JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Quantitative Analysis of Red Wine Tannins Using Fourier-Transform Mid-Infrared Spectrometry

KATHERINA FERNÁNDEZ AND EDUARDO AGOSIN*

Department of Chemical and Bioprocess Engineering, School of Engineering, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, Macul, Santiago, Chile

Tannin content and composition are critical quality components of red wines. No spectroscopic method assessing these phenols in wine has been described so far. We report here a new method using Fourier transform mid-infrared (FT-MIR) spectroscopy and chemometric techniques for the quantitative analysis of red wine tannins. Calibration models were developed using protein precipitation and phloroglucinolysis as analytical reference methods. After spectra preprocessing, six different predictive partial least-squares (PLS) models were evaluated, including the use of interval selection procedures such as iPLS and CSMWPLS. PLS regression with full-range (650−4000 cm^{-1}), second derivative of the spectra and phloroglucinolysis as the reference method gave the most accurate determination for tannin concentration (RMSEC = 2.6%, RMSEP = 9.4%, $r = 0.995$). The prediction of the mean degree of polymerization (mDP) of the tannins also gave a reasonable prediction (RMSEC = 6.7%, RMSEP = 10.3%, $r = 0.958$). These results represent the first step in the development of a spectroscopic methodology for the quantification of several phenolic compounds that are critical for wine quality.

KEYWORDS: FT-MIR; PLS; tannins; red wine; SPE; Carménère

INTRODUCTION

Tannins are polymeric flavonoid compounds containing subunits of flavan-3-ol. They have a critical role in the final quality of red wines, mainly related with long-term color stability (*1*) and astringency (*2*); additionally, they are active antioxidants and probably provide important health benefits (*3*).

Several analytical methodologies have been developed to measure tannins, such as colorimetric techniques (*4*), precipitation with proteins (*5*), or acid-catalyzed cleavage in the presence of a nucleophilic agent (*6*, *7*). More recently, efforts have focused in the development of rapid and simple chemical techniques that could be easily implemented in a routine basis in the lab. Examples of these new methodologies are tannin precipitation with methyl cellulose (*8*), or with bovine serum albumin (*9*), among others.

The use of spectrometric techniques, such as Fourier transform mid-infrared (FT-MIR) spectroscopy combined with multivariate data analysis (chemometrics) has demonstrated to be a powerful analytical tool, that allows in a short time, to quantify and predict the concentration of several specific chemical compounds (*10*, *11*). This technique consists of quantifying the absorption in the mid-infrared region of the electromagnetic spectrum, usually ranging from 400 to 4000 cm^{-1}, of molecules containing specific chemical bonds, such as C═C, C−H, C═O, N−H, and O−H (*12*).

* To whom correspondence should be addressed. E-mail: agosin@ing.puc.cl. Tel: +56 2 354 4253. Fax: +56 2 354 5803.

The potential of FT-IR for the rapid analysis of multiple wine components has been extensively reported (*13*−*15*). More specific applications of this spectroscopic technique to wine have also been published, such as the measurement of tannin−protein interactions (*16*), the classification according to wine origin (*16*, *17*), the quantification and classification of polysaccharides (*18*), or the monitoring of sugars, alcohol, and organic acids during fermentation (*19*, *20*), among others.

Nevertheless, wine characterization by FT-MIR spectroscopy and chemometrics presents two main limitations, i.e., similar IR absorption bands of most interesting compounds, and dominating absorption of major wine components, particularly ethanol and water (*13*). Both limitations are critical for the analysis of phenolic compounds, because ethanol, water, and organic acids absorb in the same MIR region, masking the characteristic IR vibrations of phenols (*21*). To overcome these problems, chemometric techniques, such as partial least-squares (PLS) regression, are commonly used as a pattern recognition tool to develop mathematical models for the parameters of wine. Moreover, the incorporation of automatic spectral selection techniques—like interval PLS (iPLS) (*22*) or changeable size moving window PLS (CSMWPLS) (*23*)—to find out the most relevant spectral features may improve the performance of the models. In the specific case of wine phenolics, the interferences of water and ethanol absorption must be removed before the acquisition of the spectrum. For this purpose, either special instrumental devices, such as attenuated total reflectance (ATR) (*24*) or flow injection manifold (*16*), or additional sample

10.1021/jf071193d CCC: $37.00
Published on Web 08/15/2007

Table 1. Statistic Values of Calibration and Validation Sets of Tannins by Protein Precipitation Assay, Phloroglucinolysis, and Mean Degree of Polymerization (mDP)

	calibration set				validation set			
methodology	N[a]	range	mean	SD[b]	N[a]	range	mean	SD[b]
protein precipitation assay (mg/L)	63	85–900	397	189	21	89–804	404	186
phloroglucinolysis (mg/L)	62	117–514	312	81	20	153–455	300	89
mean degree of polymerization (mDP)	63	2.2–6.3	3.90	0.91	21	2.6–5.3	3.91	0.93

[a] N = number of samples. [b] SD = standard deviation.

processing, such as solid-phase extraction (*25*) and/or drying (*17*, *26*), can be used.

