JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

ARTICLE

pubs.acs.org/JAFC

Quantification of Phenolic Compounds during Red Winemaking Using FT-MIR Spectroscopy and PLS-Regression

Sandra Fragoso, Laura Aceña, Josep Guasch, Montserrat Mestres,* and Olga Busto

Group of Analytical Chemistry of Wine and Food, Department of Analytical Chemistry and Organic Chemistry Campus Sescelades, Faculty of Oenology, Universitat Rovira i Virgili, 43007 Tarragona, Spain

ABSTRACT: We present a rapid method to quantify phenolic compounds all during the red winemaking process using Fourier transform mid-infrared (FT-MIR) spectroscopy and chemometrics. To get the reference values, we used the usual UV−vis spectroscopy methods, and the compounds studied were evaluated as total phenolic compounds (TPC), total anthocyanins (TA), and condensed tannins (CT). Sampling from five different grape varieties (Merlot, Tempranillo, Syrah, Cariñena, and Cabernet sauvignon), harvested at different ripening states, and monitored over 10 days of vinification produced a total of 600 spectra. These were used to build and validate four different predictive models by partial least-squares (PLS) regression. The spectral regions selected for each model were between 979 and 2989 cm^{-1}, and when selecting the most suitable one in each case, good values of performance parameters were obtained ($R^2_{val} > 0.95$ and RPD > 4.0 for TPC; $R^2_{val} > 0.90$ and RPD > 3.0 for TA; $R^2_{val} < 0.8$ and RPD < 3.0 for CT). Furthermore, also more specific PLS regression models for each phenolic parameter and each grape variety were developed using different regions with results similar to those obtained when dealing with all of the grape varieties. It is concluded that FT-MIR spectroscopy together with multivariate calibration could be a rapid and valuable tool for wineries to carry out the monitoring of phenolic compound extraction during winemaking.

KEYWORDS: FT-MIR spectroscopy, PLS regression, phenolic compounds, anthocyanins, condensed tannins, vinification, maceration

INTRODUCTION

Phenolic compounds play an important role in the organoleptic properties of red wines and also in the aptitude of the wine to age. Among the different phenolic compounds in red wines, the two most important parameters from the oenological point of view are the total anthocyanins and the condensed tannins. Thus, whereas the anthocyanins are responsible for the wine color and are located in grape skins, the condensed tannins are present in skins and seeds and are related to astringency,[1] although it has to be taken into account that tannins are also involved in the development and stabilization of wine color during its aging.[2]

The concentration of the phenolic compounds in grapes is affected by many factors including grape variety,[3] state of grape ripening,[4] climatic and soil conditions (terroir),[5,6] and viticulture techniques.[7] The transference of phenolic compounds from grapes to wines not only depends on the raw material (variety and ripening state) but also on the winemaking strategies, depending on many factors, such as fermentation temperature, skin to juice ratio, maceration time, and enzyme additions.[8,9]

During red winemaking, the phenolic extraction from the skins starts quickly and soon becomes stable, while extraction from the seeds starts later and increases gradually, when the ethanol present in the medium dissolves the waxy layer that coats the seeds and promotes the polyphenol solubilization. Indeed, the phenolic content of grapes is mainly extracted throughout the first 10 days of maceration,[10] and when skin-contact maceration is extended over this period of time, the wine color does not increase significantly, but it becomes more stable.[11] Thus, when producing a young wine, short macerations are preferred to avoid too astringent wines,[12] but for aged wines, long macerations are required to have enough phenolic compound amounts to ensure color stabilization during aging.[13] Therefore, monitoring the polyphenol content in grapes is essential to determine the wine aging ability.

The most usual analytical methods for quantifying phenolic compounds in musts/wines are based on colorimetric measurements of the diluted samples, at 520 nm for anthocyanins and at 280 nm for total phenolic compounds and condensed tannins.[1,14,15] These methods are generally used in wineries because they are simple and precise enough, but they are time-consuming and expensive when many samples must be analyzed, e.g., for process control.

To solve these problems, we propose a rapid method that combines the Fourier transform mid infrared (FT-MIR) spectroscopy and partial least squares (PLS) regression. FT-MIR spectroscopy is a powerful analytical tool that allows fast and simultaneous analysis of several parameters in a large number of samples. Moreover, it implies a minimal sample preparation,[16] being possible to know, at real time, the phenolic content all during winemaking.

However, because of the great amount of information that each FT-MIR spectrum provides, it is necessary to use chemometric tools to extract quality information, both for qualitative and for quantitative analyses. PLS regression is a multivariate calibration method, which is particularly useful when we need to predict a set of dependent variables (i.e., concentrations) from a large set of independent variables (i.e., spectra). FT-MIR

Received: May 18, 2011
Accepted: September 12, 2011
Revised: September 8, 2011
Published: September 12, 2011

dx.doi.org/10.1021/jf201973e | J. Agric. Food Chem. 2011, 59, 10795–10802

spectroscopy is an attractive technology for the food and beverage industry because of its simple, rapid, and nondestructive measurements.[16] It has been applied for routine qualitative analysis and process control in wineries, enabling immediate analysis of raw materials,[17] fermentation monitoring,[18,19] or determination of some of the main compounds in wine such as ethanol, organic acids, or sugars.[20−22] It has also been applied to the quantification of phenolic compounds in white[23] and red wines.[24−27] Related to the monitoring of the extraction of phenolic compounds during winemaking, although there are few studies carried out by using NIR spectroscopy,[28,29] up to now, there is only a previous study where FT-MIR has been used for this purpose, but its authors only considered one grape variety, and the sampling was not suitably performed to ensure representativeness.[29]

The objective of the present article was the application of FT-MIR, combined with PLS regression, for quantifying the concentration of phenolic compounds (evaluated as total phenolic compounds, total anthocyanins, and condensed tannins) in the skin-contact macerations during winemaking.

