Universiteit
Leiden
The Netherlands

NMR spectroscopy and chemometrics-based analysis of grapevine

Ali, K.

Citation

Ali, K. (2011, September 20). *NMR spectroscopy and chemometrics-based analysis of grapevine*. Retrieved from https://hdl.handle.net/1887/17843

Version:	Corrected Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/17843

Note: To cite this publication please use the final published version (if applicable).

CHAPTER 7

Metabolic characterization of Palatinate German white wines according to sensory attributes, varieties, and vintages using NMR spectroscopy and multivariate data analyses

Kashif Ali[a], Federica Maltese[a], Reinhard Toepfer[b], Young Hae Choi[a], and Robert Verpoorte[a]

[a] Natural Products Laboratory, Institute of Biology, Leiden University, The Netherlands
[b]Julius Kuehn Institute, Federal Research Centre for Cultivated Plants, Institute for Grapevine Breeding Geilweilerhof, Germany

Abstract

NMR spectroscopy has been used for metabolomic analysis of 'Riesling' and 'Mueller-Thurgau' white wines from the German Palatinate region. Diverse two-dimensional NMR techniques have been applied for the identification of metabolites, including phenolics. It is shown that sensory analysis correlates with NMR-based metabolic profiles of wine. ^{1}H NMR data in combination with multivariate data analysis methods, like principal component analysis (PCA), projections to latent structures (PLS), and bidirectional orthogonal projections to latent structures (O2PLS) analysis, were employed in an attempt to identify the metabolites responsible for the taste of wine, using a non-targeted approach. The high quality wines were characterized by elevated levels of compounds like proline, 2,3-butanediol, malic acid, quercetin, and catechin. Characterization of wine based on type and vintage was also done using orthogonal projections to latent structures (OPLS) analysis. 'Riesling' wines were characterized by higher levels of catechin, caftarate, valine, proline, malate, and citrate whereas compounds like quercetin, resveratrol, gallate, leucine, threonine, succinate, and lactate, were found discriminating for 'Mueller-Thurgau'. The wines from 2006 vintage were dominated by leucine, phenylalanine, citrate, malate, and phenolics, while valine, proline, alanine, and succinate were predominantly present in the 2007 vintage. Based on these results, it can be postulated the NMR-based metabolomics offers an easy and comprehensive analysis of wine and in combination with multivariate data analyses can be used to investigate the source and to predict certain sensory aspects of wines.

Introduction

Wine, a beverage of long tradition and high value, contains a complex mixture of compounds at a wide range of concentrations. The compounds in wine cover a diverse range of metabolites including primary (e.g. sugars, organic acids, amino acids) and secondary metabolites (e.g. flavonoids, hydroxycinnamates, hydroxybenzoates, anthocyanins) (Ali et al. 2010). Unlike ethanol and glycerol, generally low concentrations of other components are present in wine (Amaral and Caro 2005; Kosir and Kidric 2002). Many compounds influence sensory attributes of wine including color, mouth-feel, flavor, and aroma (Pickering and Nikfardjam 2010).

Different factors like grapevine variety, harvest time, vineyard environment, yeast strain, winemaking technologies, storage time, and human practices, have an affect on wine quality which is determined by the concentration and composition of the compounds present. Thus on this chemical basis, classification and characterization of wine is feasible through qualitative and quantitative information of the wine constituents (Pérez-Magariño and González-San José 2006; Castillo-Sánchez et al. 2006; Matejicek et al. 2005; Guash-Jane et al. 2004). The most important and basic factor for making wine of good quality is the grape variety and because of this many articles on grapes and wine quality relationships have been published (Gergaud and Ginsburgh 2008).