The aim of the present study was to investigate the suitability of Fourier transform mid-infrared (FT-MIR) spectroscopy as an accurate method for quantifying wine tannins, as part of a more general effort for the development of a simple spectroscopic methodology that allows the robust analysis of those phenolic compounds that are critical for wine quality. Several models were developed using multivariate PLS regression and spectral interval selection procedures (iPLS and CSMWPLS). Two reference methods, protein precipitation and phloroglucinolysis, were used to calibrate the models. The prediction of the mean degree of polymerization (mDP) was also evaluated.

MATERIALS AND METHODS

Materials. All solvents were of high-performance liquid chromatography (HPLC) grade. Acetonitrile, methanol and glacial acetic acid were purchased from J. T. Baker (Phillipsburg, NJ). Phloroglucinol, triethanolamine (TEA), ferric chloride hexahydrate, (+)-catechin (C), bovine serum albumin (BSA, fraction V), and sodium dodecyl sulfate (SDS) were purchased from Sigma (St. Louis, MO). Hydrochloric acid and anhydrous sodium acetate were purchased from E. M. Science (Gibbstown, NJ) and Mallinckrodt (Phillipsburg, NJ).

Samples. The sample set consisted of 86 red wines (vintages 2004 and 2005), of *Vitis vinifera* L. cv. Carménère, which were produced and kindly donated by different wineries situated in the central valleys of Chile. The wines contained tannins derived exclusively from the original grapes (no barrel or any other wood addition).

Sample Cleanup by Solid-Phase Extraction (SPE). In order to avoid interferences in absorption bands by other compounds present in high concentrations—ethanol, organic acids, among others (*21*)—the wines were previously purified by SPE. The C_{-18} SPE (1 g, CRS, Louisville, KY) column was activated with 10 mL of methanol followed by 10 mL of water. Four milliliters of wine was dealcoholized under reduced pressure, at 35 °C, and loaded onto the activated column. The column was washed with 10 mL of water and the cartridge dried with N_2. The phenolic compounds were eluted with 4 × 0.5 mL of methanol to complete 2 mL total. The resulting methanolic extract was divided in two: 0.5 mL for FT-MIR measurements and 1.5 mL for phloroglucinolysis assay.

Reference Methods. *Protein Precipitation Assay.* Tannins were analyzed by the modified protein precipitation assay from Harbertson et al. (*9*). Wines (200 μL) were diluted 2.5 times with a buffer of 12% (v/v) ethanol containing 5 g/L potassium bitartrate adjusted to pH 3.3 with HCl (6 M). Bovine serum albumin (BSA, fraction V) (1 mg/mL) was dissolved in a buffer containing 200 mM acetic acid and 170 mM NaCl, and adjusted to pH 4.9 with NaOH (6M). The precipitation reaction was carried out with 500 μL of diluted wine and 1 mL of BSA buffer. The mixture was incubated at room temperature for 15 min under agitation (100 rpm). Then the samples were centrifuged for 5 min at 13500*g* to separate the tannin−protein precipitate. The supernatant was poured off, and the pellet was dissolved in a buffer containing 5% (v/v) TEA and 5% (w/v) SDS and kept at room temperature for 10 min. The tube was then vortexed until the tannin−protein pellet was completely dissolved. Absorbance was read at 510 nm (background). After this reading, 125 μL of ferric chloride (10 mM) in 10 mM HCl was added, and after 10 min of incubation at room temperature, the final absorbance at 510 nm was determined. The amount of tannin−protein was calculated as the final absorbance minus the background, and expressed as catechin equivalents (mg CE/L) after comparison with a standard curve prepared with pure catechin (coefficient of variation <5%).

Phloroglucinolysis. The phenolic extract (0.5 mL) was analyzed by phloroglucinolysis followed by reversed-phase HPLC, as described in (*6*, *27*), using double strength (*28*). This method allows one to determine both the tannin concentration, as the sum of the total released units, flavan-3-ols, and adducts, and the average molecular weight (expressed as mDP) of the tannin components. HPLC analyses were carried out with a Merck-Hitachi L7100 (Darmstadt, Germany) consisting of a vacuum degasser, an autosampler, a quaternary pump, a diode array detector, and a column heater. A computer workstation with the D 7000 HSM software was used for chromatographic analysis. Tannin quantification was determined by comparing the sum of known proanthocyanidin products with catechin as standard (coefficient of variation <5%).

Infrared Spectroscopy Measurement. Acquisition of the infrared spectra of the phenolic extracts was carried out in an Avatar 360 FT-IR spectrometer (Thermo Nicolet Corporation, Madison, WI) equipped with a DTGS KBr detector. The software OMNIC version 6.0a from Thermo Nicolet was used for spectra acquisition. The data were recorded in transmission mode, with 32 scan at 8 cm^{-1} resolution, from 650 to 4000 cm^{-1}, at 26 ± 2 °C. The total number of data points was 869 for each spectrum. The phenolic extract (50 μL) was cast onto a ZeSn crystal (0.785 mm^2) and dried under reduced pressure till total removal of the methanol. A reference scan of the ZeSn crystal was taken every 10 min. The crystal was carefully cleaned with ethanol and water between measurements. All samples were scanned in triplicate. The whole procedure, including the SPE step (coefficient of variation <3%) and the drying of the extract and spectrum acquisition (coefficient of variation <2%), was highly reproducible.