It was necessary to take into account that the grape variety and ripening stage determine the phenolic content solubilization in the macerated grapes. Thus, to develop robust multivariate models, we chose the samples of the calibration set considering this natural variability by taking five red grape varieties at different stages of ripening. Moreover, the progressive change of sample matrix during maceration due to winemaking further increased the robustness of the models.

The data set reference values were obtained by common UV−vis spectroscopic methods used in cellars: absorbance at 520 nm for total anthocyanins and 280 nm for total phenolic compounds and condensed tannins (after precipitation by methylcellulose).

MATERIALS AND METHODS

Samples. Grapes from five red cultivars (Tempranillo, Merlot, Syrah, Cariñena, and Cabernet sauvignon) were obtained from the experimental vineyard, Mas dels Frares, belonging to the Faculty of Enology (Rovira i Virgili University) in Constantí (Tarragona, Spain) during the 2009 vintage. Every sampling consisted of several clusters, carefully collected as follows: Vines that combine sun and shadow were randomly selected and marked. For each grape variety, one cluster from each marked vine was collected and transported to the laboratory in a 20 kg box for winemaking. The sampling also considered the complete biological cycle of each variety, so it was done weekly for each variety, from the beginning of ripeness until vintage. Thus, the number of samplings was 20: 5 Merlot, 4 Syrah, 4 Tempranillo, 4 Cabernet sauvignon, and 3 Cariñena.

Microvinifications. We proceeded with three individual microvinifications for each variety and ripening stage (i.e., 60 microvinifications).

Thus, the grapes of each sampling were manually destemmed and crushed with a manual crusher fitted with stainless steel rollers. Then, 4 kg of crushed grapes were introduced into plastic containers of 6 L, and 0.48 g of $K_2S_2O_5$ (0.12 g/kg) was added. The alcoholic fermentation was carried out by selected yeasts (AWRI 596), which were inoculated (0.20 g per kg of crushed grape) on the same day of harvest. Fermentation activators were first added 24 h after yeast inoculation (Actiferm 1) and also when the density reached values of 1040−1050 g L^{-1} (Actiferm 2). Must/wine density and temperature were daily measured. Must temperature during fermentation was controlled by keeping the plastic containers in a thermo-controlled room in order to prevent the temperature from exceeding 30 °C.

The alcoholic fermentation finished between the fifth and the seventh day, but the skin-contact maceration was extended until the 10th day. During both periods (fermentation and maceration), the cap of each microvinification was punched down once a day, and after this action, an aliquot of 20 mL was taken to be analyzed.

Reagents and Standards. The standards of malvidin-3-glucoside (purity ≥90%) and (+)-catechin (purity ≥96%) were purchased from Fluka (Madrid, Spain). Methyl cellulose (M-0387) was supplied by Sigma Aldrich (Madrid, Spain). Gallic acid monohydrate (99.5%) and the rest of chemicals used for the study were of analytical-reagent grade and provided by Scharlab (Barcelona, Spain).

Selected yeasts (AWRI 596) were purchased from Agrovin (Ciudad Real, Spain). The fermentation activators (Actiferm 1 and 2) and the potassium bisulphite (purity ≥95%) were supplied by Martin Vialatte Oenologie (Epernay, France).

Instrumentation. Centrifugation of must and wine samples was carried out with a Hettich Universal 32 R centrifuge (Tuttlingen, Germany). The absorbance was measured using a Thermo Spectronic UV−vis spectrophotometer Model Helios γ (Thermo Electron Corporation, Cambridge, UK). All spectra of the musts/wines were collected using a FT-MIR 470 Nexus (Thermo Nicolet, USA), equipped with a Globar IR source, a KBr beam splitter, a ZnSe liquid transmission flow cell of 0.025 mm of path length (0.004 mL of liquid sample volume), and a Deuterate Triglycine Sulfate detector (DTGS). The instrument was connected to a TDI Bacchus (Gavà, Barcelona, Spain) autosampler with an online and automatic system of sample filtration (stainless steel filter of food-grade with a porous size of 50 μm, Teflon coated), which pumped a volume of 10 mL of each individual sample to the cell for spectra acquisition. The software package OMNIC, version 6.2, from Thermo Nicolet was used for spectra acquisition. The software used for data analysis and calibration was the Unscrambler package (version 9.0, CAMO ASA, Norway).

Midinfrared Scanning. Prior to MIR scanning, musts and wines were centrifugated at 8000 rpm for 10 min. All spectra were averaged from 32 scans (each spectrum acquisition takes only 30 s per sample) and collected in absorbance mode, at 4 cm^{-1} resolution, in the wavenumber range of 979−2989 cm^{-1}. To remove the environmental interferences (water vapor together with CO_2) and the possible instrumental drift over time, every 10 h the instrument collected a background spectrum that was automatically subtracted from each sample spectrum, yielding the spectrum of the compounds being analyzed. Furthermore, since water bands dominate the spectrum of liquid samples, prior to the analysis of the musts and wines, a blank of distilled water was acquired and the spectrum obtained also subtracted.

FT-MIR instruments allow one to achieve high levels of signal stability and reproducibility over time, and it is not necessary to preprocess spectral data, so we decided to work with raw spectral data.

Reference Analytical Measurements. Total phenolic compounds (mg L^{-1} gallic acid), total anthocyanins (mg L^{-1} malvidin-3-glucoside), and condensed tannins (mg L^{-1} catechin) were the oenological parameters used to monitor the phenolic extraction during winemaking. These parameters were determined by UV−vis spectroscopy as follows.

Analysis of Total Phenolic Compounds. The content of the total phenolic compounds (TPC) was determined on musts/wines by measuring the absorbance at 280 nm after dilution 1:50 with distilled water[1] in a 1 cm quartz cuvette. The quantification was carried out by the external standard method using a calibration line built with gallic acid as the standard at 6 different concentrations in the range of 2.2−18.0 mg L^{-1}.