Cell regulatory processes result in several intermediates and end products which are generally regarded as metabolites. Out of these 'secondary metabolites' are produced due to the interaction of the organism with its environment (Pereira et al. 2005). Metabolomics, generally defined as metabolic snap shots of living systems, is a holistic approach aimed at better understanding of primary and secondary metabolism of the biological systems (Clayton et al. 2006). With the advancement in the field of analytical chemistry, more powerful and sophisticated tools (like mass spectrometry and nuclear magnetic resonance) for such chemical analyses have been introduced. The development of methods and algorithms for the multivariate statistical modeling have contributed much to metabolomics as they opened the way for handling the huge datasets of large-scale metabolic analyses (Crockford et al. 2006).

The chemical analysis of complex mixtures like wine can be of great importance as it can be used for the differentiation of wines based on the grapevine variety, yeast strain, geographical origin, terroir, and vintage. Many reports have been published using methods like mass spectrometry (MS) (Giraudel et al. 2007), nuclear magnetic resonance (NMR) spectroscopy (Son et al. 2009a), mid-infrared (MIR) spectroscopy (Bevin et al. 2008), gas-chromatography (GC) (Aznar and Arroyo 2007), liquid-chromatography (LC) (Matejicek et al. 2005), atomic emission spectroscopy (AES) (Iglesias et al. 2007), inductively coupled plasma emission spectroscopy (ICP) (Moreno et al. 2008), along with the hyphenation of these methods (Guasch-Jane et al. 2004) for the wine analysis.

Since nearly no sample pretreatment is required in NMR spectroscopy, the inherent properties of the sample are well kept. As NMR is non selective, it is an ideal tool for

the profiling of broad range metabolites in the field of agriculture (Dixon et al. 2006). NMR has been already demonstrated to be a robust method and unaffected by instrumental and experimental factors as it is the case in other analytical methods. Some recent publications reported the use of NMR for the studies like quality control, authenticity or geographical characterization of different food stuffs like coffee, olive oil, tomato and orange juices, wine and beer (Nord et al. 2004; Charlton et al. 2002; D'Imperio et al. 2007; Sobolev et al. 2003; Le Gall et al. 2001).

The aim of the present work was to identify the wine phenolics using ^{1}H NMR spectroscopy with some two-dimensional NMR techniques without any chromatographic fractionation, along with the correlation of sensory data (quality score rating) of the different wine samples with the identified metabolites using different multivariate data analysis methods. The metabolic characterization of wines based on grape cultivar and vintage was also a target.

Materials and Method

Wine samples and extraction

Two types of German white wines from Palatinate region ('Riesling' and 'Mueller-Thurgau') were used in this study. The experimental design included 40 samples of 'Riesling' and 19 samples of 'Mueller-Thurgau' from the vintages 2006 and 2007. Each sample has been tasted by a panel of experts and scored/rated accordingly. On the basis of scores, samples are classified into four classes. The class number actually indicates the quality grade of wine samples i.e. wines from class 1 were low quality wines whereas wines from class 4 were high quality wines (Table S1).

To generate reference data wines were directly analyzed for their content of ethanol, glucose, fructose, tartaric acid, malic acid, lactic acid, and glycerol by a standard procedure of FTIR (Fourier Transform Infrared Spectroscopy) using WineScan^{TM} 120 FT (FossGmbH, Rellingen, Germany). For NMR spectroscopy each wine sample (20 ml) has been extracted by an equal volume of ethyl acetate. The extraction was repeated two more times and all the three extracts were pooled together. The pooled ethyl acetate fraction was treated with Na_2SO_4 to remove residual water. Then the Ethyl Acetate was evaporated using a rotary evaporator and stored at -20 °C until further use.