Model Development. Linear models for tannin and mDP quantification were developed by regressing spectral data with reference methods using PLS algorithm. Spectral preprocessing and multivariate data analysis were performed with TQ analyst (version 6.21, Thermo Nicolet) software and Matlab 6.5 (Mathworks Inc., Natick, MA) software.

Wine samples were divided into a calibration set (75% of the samples) and an external validation set (25% of the samples). **Table 1** summarizes the tannin concentration obtained by phloroglucinolysis and BSA, as well as the mDP, for each data set. Assignment of samples to either the calibration or validation set was based in the maxi-min strategy and performed by TQ software. This strategy selects a specific number of samples from a pool of candidates based in the spectral information for those samples. The software performs a principal components analysis to determine the sources of variation in the samples. Then, it plots the distribution of the candidate samples from the mean spectrum and uses the distance values to determine appropriate samples for the model.

Model Preprocessing. Prior to calibration, the spectra triplicate were averaged, the baseline was corrected, and the outliers were removed by the TQ analyst software. Then, the data matrices were mean-centered, smoothed with Savitsky and Golay filter (*29*) using a seventh order polynomial, and the first derivative and second derivative of the spectra were taken. Norris's smoothing was also evaluated in the best models, but it resulted in less accurate calibrations.

Table 2. Functional Groups and Frequency Assignments for Tannins from FT-MIR Spectra

vibrational group	group frequency wavenumber (cm^{-1})	assignment
C–H	670–900 (several)	aromatic C–H out-of-plane bend
C–H	950–1225 (several)	aromatic C–H in-plane bend
C–O	1230–1320	C–O stretch from pyran-derived ring structure
C═C–C	1450–1510; 1580–1615	aromatic ring stretch
combi	1660–2000	aromatic combination bands
C–H	3070–3130	aromatic C–H stretch
O–H	3200–3400; 3200–3570 (broad)	normal polymeric OH, stretch; hydroxy group, H-bonded OH stretch

Wavelength Selection. Full, partial, and optimized local PLS models were assessed. Local PLS models were coded in Matlab using two interval selection procedures: interval partial least-squares (iPLS) (*22*) and changeable size moving window partial least-squares (CSMWPLS) (*23*). These procedures search for the most informative regions in the spectra with important covarying spectral regions, to build models on fewer variables (*30*).

In total, six models were built based in the following spectral regions: Region one included the main range frequencies observed in the FT-MIR spectrum of tannins (**Table 2**). Region two ranged from 650 to 4000 cm^{-1} wavelength (full-range). Region three ranged from 800 to 1800 cm^{-1} (fingerprint zone). Region four corresponded to a zone suggested by the PLS software package of the TQ analyst, based on the spectral information and the analysis of path length parameters. Regions five and six were spectral regions suggested from the iPLS and CSMWPLS procedures, respectively. The optimal spectral zones for regions 4, 5, and 6 are specified in **Tables 3**, **4**, and **5**, respectively.

Model Evaluation. PLS models with 1−10 factors were investigated. The optimal number of PLS factors to use in the PLS models was obtained by the cross-validation method with the leave-one-out option (*31*), based on the minimum predicted residual sum of squares (PRESS) and an optimum correlation coefficient (r) value, which should be as high as possible. The predictive ability of the models was tested by computing the root mean standard error of calibration (RMSEC) and root mean standard error of prediction (RMSEP). The % RMSEP was calculated as follows (*32*):

$$\text{RMSEP (\%)} = \frac{100}{\bar{y}}\left(\sqrt{\frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{N}}\right)$$

where N = number of samples, y_i = actual concentration, $\hat{y}_i$ = predicted concentration, and $\bar{y}$ = average of actual concentration values. The % RMSEC was calculated to replace in the equation the values of calibration.

RESULTS AND DISCUSSION

Spectral Features. The full spectra of the dry extracts presented main differences in the wavelength regions between 2800−3700 cm^{-1} and 800−1800 cm^{-1} of the electromagnetic spectrum (**Figure 1**). Dried wine extracts are complex mixtures of various phenolic compounds, and full assignment of the spectral bands is challenging. In spite of this, the sharp absorption bands and shoulders at different frequencies could be attributed to the functional groups given in **Table 2**. The prominent absorption band around 3400 cm^{-1} can be associated with O−H stretching and C−H stretching vibrations (*33*). The peaks between 800 cm^{-1} and 1800 cm^{-1} (**Figure 2**), the fingerprint zone, could be attributed to C═C−C aromatic ring stretching (1580−1615 cm^{-1}; 1450−1510 cm^{-1}) and several aromatic C−H out-of-plane (670−900 cm^{-1}) and in-plane (950−1225 cm^{-1}) bending vibrations, among others (*26*, *33*). In particular, the peak around 1285 cm^{-1} indicates a characteristic feature for the flavonoid-based tannins (*16*). This peak was assigned to the ethereal C−O stretching vibration arising from the pyran-derived ring structure of this class of tannins (*34*).