Analysis of Total Anthocyanins. The content of the total anthocyanins (TA) was determined on musts/wines by measuring the absorbance at 520 nm after dilution 1:25 with 0.1 M HCl to get pH ∼1.0 in a 1 cm plastic cuvette.[14] The quantification was carried out by the external standard method using a calibration line built with malvidin-3-glucoside as the standard at 6 different concentrations in the range of 2.4−20.0 mg L^{-1}.

Table 1. Descriptive Statistics for Total Phenolic Compounds, Total Anthocyanins, and Condensed Tannins of Training and Test Sets[a]

	training set (samples = 400)				test set (samples = 200)			
phenolic parameter	min	max	mean	SD	min	max	mean	SD
TPC	250	1189	778	233	250	1172	746	252
TA	98	577	349	126	99	569	333	131
CT	80	742	389	152	91	761	388	150

[a] TPC, total phenolic compounds (mg gallic acid L^{-1}); TA, total anthocyanins (mg malvidin-3-glucoside L^{-1}); CT, condensed tannins (mg catechin L^{-1}); min, the minimum value; max, the maximum value; SD, standard deviation.

Table 2. Descriptive Statistics for the Total Phenolic Compounds (TPC), Total Anthocyanins (TA), and Condensed Tannins (CT) for Each Varietal Must/Wine[a]

cultivar	samples	phenolic parameter	min	max	mean	SD
Syrah	120	TPC	391	1122	921	183
		TA	238	566	461	74
		CT	80	521	297	96
Merlot	150	TPC	378	1189	896	216
		TA	192	577	444	93
		CT	84	502	294	95
Tempranillo	120	TPC	278	1010	801	192
		TA	119	384	300	57
		CT	139	761	513	128
Cariñena	90	TPC	250	582	476	89
		TA	98	208	163	27
		CT	100	664	440	143
Cabernet sauvignon	120	TPC	252	904	638	165
		TA	99	377	282	70
		CT	91	690	437	150

[a] TPC, total phenolic compounds (mg gallic acid L^{-1}); TA, total anthocyanins (mg malvidin-3-glucoside L^{-1}); CT, condensed tannins (mg catechin L^{-1}); min, the minimum value; max, the maximum value; SD, standard deviation.

Analysis of Condensed Tannins. The content of the condensed tannins (CT) was determined on musts/wines by the indirect method of precipitation with methyl-cellulose precipitable (MCP),[15] measuring the absorbance at 280 nm in a 1 cm quartz cuvette. The quantification was carried out by the external standard method using a calibration line built with (+)-catechin as the standard at 6 different concentrations in the range of 19.2 to 76.9 mg L^{-1}.

Development and Validation of FT-MIR Models. The data set was composed of a total of 600 samples, corresponding to musts/wines from 60 vinifications (15 Merlot, 12 Syrah, 12 Tempranillo, 12 Cabernet sauvignon, and 9 Cariñena) analyzed daily during 10 days of skin-contact maceration.

First, a descriptive analysis of the data was performed by principal components analysis (PCA). Then, quantitative analysis was performed with partial least-squares (PLS) regression by relating the sample spectra and the reference values.[30] Before proceeding with the PLS, the FT-MIR spectra data were automatically processed by mean centering.

The whole data set was split using the Kennard−Stone algorithm,[31] into a training set (N_C = 400 samples), to build and validate the model, and a test set (N_T = 200 samples), to evaluate the prediction ability of the model.

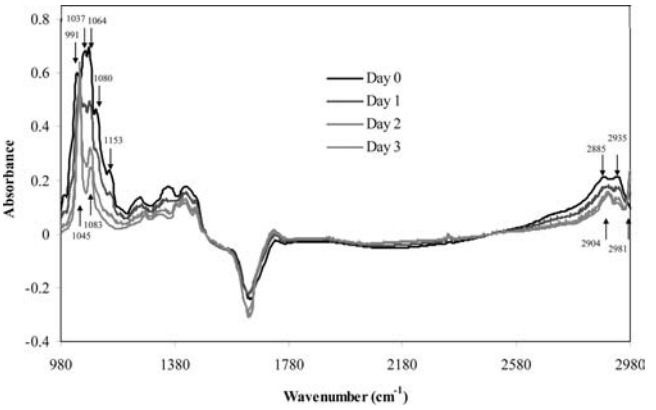

Figure 1. FT-MIR spectra belonging to the fermentation of a must sample daily monitored during the first 3 days. Main glucose peaks (↓): 991, 1037, 1064, 1080, 1103, 1153, 2885, and 2935 cm^{-1}. Main ethanol peaks (↑): 1045, 1083, 2904, and 2981 cm^{-1}.

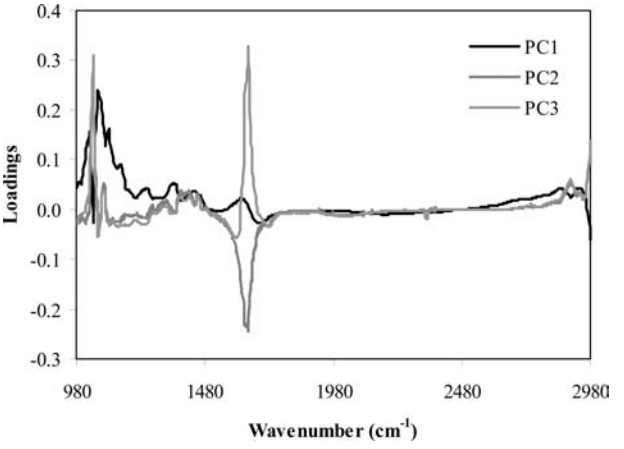

Figure 2. Loadings for the first three PC's of the full-range (979−2989 cm^{-1}) spectral data.