NMR spectroscopy

The evaporated ethyl acetate extract was redissolved in 1 ml of deutrated methanol (CH_3OH-d_4). An aliquot of 800 µl of sample was transferred to the 5 mm NMR tube and used for the NMR analysis. ^{1}H NMR spectra were recorded at 25 °C on a 500 MHz Bruker DMX-500 spectrometer (Bruker, Karlsruhe, Germany) operating at a proton NMR frequency of 500.13 MHz. CH_3OH-d_4 was used as the internal lock. Each ^{1}H NMR spectrum consisted of 128 scans requiring 10 min and 26 s acquisition time with the following parameters: 0.16 Hz/point, pulse width (PW) = 30° (11.3 µs), and relaxation delay (RD) = 1.5 s. A pre-saturation sequence was used to suppress the residual H_2O signal with low power selective irradiation at the H_2O frequency during the recycle delay. FIDs were Fourier transformed with LB = 0.3 Hz. The resulting spectra were manually phased and baseline corrected, and calibrated to methanol at 3.3 ppm, using XWIN NMR (version 3.5, Bruker). 2D *J*-resolved NMR spectra were acquired using 600 MHz (Bruker DMX-600) 8 scans per 128 increments for F1 and 8 k for F2 using spectral widths of 5000 Hz in F2 (chemical shift axis) and 66 Hz in F1 (spin–spin coupling constant axis). A 1.5 s relaxation delay was employed, giving a total acquisition time of 56 min. Datasets were zero-filled to 512 points in F1 and both dimensions were multiplied by sine-bell functions (SSB = 0) prior to double complex FT. *J*-Resolved spectra tilted by 45°, were symmetrized about F1, and then calibrated, using XWIN NMR (version 3.5, Bruker). ^{1}H–^{1}H correlated spectroscopy (COSY) and heteronuclear multiple bonds coherence (HMBC) spectra were recorded on a 600 MHz Bruker DMX-600 spectrometer (Bruker). The COSY spectra were acquired with 1.0 s relaxation delay, 6361 Hz spectral width in both dimensions. Window function for COSY spectra was sine-bell (SSB = 0). The HSQC spectra were obtained with 1.0 s relaxation delay, 6361 Hz spectral width in F2 and 27,164 Hz in F1. Qsine (SSB = 2.0) was used for the window function of the HSQC. The HMBC spectra were recorded with the same parameters as the HSQC spectra except for 30,183 Hz of spectral width in F2. The optimized coupling constants for HSQC and HMBC were 145 Hz and 8 Hz, respectively.

Statistical analyses

The ^{1}H NMR spectra were automatically reduced to ASCII files. Spectral intensities were scaled to total intensity and reduced to integrated regions of equal width (δ 0.04) corresponding to the region of δ 0.0–10.0. The regions of δ 4.85–4.95 and δ 3.20–3.40 were excluded from the analysis because of the residual signal of D_2O and CD_3OD, respectively. Bucketing was performed by AMIX software (Bruker). Principal component analysis (PCA) based on Pareto scaling while projections to latent structures (PLS), orthogonal projections to latent structures (OPLS), and bidirectional orthogonal projections to latent structures (O2PLS) analysis based on UV scaling were performed with the SIMCA-P+ software (v. 12.0, Umetrics, Umeå, Sweden).

Results and Discussion

Extraction and visual analysis of ^{1}H NMR spectra

A total of fifty-eight different wine samples belonging to two wine types from two vintages were analyzed. The extraction procedure followed in this study was found quite effective as compared to most common methods in NMR based wine analyses. Methods like direct analysis of wine (Lopez-Rituerto et al. 2009), lyophilization (Baderschneider and Winterhalter 2001), and the use of nitrogen-flow (Amral and Caro 2005), are associated with low reproducibility, time consuming, and signal overlapping by dominant compounds in wine, e.g. ethanol, and shifting of signals due to differences in pH of wine, respectively. The ethyl acetate extraction used in this study was found to be fast, does not require a pH control, and has less signal intensity for wine dominating metabolites like ethanol and lactic acid. Around fifty metabolites were identified in the different wine samples using different 2D NMR techniques like *J*-resolved, COSY (correlated spectroscopy), HMBC (heteronuclear multiple bonds coherence), and HSQC (heteronuclear single quantum coherence) spectra. The metabolites identified cover a wide diversity including amino acids, organic acids, carbohydrates, hydroxycinnamates, hydroxybenzoates, stilbenes, and flavonoids. All assignments were done by comparing the spectra with previous reports (Ali et al. 2009; Liang et al. 2006), and with 1D and 2D NMR spectra of more than 500 common metabolites in our in-house library.