Data Preprocessing. An important decision when working with FTIR data is to determine whether a preprocessing method is necessary before the data analysis. Noise and bias are undesirable features in the spectra and should be removed or reduced. Bias can be removed by applying first or second derivative, while noise can be attenuated with filtering or smoothing algorithms (*31*). We therefore smoothed the data and compared the raw transmission spectra with their first and second derivatives. Several sharp absorption bands were observed in the fingerprint region (800−1800 cm^{-1}) of the derived spectra, particularly for the second derivative one (**Figure 1** vs **Figure 3**).

Tannin Models. We calibrated the FT-MIR raw spectra, as well as their corresponding first and second derivatives against the chemical reference methods, using partial least-square (PLS) regressions. The values of RMSEC, RMSEP, r, and their respective percentages indicate the precision achieved in the calibration and validation models. These values were more accurate for the second derivative of the spectra than for the other pretreatments of the data (first derivative and raw spectra).

Table 3 shows the statistical parameters for the models calibrated with the protein precipitation assay. The six models tested provided good calibration models ($r > 0.9$), although neither of them presented an accurate prediction (% RMSEP > 10%). In particular, the main frequencies model had the highest number of factors and its % RMSEC and % RMSEP were over 12%. The identification of the main spectrum features did not help to correlate the spectral information with the tannin concentration. The full-range model had the lowest number of PLS factors and gave the most accurate values, i.e., RMSEC = 9.0%, and $r = 0.981$. **Figure 4** illustrates tannin concentration determined by protein precipitation assay versus tannin concentration by FT-MIR for this model. The optimized local PLS models tested (Fingerprint, TQ analyst, iPLS, CSMWPLS) did not show superior calibrations, or better prediction capacity.

When phloroglucinolysis was used as the reference method, even more accurate calibration models were obtained. **Table 4** shows the statistical parameters obtained for the models. Here, the PLS correlation based on the main frequencies at which the chemical groups absorb gave a suitable calibration and prediction of the tannin concentration (RMSEC = 7.7%, RMSEP = 9.9%, and $r = 0.952$). This model was the second more accurate of the models tested. The best PLS model (RMSEC = 2.6%, RMSEP = 9.4% and $r = 0.995$) was achieved when the full-range was used. **Figure 5** illustrates the tannin concentration by phloroglucinolysis versus the tannin concentration by FT-MIR for this model. Even though the fingerprint, TQ Analyst, iPLS, and CSMWPLS models had lower number of PLS, their statistics values were not good enough to be considered accurate. In general, the model errors observed were even lower than the calibration errors performed with protein precipitation assay, which points to a better tannin prediction with FT-MIR calibrated using phloroglucinolysis.

The differences observed between the models for tannin quantification could be associated with variations in the nature of the reference methods. Protein precipitation assay was applied to the whole wine; this method determines only the phenols

Table 3. Summary of Calibration and Prediction Statistics[a] for Tannin (mg/L) Determination[b]

model	region (cm^{-1})	PLS factors	RMSEC	(%) RMSEC	RMSEP	(%) RMSEP	*r*	outliers
main frequencies	670–900; 950–1225; 1230–1320; 1450–1510; 1580–1615; 1660–2000	10	49.6	12.5	50.7	12.5	0.964	2
full-range	650–4000	8	35.8	9.0	53.7	13.2	0.981	2
Fingerprint	800–1800	9	47.9	12.1	51.9	12.8	0.966	2
TQ Analyst	791–2989	9	42.5	10.7	51.9	12.8	0.974	2
iPLS	953–1751	9	47.8	12.0	51.1	12.6	0.967	2
CSMWPLS	1026–1196; 1489–1751	8	46.8	11.8	64.6	15.9	0.968	2

[a] RMSEC, root mean standard error of calibration; RMSEP, root mean standard error of prediction; *r*, coefficient of correlation. [b] Calibration was carried out with protein precipitation assay as reference method and using the second derivative of the spectra.

Table 4. Summary of Calibration and Prediction Statistics[a] for Tannin (mg/L) Determination[b]

model	region (cm^{-1})	PLS factors	RMSEC	(%) RMSEC	RMSEP	(%) RMSEP	*r*	outliers
main frequencies	670–900; 950–1225;1230–1320; 1450–1510; 1580–1615; 1660–2000	10	24.3	7.7	29.8	9.9	0.952	4
full-range	650–4000	9	7.97	2.6	29.9	9.4	0.995	4
Fingerprint	800–1800	6	32.9	10.5	34.6	11.5	0.911	3
TQ Analyst	791–1300	6	38.0	12.2	50.9	16.9	0.879	3
iPLS	1003–1250	4	31.3	10.0	34.4	11.5	0.902	3
CSMWPLS	1084–1207	5	34.8	11.1	36.3	12.1	0.890	3

[a] RMSEC, root mean standard error of calibration; RMSEP, root mean standard error of prediction; *r*, coefficient of correlation. [b] Calibration was carried out with phloroglucinolysis as reference method and using the second derivative of the spectra.