The Kennard-Stone algorithm proposed a sequential method that should cover the experimental region uniformly. The procedure consists of selecting the first two samples with the largest Euclidean distance for the training set, considering the values of the spectral variables. Then, from the rest of all possible samples, the one that is most distant from those already selected was chosen, and it was included in the training set. This selection process continued until the desired number of samples for the training set was reached. The remaining samples in the data set are used to create the validation set. The Kennard-Stone algorithm ensures

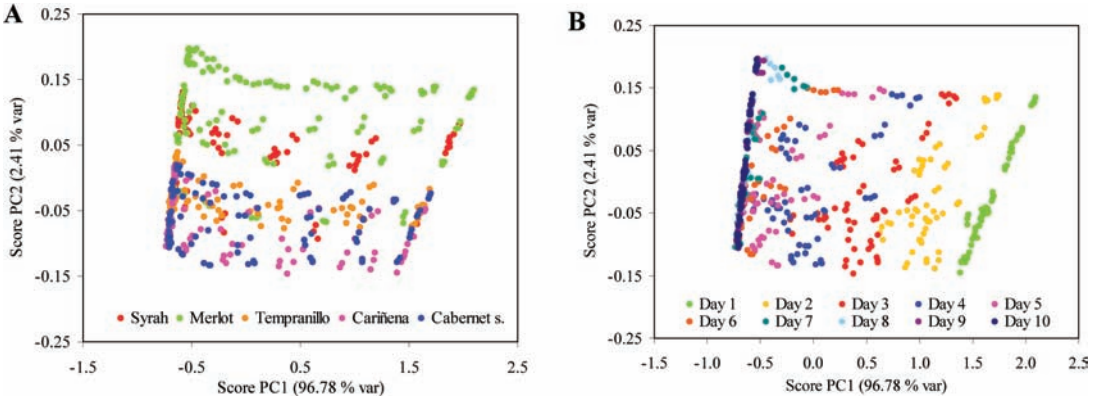

Figure 3. Score plots (PC1−PC2) obtained by applying PCA to the FT-MIR spectra in the region 979−1477 cm^{-1} (A) when considering the cultivars and (B) when considering the vinification stage.

Table 3. Analytical Performance Parameters of the Multivariate Calibration Models Built by PLS Regression Using Different Spectral Regions of Musts/Wines[a]

dependent variable	model	region (cm^{-1})	PLS factors	RMSEC	RMSEC (%)	R^2_{cal}	RMSEP	RMSEP (%)	R^2_{val}	$RMSEP_{cor}$ (%)	RPD
TPC	full-range	979−2989	9	48.6	6.2	0.956	52.0	7.0	0.958	6.9	4.9
	fingerprint	979−1477	10	39.2	5.0	0.972	42.4	5.7	0.972	5.6	6.0
	main phenolic region	1133−1457	9	39.4	5.1	0.971	43.0	5.8	0.971	5.7	5.9
	selected region	1168−1457	9	40.4	5.2	0.970	45.7	6.1	0.967	6.1	5.5
TA	full-range	979−2989	10	31.9	9.1	0.936	39.2	11.8	0.912	11.8	3.3
	fingerprint	979−1477	12	20.6	5.9	0.973	25.1	7.5	0.963	7.5	5.2
	main phenolic region	1133−1457	9	25.1	7.2	0.960	28.6	8.6	0.952	8.6	4.6
	selected region	1168−1457	9	25.0	7.2	0.960	29.4	8.8	0.950	8.8	4.5
CT	full-range	979−2989	15	57.7	14.8	0.858	70.6	18.2	0.777	*b*	2.1
	fingerprint	979−1477	13	64.0	16.4	0.822	66.9	17.4	0.799	*b*	2.2
	main phenolic region	1133−1457	15	62.5	15.9	0.830	71.2	18.4	0.772	*b*	2.1
	selected region	1168−1457	13	63.8	16.3	0.824	77.8	20.1	0.728	*b*	1.9

[a] RMSEC, root mean square error of calibration; R^2_{cal}, coefficient of determination of calibration; RMSEP, root mean square error of prediction; R^2_{val}, coefficient of determination of validation; $RMSEP_{cor}$, bias-corrected RMSEP; RPD, residual predictive deviation. TPC, total phenolic compounds (mg gallic acid L^{-1}); TA, total anthocyanins (mg malvidin-3-glucoside L^{-1}); CT, condensed tannins (mg catechin L^{-1}). [b] It was impossible to calculate this value as the correction of the RMSEP % due to the negative values inside the square roots.

that the validation samples are in the experimental space of the training set, minimizing the extrapolation when the validation samples are predicted.

A critical step in model building is the selection of the optimum number of factors to ensure the prediction ability but also to avoid overfitting. In this study, the number of factors was determined by means of leave-one-out cross-validation[32] considering the lowest root-mean-square error of cross validation (RMSECV).[33]

Related to the calibration step, we evaluated the model fit to the data with the root-mean-square error of calibration (RMSEC) expressed as a percentage (RMSEC %),[33] which can be defined as the mean error of the model.

To test the predictive accuracy of the calibration models built, the external test set was used to determine the root-mean-square error of prediction (RMSEP) expressed as a percentage (RMSEP %).[33]

Additionally, to standardize the predictive accuracy, the residual predictive deviation (RPD) was calculated for each model as the ratio between the standard deviation (SD) of the TPC, TA, and CT values of the validation samples (test set) and the RMSEP results.[34] An RPD value greater than 3.0 indicates a suitable calibration model for prediction purposes.[17,35]

Estimation of the True Prediction Error. The reference values used in the development of the models lead to an intrinsic associated uncertainty, so they are obtained with a standard error. Therefore, the validation of multivariate calibration models using these reference values provides a systematic overestimation of the true prediction error, which is the so-called apparent prediction error. To yield a more realistic estimation of the true prediction error of the models, we used the simple correction procedure proposed by Faber et al.[36] (eq 1)

$$MSEP_{cor} = MSEP_{app} - \sigma^2 \quad (1)$$

where $MSEP_{cor}$ is the bias-corrected MSEP, $MSEP_{app}$ is the apparent MSEP (i.e., the value obtained with the test set), and σ^2 is the variance of the measurement error in the reference values.