Conventionally a ^{1}H NMR spectrum is divided into three regions. The area between δ 0.8-4.0 corresponds to amino acids with few organic acids signals as well. The region of δ 4.0-5.5 is considered to be the region for anomeric protons of carbohydrate and the

remaining part, i.e. δ 5.5-8.5, is known as the phenolic region. Figure 1 presents a comparison of ^{1}H NMR spectra of two different wines, 'Riesling' and Mueller-Thurgau (from the 2007 vintage), showing differences not only in the phenolic contents but also in the organic and amino acids profile of the wines. Different vintages of 'Riesling' wine were also compared and the spectral differences among the wine types and vintages are clearly observed and are discussed below.

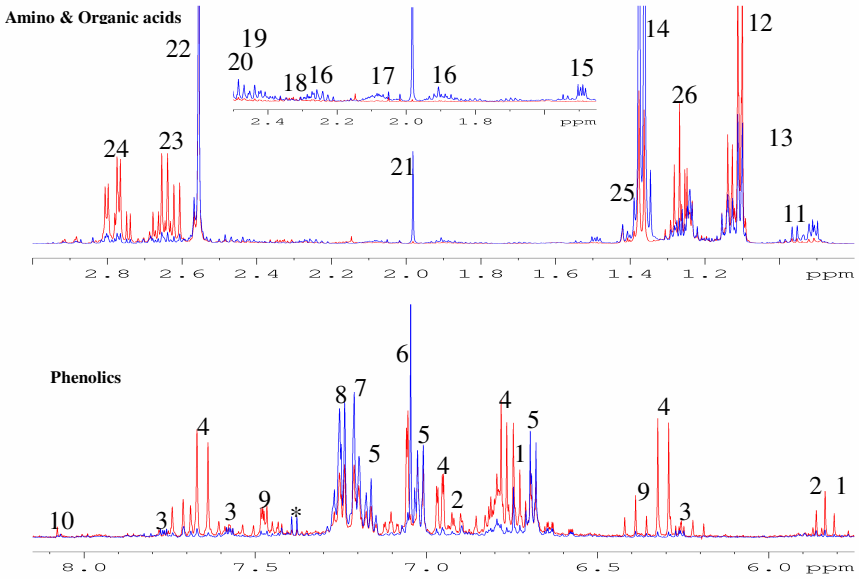

Figure 1. ^{1}H NMR spectra of 'Riesling' (red) and 'Mueller-Thurgau' (blue) from 2007 vintage. 1: (+)-catechin, 2: (-)-epicatechin, 3: quercetin, 4: caffeoyl moiety, 5: *cis*-resveratrol, 6: gallic acid, 7: tyrosine, 8: phenylalanine, 9: coumaroyl moiety, 10: kaempferol, 11: leucine, 12: valine, 13: 2,3-butanediol, 14: threonine, 15: alanine, 16: GABA, 17: methionine, 18: proline, 19: glutamate, 20: glutamic acid, 21: acetate, 22: succinate, 23: citrate, 24: malate, 25: lactate, 26: ethanol.

Compound identification

The use of NMR in metabolomics studies offers many advantages and is considered as a powerful analytical tool for such studies. A problem associated with NMR based metabolomics is the congestion of signals in the spectra especially in case of complex mixtures like wine. This signal overlapping represents the major difficulty in compound identification. This problem is overcome by the use of different 2D techniques like *J*-resolved, ^{1}H-^{1}H COSY, ^{1}H-^{13}C HMBC, and ^{1}H-^{13}C HSQC, which provide additional information regarding the structure of metabolites. Among the 2D NMR techniques

used in this study, the *J*-resolved is characterized by short measuring time and good quantitative features as compared to other 2D NMR experiments (Viant 2003). The ^{1}H-^{1}H COSY was found very effective in terms of confirming some characteristic correlations among the protons, especially for phenylpropanoids and flavonoids as shown in many metabolomics studies (Ali et al. 2009; Liang et al 2006). The ^{13}C-related NMR techniques, like ^{1}H-^{13}C HMBC, and ^{1}H-^{13}C HSQC, are not very common in metabolomics studies due to their long measuring time but recently the potential of these techniques in metabolomics has been discussed (Hyberts et al. 2007; Lewis et al. 2007). The assignments of signals to various metabolites identified in the present study are based on our in-house library of NMR data of common metabolites.