Table 5. Summary of Calibration and Prediction Statistics[a] for mDP Determination[b]

model	region (cm^{-1})	PLS factors	RMSEC	(%) RMSEC	RMSEP	(%) RMSEP	*r*	outliers
main frequencies	670–900; 950–1225; 1230–1320; 1450–1510; 1580–1615; 1660–2000	10	0.351	9.0	0.632	16.1	0.922	2
full-range	650–4000	8	0.261	6.7	0.405	10.3	0.958	2
Fingerprint	800–1800	7	0.450	11.5	0.628	16.0	0.869	2
TQ Analyst	1084–2958	6	0.517	13.2	0.550	14.0	0.823	2
iPLS	1170–1500	6	0.506	12.9	0.649	16.6	0.831	2
CSMWPLS	1686–1732	7	0.412	10.6	0.543	13.9	0.861	2

[a] RMSEC, root mean standard error of calibration; RMSEP, root mean standard error of prediction; *r*, coefficient of correlation. [b] Calibration was carried out with phloroglucinolysis as reference method and using the second derivative of the spectra.

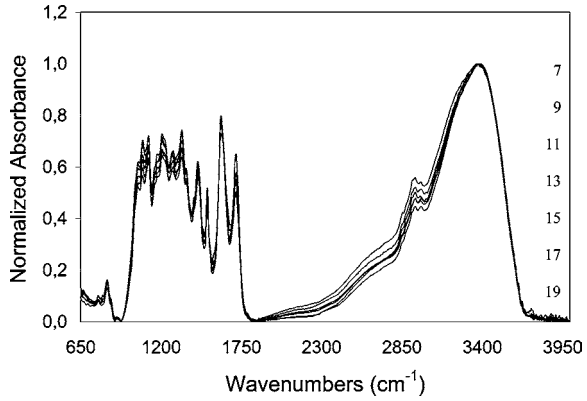

Figure 1. FT-MIR spectra for seven phenolic extracts (650–4000 cm^{-1}) after baseline correction and normalization scale.

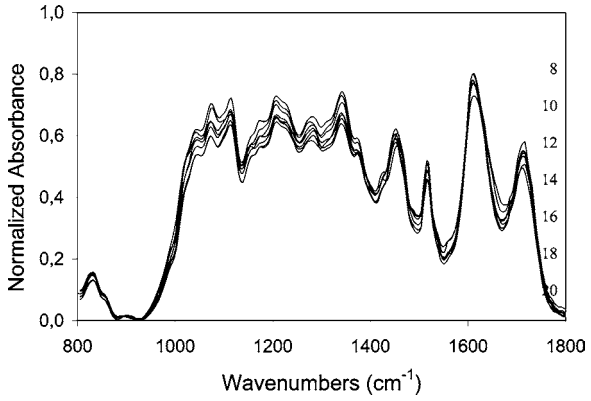

Figure 2. FT-MIR spectra for seven phenolic extracts, fingerprint zone (800–1800 cm^{-1}) after baseline correction and normalization scale.

that are bound to the protein and precipitate. On the other hand, phloroglucinolysis was applied to the phenolic extracts and it quantifies the flavan-3-ols and adducts released after acid catalysis. Finally, the spectral acquisition was carried out from the phenolic extracts. Thus, the lower accuracy of the model calibrated with protein precipitation assay could be associated with small differences between the compounds measured in the wine and the compounds present in the spectrum (phenolic extract). Moreover, the comparison of both main frequencies models (**Tables 3** and **4**) could also evidence these differences.

Among all models tested, the superior performance presented for the full-range model seems to contradict the relationship between the spectral features and the tannin concentration found

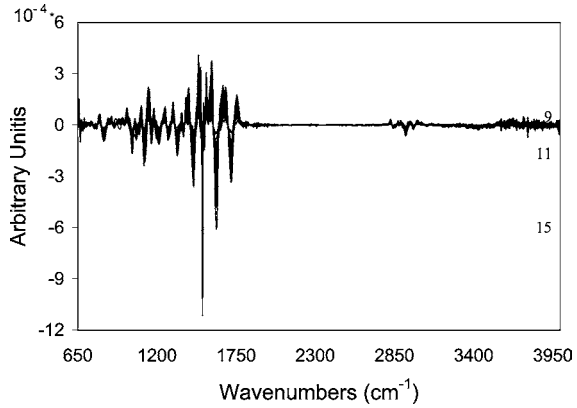

Figure 3. Second derivative spectra of phenolic extracts.