RESULTS AND DISCUSSION

Reference Values. Table 1 summarizes the reference values of the contents of the total phenolic compounds (TPC), total anthocyanins (TA), and condensed tannins (CT) for both data sets determined by using the reference analytical measurements. As shown, a wide range of composition was covered due to the changes that occurred during winemaking.

Table 2 summarizes the concentration of the phenolic compounds evaluated considering each variety separately. It can be seen, for example, that Cariñena, which is characterized for its

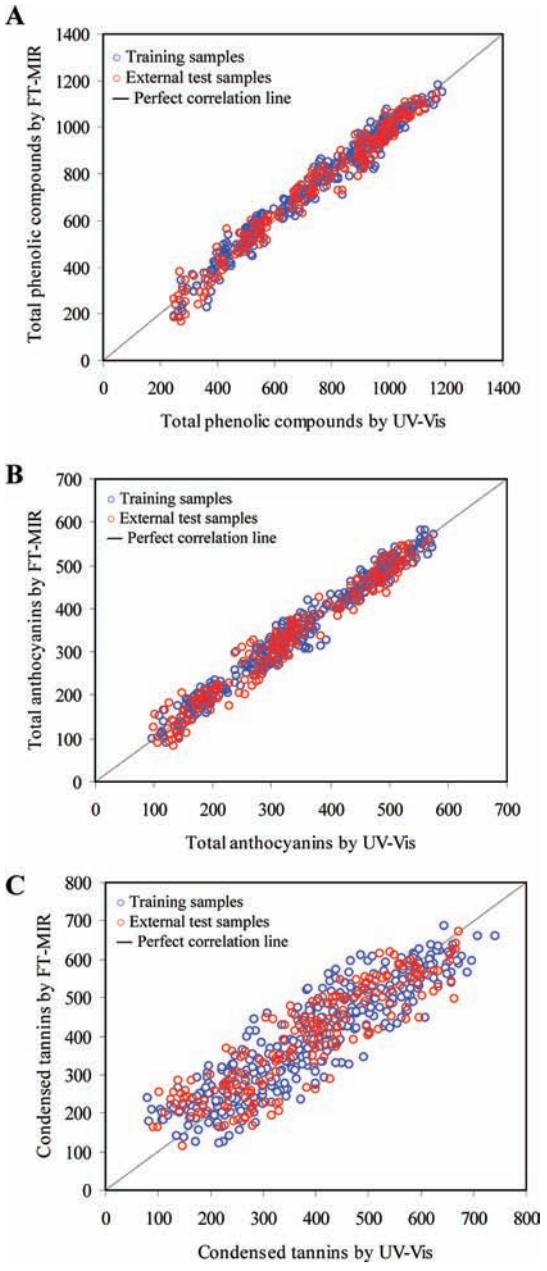

Figure 4. Correlation plot for the prediction of (A) total phenolic compounds (expressed as mg L^{-1} of gallic acid), (B) total anthocyanins (expressed as mg L^{-1} of malvidin-3-glucoside), and (C) condensed tannins (expressed as mg L^{-1} of catechin) using the model for the selected region (979−1477 cm^{-1}).

low anthocyanic content, presented the lowest TA value, while Merlot and Syrah varieties, which are known for their intense red color, presented the highest TA values. The uncommon low TA value found for a high colored variety such as Cabernet sauvignon was due to the fact that the field zone where this variety grows is very wet, and this high humidity makes a proper ripening process difficult. However, as shown, Merlot and Syrah provided the highest standard deviation (SD) values. This is because these grape varieties are characterized for showing a progressive accumulation of TPC and TA during long ripening periods.

Spectra of Musts and Wines. As expected, the dominating absorption bands belonged to the major components of grape, i.e, to organic acids, sugars, and also to ethanol generated during the fermentation of sugars. All these bands masked the characteristic IR vibrations of phenolic compounds.[19,37−39]

Moreover, it has to be pointed out that, since during the fermentation the contents of these compounds changed, the spectra obtained were different depending on the fermentation stage (Figure 1). These differences could be observed mainly in the 979−1477 cm^{-1} region where the bands were mainly generated by the contribution of the C−O stretch belonging to the primary alcohol of ethanol[29,40] and also of the C−O valence vibrations and C−O−C stretching vibrations belonging to carbohydrates, including fructose and glucose.[19,39,41,42] However, although in less extent, the differences can also be detected in the 2800−2960 cm^{-1} region where the bands were due to the C−H stretch of CH_3 and CH_2 from ethanol.[29,40] However, since the first region contained much more information about the organic compounds, it was called the fingerprint region.

The negative absorption band at 1500−1740 cm^{-1}, assigned to water, was due to the automatic subtraction of the aqueous blank absorbance that the instrument carried out prior to the sample analysis.[37] The peak at 2341 cm^{-1} corresponds to the CO_2 released during fermentation. The water and CO_2 absorption regions do not contain useful information, so they are isolated from the spectral data for calibrating purposes.

Because of the great number of absorption bands obtained, it was not easy to find the spectral regions associated with the absorptions of phenolic compounds. However, there are some studies related to the identification of the spectral regions for the wine tannins,[27] and we also carried out some studies with the spectra from grape extracts enriched with different phenolics of different chemical structure in a previous study.[33] These studies revealed that, when dealing with phenolic compounds, the best spectral regions for calibration purposes ranged from 1168 to 1457 cm^{-1} or from 1133 to 1457 cm^{-1}.

Principal Components Analysis (PCA). We applied a PCA to the full range (979−2989 cm^{-1}) spectral data set. The results showed that three principal components (PC's) explained the 99.68% of the spectral variation of the samples (first PC 96.78%; second PC 2.41%; and third PC 0.50%).

Figure 2 shows the loadings for the first three PCs, that is, the influence of each wavelength on the variance along this region. The main regions with the highest influence on the spectral variance coincided with the regions where the signal is higher: 979−1477 cm^{-1} and 1500−1740 cm^{-1}. However, since the latest bands were due to water absorption, we could conclude that the region 979−1477 cm^{-1} contained almost all the information that characterizes the samples. Indeed, a PCA analysis using this region showed that the first PC explained the 98.76% of variance and that the first three PCs explained the 99.98%.

