, D. J. (2001). Wall structure and wall loosening. A look backwards and forwards. *Plant Physiology*, *125*, 131–134.

Durbin, J., & Watson, G. S. (1950). Testing for serial correlations in least squares regression. *Biometrika*, *37*, 409–428.

Femenia, A., Sanches, E. S., Simal, S., & Rosselló, C. (1998). Effects of drying pretreatments on the cell wall composition of grape tissues. *Journal of Agricultural and Food Chemistry*, *46*, 271–276.

Ferreira, D., Barros, A., Coimbra, M. A., & Delgadillo, I. (2001). Use of FT-IR spectroscopy to follow the effect of processing in cell wall polysaccharide extracts of a sun-dried pear. *Carbohydrate Polymers*, *45*, 175–182.

Geladi, P., & Kowalski, B. (1986). Partial least-squares regression: A tutorial. *Analytica Chimica Acta*, *198*, 1–17.

Grant, G. T., Morris, E. R., Rees, D. A., Smith, P. J. C., & Thom, D. (1973). Biological interactions between polysaccharides and divalent cations: The egg-box model. *FEBS Letters*, *32*, 195–198.

Grasdalen, H., Bakoy, O. E., & Larsen, B. (1988). Determination of the degree of esterification and the distribution of methylated and free carboxyl groups in pectins by ^{1}H-N.M.R. spectroscopy. *Carbohydrate Research*, *184*, 183–191.

Haas, U., & Jager, M. (1986). Degree of esterification of pectins determined by photoacoustic near infrared spectroscopy. *Journal of Food Science*, *51*, 1087–1088.

Harris, P. J., Blakeney, A. B., Henry, R. J., & Stone, B. A. (1988). Gas chromatographic determination of the monosaccharide composition of plant cell wall preparations. *Journal of AOAC International*, *71*, 272–275.

Jiménez, A., Rodríguez, R., Fernández-Caro, I., Guillén, R., Fernández-Bolaños, J., & Heredia, A. (2001). Olive fruit cell wall: Degradation of pectic polysaccharides during ripening. *Journal of Agricultural and Food Chemistry*, *49*, 409–415.

Kacuráková, M., Capek, P., Sasinková, V., Wellner, N., & Ebringerová, A. (2000). FT-IR study of plant cell wall model compounds: Pectic polysaccharides and hemicelluloses. *Carbohydrate Polymers*, *43*, 195–203.

Klavons, J. A., & Bennet, R. D. (1986). Determination of methanol using alcohol oxidase and its application to methyl ester content of pectins. *Journal of Agricultural and Food Chemistry*, *34*, 597–599.

Knee, M. (1978). Properties of polygalacturonate and cell cohesion in apple fruit cortical tissue. *Phytochemistry*, *17*, 1257–1260.

Lopes da Silva, J. A., Gonçalves, M. P., & Rao, M. A. (1995). Kinetics and thermal behaviour of the structure formation process in high-methoxyl pectin/sucrose gelation. *International Journal of Biological Macromolecules*, *17*, 25–32.

Mafra, I., Lanza, B., Reis, A., Marsilio, V., Campestre, C., De Angelis, M., & Coimbra, M. A. (2001). Effect of ripening on texture, microstructure and cell wall polysaccharide composition of olive fruit (*Olea europaea*). *Physiologia Plantarum*, *111*, 439–447.

Maness, O. N., Ryan, J. D., & Mort, A. J. (1990). Determination of the degree of methyl esterification of pectins in small samples by selective reduction of esterified galacturonic acid to galactose. *Analytical Biochemistry*, *185*, 346–352.

Martin-Cabrejas, M. A., Waldron, K. W., Selvendran, R. R., Parker, M. L., & Moates, G. K. (1994). Ripening-related changes in the cell walls of Spanish pear (*Pyrus communis*). *Physiologia Plantarum*, *91*, 671–679.