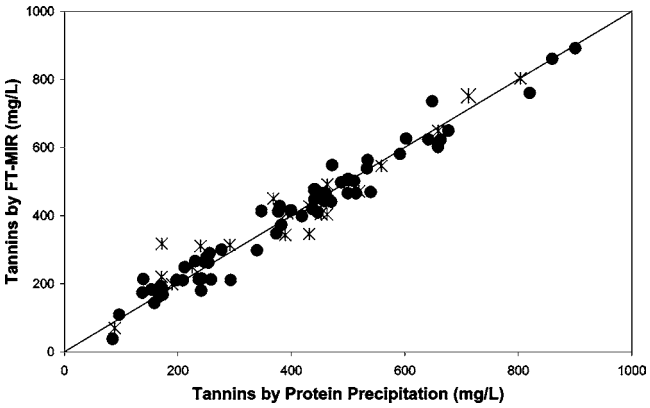

Figure 4. Correlation plot of best calibration and validation for the prediction of tannins (mg/L) using FT-MIR and protein precipitation assay. (●) Calibration samples, (✳) external validation samples, and (—) line of perfect correlation. Wavelength range from 650 cm^{-1} to 4000 cm^{-1}; second derivative, PLS calibration statistics RMSEC = 9.0%, RMSEP = 13.2%, $r = 0.981$.

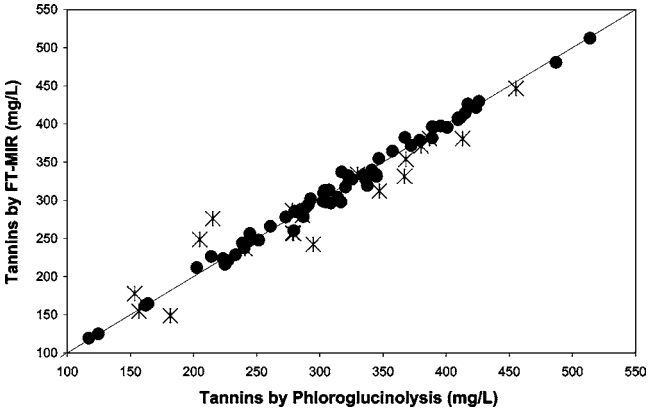

Figure 5. Correlation plot of best calibration and validation for the prediction of tannins (mg/L) using FT-MIR and phloroglucinolysis. (●) Calibration samples, (✳) external validation samples, and (—) line of perfect correlation. Wavelength range from 650 cm^{-1} to 4000 cm^{-1}; second derivative, PLS calibration statistics RMSEC = 2.6%, RMSEP = 9.4%, $r = 0.995$.

for the PLS model. We believe that the heterogeneous nature of the sample matrix, given by the presence of polymeric pigments in the phenolic extracts, which were not completely observed with the main frequencies model, could have produced better calibration results taking into consideration more of the spectral information (full-range).

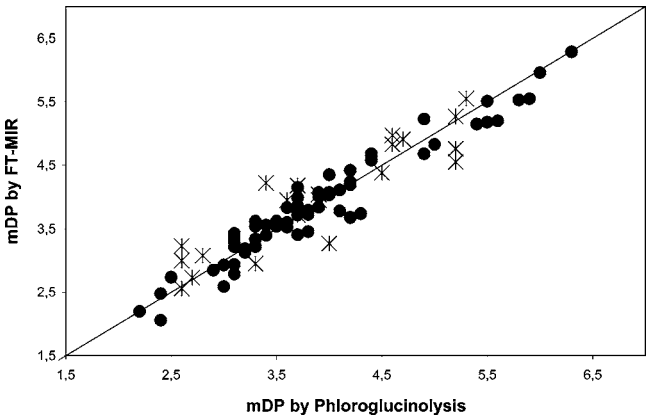

Figure 6. Correlation plot of best calibration and validation for the prediction of mDP using FT-MIR. (●) Calibration samples, (✳) external validation samples, and (—) line of perfect correlation. Wavelength range from 650 cm^{-1} to 4000 cm^{-1}; second derivative, PLS calibration statistics RMSEC = 6.7%, RMSEP = 10.3%, $r = 0.958$.

In addition, to obtain more accurate results, we developed the models on a single red wine variety, given that more accurate calibrations are usually obtained when the wines are segregated by variety (*20*, *21*). Thus, to generalize the models to any red wine, many more samples will be necessary.

It is worthwhile mentioning that the strong correlation existing between tannin concentration and wine astringency (*35*) opens the possibility of using the models developed here to predict wine astringency.

Mean Degree of Polymerization (mDP). Valuable information regarding the structure of proanthocyanidins is mDP, that is, the average number of monomers per polymer chain molecule. We estimated the mDP of the wine samples by phloroglucinolysis, and using FT-MIR together with PLS regression. The statistical parameters are presented in **Table 5.** The results indicated that the correlation of mDP with main frequencies of the spectra gave an accurate calibration (RMSEC = 9.0% and $r = 0.922$), although the number of PLS factors was the higher and the RMSEP = 16.1% was not good enough. The full-range model fitted best ($r = 0.958$), showing the lowest RMSEC = 6.7% and RMSEP = 10.3%. **Figure 6** shows the correlation between the mDP determined by phloroglucinolysis and the mDP by FT-MIR for the best calibration model. Inspection of the scatter plots and statistics (**Table 5**) shows a good fit of the model. The local PLS models (iPLS and CMWPLS) resulted in even lower accuracy than the fingerprint and than the zone suggested by the TQ software.