Moreover, when examining the score plots in the area defined by the first two principal components, it was observed that the samples did not form clusters or groups, but they were distributed all along the space (Figure 3A,B), so global models with total calibration samples could be built. Therefore, we could check that the fingerprint region was very useful for calibration purposes.

Quantitative Analysis: PLS Models. The PLS regression method was used to build four different multivariate calibration models between the FT-MIR spectra (using different wavenumber ranges selections) and the reference values of the TPC, TA, and CT phenolic concentrations, respectively. The four different wavenumber range selections corresponded to the full spectrum

Table 4. Results of PLS Models for Phenolic Compounds in Varietal Musts/Wines Using the Best Calibration Regions[a]

cultivar	dependent variable	samples	region (cm^{-1})	PLS factors	RMSEC	RMSEC (%)	R^2_{cal}	RMSECV (%)	R^2_{val}	RPD
Merlot	TPC	150	1133−1457	3	27.8	3.1	0.983	3.2	0.982	7.6
	TA	150	1060−1457	6	19.5	4.4	0.956	4.8	0.946	4.3
	CT	143	1060−1457	12	42.3	14.5	0.799	17.3	0.715	1.9
Tempranillo	TPC	120	1168−1457	6	26.8	3.3	0.980	3.7	0.975	6.4
	TA	120	1133−1457	8	11.8	3.9	0.957	4.4	0.946	4.3
	CT	120	979−1477	9	41.6	8.1	0.893	9.6	0.850	2.6
Syrah	TPC	120	979−1477	5	24.3	2.6	0.982	2.8	0.980	7.0
	TA	120	1060−1457	7	16.0	3.5	0.953	3.9	0.941	4.1
	CT	114	1133−1457	11	40.2	13.4	0.815	16.3	0.730	1.9
Cariñena	TPC	90	1168−1457	4	15.5	3.2	0.970	3.7	0.961	5.1
	TA	90	1168−1457	9	6.0	3.7	0.951	4.5	0.927	3.7
	CT	90	979−1477	9	44.0	10.0	0.905	12.5	0.852	2.6
Cabernet sauvignon	TPC	118	1133−1457	2	22.4	3.5	0.981	3.6	0.980	7.1
	TA	117	1133−1457	6	11.2	4.0	0.973	4.4	0.966	5.4
	CT	120	1133−1457	6	45.4	10.4	0.908	11.4	0.889	3.0

[a] RMSEC, root mean square error of calibration; R^2_{cal}, coefficient of determination of calibration; RMSECV, root mean square error of cross-validation; R^2_{val}, coefficient of determination of validation; RPD, residual predictive deviation. TPC, total phenolic compounds (mg gallic acid L^{-1}); TA, total anthocyanins (mg malvidin-3-glucoside L^{-1}); CT, condensed tannins (mg catechin L^{-1}).

(979−2989 cm^{-1}), the fingerprint region (979−1477 cm^{-1}), and the two regions revealed by the previous enrichment experiment:[33] 1133−1457 cm^{-1} (called the main phenolic region) and 1168−1457 cm^{-1} (called the selected region). The results of each model are reported in Table 3, and an example of the prediction correlation plots obtained when using the fingerprint region is shown in Figure 4.

TPC Quantification. Among the different models, the ones developed when using the fingerprint and the main phenolic regions presented the best TPC prediction results with highest values of R^2_{val} (0.972 and 0.971) and RPD (6.0 and 5.9) and also with lowest RMSEP % values (5.7% and 5.8%), respectively. However, the number of factors to describe the spectral variance when working with the main phenolic region was slightly lower. All these results indicated that when some bands inside MIR spectra were selected more satisfactory prediction results were obtained than when the full spectrum region was used. Moreover, all the models obtained in this study presented lower prediction error values than the ones provided by other researchers on the quantification of total phenolic compounds in red wine fermentations by MIR spectroscopy.[29]

To verify the absence of bias between the reference values and the predicted values calculated by each model, we proceed with the joint confidence region test. We concluded that the results of the four FT-MIR models were unbiased because the slope and intercept of the four regression lines were not significantly different (with a significance level of 0.05) from 1 and 0, respectively.

TA Quantification. After applying the joint confidence region test to the data of this parameter, we detected a significant bias when using the full range of spectrum. Indeed the models obtained working in this region presented the lowest R^2_{val} (0.912) and RPD (3.3) values and the highest RMSEP % (11.8%). On the contrary, the model developed using the fingerprint region showed the highest R^2_{val} (0.963) and RPD (5.2) and the lowest RMSEP (7.5%). The models built with the other two regions considered also provided satisfactory prediction results, with R^2_{val} = 0.952 and 0.950, RPD = 4.6 and 4.5, and RMSEP % = 8.6% and 8.8%, respectively.

In comparison with those obtained for the TPC results, the RMSEP values obtained for TA predictions were higher. This different behavior could be due to the chemical similarity of anthocyanins with other phenolic compounds, which makes it difficult to find their total relevant spectral region. However, whatever the region selected, whereas for TA the $RMSEP_{cor}$ values obtained were equal to the RMSEP, these values are slightly different for TPC. This different trend is due to the fact that the reference method to quantify TA is more precise and presents a lower standard deviation. In any case, because all the models obtained for both parameters provided $RMSEP_{cor}$ lower than 10%, we considered that the uncertainty values are suitable for predictive purposes.

CT Quantification. The calibration models obtained for CT were the least robust of all. In fact, already in the joint confidence test, we checked that the four models were significantly biased, so they did not provide comparable results with the reference values. Moreover, the models obtained showed higher values of RMSEP % than those for TPC and TA parameters whatever the spectral region selected, so the predictive ability was not good enough even when we considered the best model, which was obtained with the fingerprint region. This was mainly attributed to the high relative standard deviation (>20%) of the chosen reference analytical method. This lack of accuracy of the reference method can be explained because it is an indirect quantification method and also because wine tannins are structurally very complex and diverse, which implies a poorly repetitive interaction with the methyl cellulose. Therefore, in future studies the reference method should be replaced.