McFeeters, R. F., & Armstrong, S. A. (1984). Measurement of pectin methylation in plant cell walls. *Analytical Biochemistry*, *139*, 212–217.

Mizote, A., Odagiri, H., Toei, K., & Tanaka, K. (1975). Determination of residues of carboxilic acids (mainly galacturonic acid) and their degree of esterification in industrial pectins by colloid titration with cat-floc. *Analyst*, *100*, 822–826.

Paull, R. E., Gross, K., & Qiu, Y. (1999). Changes in papaya cell walls during fruit ripening. *Postharvested Biology and Technology*, *16*, 79–89.

Pfeffer, P. E., Doner, L. W., Hoagland, P. D., & McDonald, G. G. (1981). Molecular interactions with dietary fiber components. Investigation of the possible association of pectin and bile acids. *Journal of Agricultural and Food Chemistry*, *29*, 455–461.

Reintjes, M., Musco, D. D., & Joseph, G. H. (1962). Infrared spectra of some pectic substances. *Journal of Food Science*, *27*, 441–445.

Renard, C. M. G. C., & Jarvis, M. C. (1999). Acetylation and methylation of homogalacturonans 1: Optimisation of the reaction and characterisation of the products. *Carbohydrate Polymers*, *39*, 201–207.

Roberts, K. (2001). How the cell wall acquired a cellular context. *Plant Physiology*, *125*, 127–130.

Rutledge, D. N., & McIntyre, P. (1992). A proposed European implementation of the JCAMP-DX format. *Chemometrics Intelligent Laboratory Systems*, *16*, 95–101.

Rutledge, D. N., Barros, A. S., & Giangiacomo, R. (2001). Interpreting near infrared spectra of solutions by outer product analysis with time domain—NMR. In P. Belton, I. Delgadillo, A. M. Gil & G. A. Webb, *Advances in magnetic resonance in food science* (pp. 179–192). Royal Society of Chemistry, Cambridge.

Selvendran, R. R., March, J. F., & Ring, S. G. (1979). Determination of aldoses and uronic acid content of vegetable fibre. *Analytical Biochemistry*, *96*, 282–292.

Sun, R. C., & Hughes, S. (1999). Fractional isolation and physico-chemical characterization of alkali-soluble polysaccharides from sugar beet pulp. *Carbohydrate Polymers*, *38*, 273–281.

Sun, R. C., Fang, J. M., Goodwin, A., Lawther, J. M., & Bolton, A. J. (1998). Isolation and charcaterization of polysaccahrides from abaca fibre. *Journal of Agricultural and Food Chemistry*, *46*, 2817–2822.

Vierhuis, E., Schols, H. A., Beldman, G., & Voragen, A. G. J. (2000).

Isolation and characterization of cell wall material from olive fruit (*Olea europaea* cv koronieki) at different ripening stages. *Carbohydrate Polymers*, *43*, 11–21.

Voragen, A. G. J., Schols, H. A., & Pilnik, W. (1986). Determination of the degree of methylation and acetylation of pectins by h.p.l.c. *Food Hydrocolloids*, *1*, 65–70.

Voragen, A. G. J., Pilnik, W., Thibault, J. -F., Axelos, M. A. V., & Renard, C. M. G. C. (1995). Pectins. In A. M. Sephen, *Food polysaccharides and their applications* (pp. 287–339). New York: Marcel Dekker.

Waldron, K. W., & Selvendran, R. R. (1990). Composition of the cell walls of different asparagus (*Asparagus officinalis*) tissues. *Physiologia Plantarum*, *80*, 568–575.

Walkinshaw, M. D., & Arnott, S. (1981). Conformations and interactions of pectins. II. Models for junction zones in pectinic acid and calcium pectate gels. *Journal of Molecular Biology*, *153*, 1075–1085.

Wellner, N., Kacuráková, M., Malovíková, A., Wilson, R. H., & Belton, P. S. (1998). FT-IR study of pectate and pectinate gels formed by divalent cations. *Carbohydrate Research*, *308*, 123–131.