Even though PLS automatically gives high priority to wavelengths that covary with the variables to calibrate, it is argued that the predictive ability often increases when the region is based only on the significant variable intervals (*22*, *23*). We tried to prove this hypothesis using iPLS and CSMWPLS as automatic variable selection methods. However, our calibration and prediction models did not improve, although the number of factors found was always the lowest obtained. This agrees with the principles of variable reduction postulated in the selection methods (*30*). It is possible that the chemical information of the present study (the signal-to-noise ratio of tannin information relative to the background) was good enough for simpler models to adequately describe the variations. Therefore, and as a general strategy, it might be better to stick to simpler models to improve the chances of building more robust models, which would be more useful over the extended experimental time periods required in industrial applications.

As a whole, this work showed that FT-MIR spectrometry combined with multivariate data analysis allows an accurate determination of tannin concentration and mDP in wines. The possibility of predicting these characteristics of wines is an invaluable tool when a large number of tannin samples need to be analyzed. However, further work is needed to extend these results to other important grape and wine phenolic compounds, such as anthocyanins and flavanols, as well as to the development of similar methods for their assessment in other grape varieties. In the future, we expect that the models developed here, together with other mid-infrared spectral calibrations for other important phenolic compounds (anthocyanins, copigmentation, total phenols, etc.) in wine, currently under development in our group, could be successfully transferred and implemented as routine analyses in wineries.

ACKNOWLEDGMENT

The authors thank the following Chilean wineries for the Carménère varietal wines kindly donated for this study: Misiones de Rengo, Luis Felipe Edwards, Bisquert, Requingua, La Ronciere, Morande, Conosur, La Fortuna, La Estampa, Viu Manet, Mont Grass, San Pedro, Casa Silva, Tarapacá, Aresti, Undurraga, de Martino, and Carmen. We are also grateful to Dr. Ricardo Pérez and Ana Maria Molina for their constructive comments on the manuscript.