Models for Individual Cultivars. Separate calibrations for each cultivar and spectral region were performed and the models cross-validated to evaluate the predictive ability of the calibration equations. Table 4 shows the models built with the spectral regions that provided the best results for each phenolic parameter and, as can be seen, every variety follows a specific trend.

In comparison with those of the global calibrations, these models showed better RMSEC % values, mainly for TPC and TA (lower than 5.0%). Moreover, less PLS factors were required,

which is what is in accordance with the results obtained by other researchers.[29]

From all the results presented, we conclude that the FT-MIR spectrometry combined with multivariate calibration enables an easy and reliable measurement of the content of phenolic compounds at the same time, regardless of the stage of skin-contact maceration. Indeed, although the sample matrixes were constantly change during winemaking, the phenolic compounds studied can be well predicted avoiding sample preparation or spectral pretreatment. This implies a very short analytical time, so the methodology proposed becomes an invaluable tool when a large number of samples have to be analyzed.

This allows one to determine the suitability of a grape to get a specific style of wine because the quick establishment of the extraction ending moment can prevent negative sensory attributes in red wines.

Moreover, the FT-MIR instruments could be applied to predict phenolic compounds on specific cultivars. This would allow one to design specific models to fulfill each winery's requirements.

AUTHOR INFORMATION

Corresponding Author

*Tel: +34977558494. Fax: +34977558446. E-mail: montserrat.mestres@urv.cat.

Funding Sources

We acknowledge the Spanish Ministry of Science and Technology (projects AGL2007-61550 and AGL2010-19688) for the financial support.

ACKNOWLEDGMENT

We are grateful to Joe Luis Villa Medina for his valuable help.

REFERENCES

(1) Ribéreau-Gayon, P.; Glories, Y.; Maujean, A.; Dubourdieu, D. *Handbook of Enology, Vol 2. The 395 Chemistry of Wine, Stabilization and Treatments*; John Wiley & Sons: Chichester, West Sussex, 1999.

(2) Cheynier, V.; Duenas-Paton, M.; Salas, E.; Maury, C.; Souquet, J.; Sarni-Manchado, P.; Fulcrand, H. Structure and properties of wine pigments and tannins. *Am. J. Enol. Vitic.* **2006**, *57*, 298–305.

(3) Romero-Cascales, I.; Ortega-Regules, A.; Lopez-Roca, J. M.; Fernandez-Fernandez, J. I.; Gomez-Plaza, E. Differences in anthocyanin extractability from grapes to wines according to variety. *Am. J. Enol. Vitic.* **2005**, *56*, 212–219.

(4) Pérez-Magariño, S.; González-San José, M. L. Polyphenols and colour variability of red wines made from grapes harvested at different ripeness grade. *Food Chem.* **2006**, *96*, 197–208.

(5) Rodríguez Montealegre, R.; Romero Peces, R.; Chacón Vozmediano, J. L.; Martínez Gascueña, J.; García Romero, E. Phenolic compounds in skins and seeds of ten grape *Vitis vinifera* varieties grown in a warm climate. *J. Food Compos. Anal.* **2006**, *19*, 687–693.

(6) De Andres-De Prado, R.; Yuste-Rojas, M.; Sort, X.; Andres-Lacueva, C.; Torres, M.; Lamuela-Raventos, R. Effect of soil type on wines produced from *Vitis vinifera* L. Cv. Grenache in commercial vineyards. *J. Agric. Food Chem.* **2007**, *55*, 779–786.

(7) Perez-Lamela, C.; Garcia-Falcon, M.; Simal-Gandara, J.; Orriols-Fernandez, I. Influence of grape variety, vine system and enological treatments on the colour stability of young red wines. *Food Chem.* **2007**, *101*, 601–606.

(8) Gomez-Miguez, M.; Heredia, F. Effect of the maceration technique on the relationships between anthocyanin composition and objective color of Syrah wines. *J. Agric. Food Chem.* **2004**, *52*, 5117–5123.

(9) Sacchi, K. L.; Bisson, L. F.; Adams, D. O. A review of the effect of winemaking techniques on phenolic extraction in red wines. *Am. J. Enol. Vitic.* **2005**, *56*, 197–206.

(10) González-Manzano, S.; Rivas-Gonzalo, J. C.; Santos-Buelga, C. Extraction of flavan-3-ols from grape seed and skin into wine using simulated maceration. *Anal. Chim. Acta* **2004**, *513*, 283–289.

(11) Haslam, E. Taste, Bitterness and Astringency. In *Practical Polyphenolics. From Structure to Molecular Recognition and Physiological Action*; Haslam, E., Ed.;Cambridge University Press: Cambridge, U.K., 1998.

(12) Samappito, S.; Butkhup, L. Effect of skin contact treatments on the aroma profile and chemical components of mulberry (*Morus alba* Linn.) wines. *Afr. J. Food Sci.* **2010**, *4*, 52–61.

(13) Llaudy, M.d.C.; Canals, R.; Canals, J. M.; Zamora, F. Influence of ripening stage and maceration length on the contribution of grape skins, seeds and stems to phenolic composition and astringency in wine-simulated macerations. *Eur. Food Res. Technol.* **2008**, *226*, 337–344.

(14) Blouin, J. *Techniques d'analyses des moûts et des vins*; Dujardin-Salleron: Paris, 1992.

(15) Mercurio, M. D.; Dambergs, R. G.; Herderich, M. J.; Smith, P. A. High throughput analysis of red wine and grape phenolics-Adaptation and validation of methyl cellulose precipitable tannin assay and modified Somers color assay to a rapid 96 well plate format. *J. Agric. Food Chem.* **2007**, *55*, 4651–4657.

(16) Sun, D. *Infrared Spectroscopy for Food Quality Analysis and Control*; Academic Press: Amsterdam, The Netherlands, 2009.

(17) Versari, A.; Parpinello, G. P.; Mattioli, A. U.; Galassi, S. Determination of grape quality at harvest using Fourier-Transform Mid-Infrared spectroscopy and multivariate analysis. *Am. J. Enol. Vitic.* **2008**, *59*, 317–322.

(18) Urtubia, A.; Pérez-Correa, J. R.; Meurens, M.; Agosin, E. Monitoring large scale wine fermentations with infrared spectroscopy. *Talanta* **2004**, *64*, 778–784.

(19) Mazarevica, G.; Diewok, J.; Baena, J. R.; Rosenberg, E.; Lendl, B. On-line fermentation monitoring by Mid-infrared spectroscopy. *Appl. Spectrosc.* **2004**, *58*, 804–810.

(20) Vonach, R.; Lendl, B.; Kellner, R. High-performance liquid chromatography with real-time Fourier-transform infrared detection for the determination of carbohydrates, alcohols and organic acids in wines. *J. Chromatogr., A* **1998**, *824*, 159–167.

(21) Moreira, J. L.; Santos, L. Analysis of organic acids in wines by Fourier-transform infrared spectroscopy. *Anal. Bioanal. Chem.* **2005**, *382*, 421–425.

(22) Patz, C. D.; Blieke, A.; Ristow, R.; Dietrich, H. Application of FT-MIR spectrometry in wine analysis. *Anal. Chim. Acta* **2004**, *513*, 81–89.

(23) Gorinstein, S.; Weisz, M.; Zemser, M.; Tilis, K.; Stiller, A.; Flam, I.; Gat, Y. Spectroscopic analysis of polyphenols in white wines. *J. Ferment. Bioeng.* **1993**, *75*, 115–120.

(24) Soriano, A.; Pérez-Juan, P. M.; Vicario, A.; González, J. M.; Pérez-Coello, M. S. Determination of anthocyanins in red wine using a newly developed method based on Fourier transform infrared spectroscopy. *Food Chem.* **2007**, *104*, 1295–1303.

(25) Versari, A.; Boulton, R.; Thorngate, J. The prediction of the color components of red wines using FTIR, wine analyses, and the method of Partial Least Squares. *ACS Symp. Ser.* **2004**, *886*, 53–67.

(26) Fernandez, K.; Agosin, E. Quantitative analysis of red wine tannins using Fourier-transform mid-infrared spectrometry. *J. Agric. Food Chem.* **2007**, *55*, 7294–7300.

(27) Jensen, J. S.; Egebo, M.; Meyer, A. Identification of spectral regions for the quantification of red wine tannins with Fourier transform mid-infrared spectroscopy. *J. Agric. Food Chem.* **2008**, *56*, 3493–3499.

(28) Cozzolino, D.; Kwiatkowski, M. J.; Parker, M.; Cynkar, W. U.; Dambergs, R. G.; Gishen, M.; Herderich, M. J. Prediction of phenolic compounds in red wine fermentations by visible and near infrared spectroscopy. *Anal. Chim. Acta* **2004**, *513*, 73–80.

(29) Di Egidio, V.; Sinelli, N.; Giovanelli, G.; Moles, A.; Casiraghi, E. NIR and MIR spectroscopy as rapid methods to monitor red wine fermentation. *Eur. Food Res. Technol.* **2010**, *230*, 947–955.

(30) Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: a basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–130.
(31) Kennard, R. W.; Stone, L. A. Computer aided design of experiments. *Technometrics* **1969**, *11*, 137–148.
(32) Martens, H.; Naes, T. *Multivariate Calibration*; Wiley: New York, 1989.
(33) Fragoso, S.; Aceña, L.; Guasch, J.; Busto, O.; Mestres, M. Application of FT-MIR spectroscopy for fast control of red grape phenolic ripening. *J. Agric. Food Chem.* **2011**, *59* (6), 2175–2183.
(34) Wu, D.; He, Y.; Shi, J.; Feng, S. Exploring near and midinfrared spectroscopy to predict trace iron and zinc contents in powdered milk. *J. Agric. Food Chem.* **2009**, *57*, 1697–1704.
(35) Williams, P. C. *Implementation of Near-Infrared Technology*; Williams, P. C., Norris, K. H., Eds.; American Association of Cereal Chemists: St. Paul, MN, 2001, pp 145−169.
(36) Faber, K.; Kowalski, B. Improved prediction error estimates for multivariate calibration by correcting for the measurement error in the reference values. *Appl. Spectrosc.* **1997**, *51*, 660–665.
(37) Moreira, J. L.; Santos, L. Spectroscopic interferences in fourier transform infrared wine analysis. *Anal. Chim. Acta* **2004**, *513*, 263–268.
(38) Boulet, J. C.; Williams, P.; Doco, T. A Fourier transform infrared spectroscopy study of wine polysaccharides. *Carbohydr. Polym.* **2007**, *69*, 79–85.
(39) Armenta, S.; Garrigues, S.; de la Guardia, M.; Rondeau, P. Attenuated total reflection-Fourier transform infrared analysis of the fermentation process of pineapple. *Anal. Chim. Acta* **2005**, *545*, 99–106.
(40) Bevin, C. J.; Fergusson, A. J.; Perry, W. B.; Janik, L. J.; Cozzolino, D. Development of a rapid "fingerprinting" system for wine authenticity by mid-infrared spectroscopy. *J. Agric. Food Chem.* **2006**, *54*, 9713–9718.
(41) Edelmann, A.; Diewok, J.; Schuster, K. C.; Lendl, B. Rapid method for the discrimination of red wine cultivars based on mid-infrared spectroscopy of phenolic wine extracts. *J. Agric. Food Chem.* **2001**, *49*, 1139–1145.
(42) Schindler, R.; Vonach, R.; Lendl, B.; Kellner, R. A rapid automated method for wine analysis based upon sequential injection (SI)-FTIR spectrometry. *Fresenius J. Anal. Chem.* **1998**, *362*, 130–136.