LITERATURE CITED

(1) Somers, T. C. The polymeric nature of wine pigments. *Phytochemistry* **1971**, *10*, 2175−2186.
(2) Gawel, R. Red wine astringency: A review. *Aust. J. Grape Wine Res.* **1998**, *4*, 74−95.
(3) Scalbert, A.; Manach, C.; Morand, C.; Remesy, C.; Jimenez, L. Dietary polyphenols and the prevention of diseases. *Crit. Rev. Food Sci. Nutr.* **2005**, *45*, 287−306.
(4) Singleton, V. L.; Orthofer, R.; Lamuela-Raventos, R. M. Analysis of total phenols and other oxidation substrates and antioxidants by means of Folin-Ciocalteu reagent. *Methods Enzymol.* **1999**, *299*,152−178
(5) Hagerman, A. E.; Butler, L. G. Determination of protein in tannin-protein precipitates. *J. Agric. Food Chem.* **1980**, *28*, 944−947.
(6) Kennedy, J. A.; Jones, G. P. Analysis of proanthocyanidin cleavage products following acid-catalysis in the presence of excess phloroglucinol. *J. Agric. Food Chem.* **2001**, *49*, 1740−1746.
(7) Guyot, S.; Marnet, N.; Laraba, D.; Sanoner, P.; Drilleau, J. P. Reversed-Phase HPLC following thiolysis for quantitative estimation and characterization of the four main classes of phenolic compounds in different tissue zones of a french cider apple variety (Malus domestica Var. Kermerrien). *J. Agric. Food Chem.* **1998**, *46*, 1698−1705.
(8) Sarneckis, C. J.; Dambergs, R. G.; Jones, P.; Mercurio, M.; Herderich, M. J.; Smith, P. A. Quantification of condensed tannins by precipitation with methyl cellulose: development and validation of an optimised tool for grape and wine analysis. *Aust. J. Grape Wine Res.* **2006**, *12*, 39−49.
(9) Harbertson, J. F.; Kennedy, J. A.; Adams, D. O. Tannin in skins and seeds of Cabernet Sauvignon, Syrah, and Pinot noir berries during ripening. *Am. J. Enol. Vitic.* **2002**, *53*, 54−59.
(10) 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.
(11) Dixit, V.; Tewari, J. C.; Cho, B. K.; Irudayaraj, J. M. K. Identification and quantification of industrial grade glycerol adulteration in red wine with Fourier transform infrared spectroscopy using chemometrics and artificial neural networks. *Appl. Spectrosc.* **2005**, *59*, 1553−1561.
(12) Silverstein, R. M.; Bassler, G. C.; Morrill, T. C. *Spectrometric identification of organic compounds*; John Wiley & Sons: New York, 1981; pp 95−137.
(13) Patz, C. D.; Blieke, A.; Ristow, R.; Dietrich, H. Application of FT-MIR spectrometry in wine analysis. *Anal. Chim. Acta* **2004**, *513*, 81−89.
(14) Urbano Cuadrado, M.; Luque de Castro, M. D.; Pérez Juan, P. M.; Gomez-Nieto, M. A. Comparison and joint use of near-infrared spectroscopy and Fourier transform mid infrared spectroscopy for the determination of wine parameters. *Talanta* **2005**, *66*, 218−224.
(15) Moreira, J. L.; Santos, L. Analysis of organic acids in wines by Fourier-transform infrared spectroscopy. *Anal. Bioanal. Chem.* **2005**, *382*, 421−425.
(16) Edelmann, A.; Lendl, B. Toward the optical tongue: Flow-through sensing of tannin-protein interactions based on FTIR spectroscopy. *J. Am. Chem. Soc.* **2002**, *124*, 14741−14747.
(17) Picque, D.; Cattenoz, T.; Corrieu, G.; Berger, J. L. Discrimination of red wines according to their geographical origin and vintage year by the use of mid-infrared spectroscopy. *Sci. Aliments* **2005**, *25*, 207−220.
(18) Coimbra, M. A.; Goncalves, F.; Barros, A. S.; Delgadillo, I. Fourier transform infrared spectroscopy and chemometric analysis of white wine polysaccharide extracts. *J. Agric Food Chem.* **2002**, *50*, 3405−3411.
(19) Cozzolino, D.; Parker, M.; Dambergs, R. G.; Herderich, M.; Gishen, M. Chemometrics and visible-near-infrared spectroscopic monitoring of red wine fermentation in a pilot scale. *Biotech. Bioeng.* **2006**, *95*, 1101−1107.
(20) Urtubia, A.; Perez-Correa, J. R.; Meurens, M.; Agosin, E. Monitoring large scale wine fermentations with infrared spectroscopy. *Talanta* **2004**, 64, 778−784.
(21) Moreira, J. L.; Santos, L. Spectroscopic interferences in Fourier transform infrared wine analysis. *Anal. Chim. Acta* **2004**, *513*, 263−268.
(22) Norgaard, L.; Saudland, A.; Wagner, J.; Nielsen, J. P.; Munck, L.; Engelsen, S. B. Interval partial least-squares regression (iPLS): A comparative chemometric study with an example from near-infrared spectroscopy. *Appl. Spectrosc.* **2000**, *54*, 413−419.
(23) Du, Y. P.; Liang, Y. Z.; Jiang, J. H.; Berry, R. J.; Ozaki, Y. Spectral regions selection to improve prediction ability of PLS models by changeable size moving window partial least-squares and searching combination moving window partial least-squares. *Anal. Chim. Acta* **2004**, *501*, 183−191.
(24) 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.
(25) Picque, D.; Lieben, P.; Corrieu, G.; Cantagrel, R.; Lablanquie, O.; Snakkers, G. Discrimination of cognacs and other distilled drinks by mid-infrared spectropscopy. *J. Agric. Food Chem.* **2006**, *54*, 5220−5226.
(26) Picque, D.; Cattenoz, T.; Corrieu, G. Classification of red wines analysed by middle infrared spectroscopy of dry extract according to their geographical origin. *J. Int. Sci. Vigne Vin* **2001**, *35*, 165−170.
(27) Kennedy, J. A.; Taylor, A. W. Analysis of proanthocyanidins by high-performance gel permeation chromatography. *J Chromatogr. A* **2003**, *995*, 99−107.
(28) Peyrot des Gachons, C. P.; Kennedy, J. A. Direct method for determining seed and skin proanthocyanidin extraction into red wine. *J. Agric. Food Chem.* **2003**, *51*, 5877−5881.
(29) Savitsky, A.; Golay M. J. E. Smoothing and differentiation of data by simplified Least-Squares Procedures. *Anal. Chem.* **1964**, *36*, 1627−1639.
(30) Jiang, J. H.; Berry, R. J.; Siesler, H. W.; Ozaki, Y. Wavelength interval selection in multicomponent spectral analysis by moving window partial least-squares regression with applications to mid-infrared and near-infrared spectroscopic data. *Anal. Chem.* **2002**, *74*, 3555−3565.

(31) Martens, H.; Tormod, N. *Multivariate calibration*; John Wiley and Sons: Chichester, U.K., 1993.
(32) Duarte, I. F.; Barros, A.; Delgadillo, I.; Almeida, C.; Gil, A. M. Application of FTIR spectroscopy for the quantification of sugars in mango juice as a function of ripening. *J. Agric. Food Chem.* **2002**, *50*, 3104−3111.
(33) Cotes, J. *Interpretation of infrared spectra, a practical approach*; John Wiley & Sons: Chichester, U.K., 2000.
(34) Socrates, G. *Infrared and Raman characteristic group frequencies: tables and charts*; John Wiley and Sons: Chichester, U.K., 2001.
(35) Kennedy, J. A.; Ferrier, J.; Harbertson, J. F. and Peyrot des Gachons, C. Analysis of tannins in red wine using methods: correlation with perceived astringency. *Am. J. Enol. Vitic.* **2006**, *57*, 481−485.

Received for review April 23, 2007. Revised manuscript received June 19, 2007. Accepted July 3, 2007. This research was funded by the Chilean National Council for Scientific and Technological Research, CONICYT, Grant No. AT 24050180.

JF071193D