Amino acids, carbohydrates, and organic acids in wine

The amino acids not only contribute to wine taste and appearance (Hernendez-Orte et al. 2002) but also have an influence on aromas during the maturing process (Escudero et al. 2000). In wine, nearly all the sugars (from grapes) are consumed by the yeast during fermentation unless fermentation is arrested intentionally to make fortified, dry, late harvested or semisweet wines. Organic acids composition in wine and grape juice is of high importance as it influences the organoleptic properties, involved in control of microbiological growth, and is a critical parameter in wine stabilization. These acids originated directly from the grapes or are formed in processes like alcoholic fermentation, oxidation of ethanol, and malolactic fermentation. Compounds like tartarate, malate, citrate (from grapes) along with succinate, oxalate, fumarate, citrate (from fermentation) are known to influence the pH of wine. Small amounts of other acids like galacturonate, citramalate, pyruvate, and ketoglutarate, are also present in wine (Castellari et al. 2000).

The high signal intensities in the amino acid region were helpful to elucidate a number of amino and organic acid signals. Amino acids like alanine, threonine, valine, proline, methionine, tyrosine, phenylalanine, glutamine, glutamate, arginine, and aspartate, were identified in wine by comparison with the reference spectra of these compounds. The signals in the carbohydrate regions were highly clustered and overlapped. This region showed the signals of the anomeric protons of β-glucose, α-glucose, and sucrose. Resonances for some other compounds like GABA, choline, glycerol, and 2,3-

butanediol were also identified in the same region. A number of signals have been elucidated as organic acids like acetic acid, succinic acid, fumaric acid, formic acid, citric acid, malic acid, lactic acid, and tartaric acid. The ^{1}H NMR chemical shifts for these metabolites are shown in Table 1.

Table 1 ^{1}H NMR chemical shifts (δ) and coupling constants (Hz) of wine phenolics identified by references and using 1D and 2D NMR spectra.

Compounds	**Chemical Shifts (δ)**
Alanine	1.48 (d, *J*=7.4)
Threonine	1.32 (d, *J*=6.5), 3.51 (d, *J*=5.0), 4.27 (m)
Valine	1.01 (d, *J*=7.0), 1.06 (d, *J*=7.0), 2.28 (m)
Proline	2.35 (m), 3.37 (m)
Methionine	2.15 (m), 2.65 (t, *J*=8.0)
Tyrosine	6.85 (d, *J*=8.5), 7.19 (d, *J*=8.5)
Phenylalanine	3.15 (dd, *J*=8.2, 14.5), 3.91 (t, *J*=9.6)
Glutamine	2.14 (m), 2.41 (td, *J*=16.2, 7.5)
Glutamate	2.13 (m), 2.42 (m), 3.71 (dd, *J*=7.0, 1.9)
Arginine	1.75 (m), 3.75 (t, *J*=5.5)
Aspartate	2.80 (m), 3.80 (m)
β-glucose	4.58 (d, *J* = 7.8)
α-glucose	5.17 (d, *J* = 3.7)
Sucrose	5.39 (d, *J* = 3.9)
GABA	1.90 (m), 2.31(t, *J*=7.5), 3.01 (t, *J*=7.5)
Choline	3.20 (s)
Glycerol	3.56 (m), 3.64 (m)
2,3-butanediol	1.14 (d, *J*=6.5)
Acetic acid	1.94 (s)
Succinic acid	2.53 (s)
Fumaric acid	6.52 (s)
Formic acid	8.45 (s)
Citric acid	2.56 (d, *J* = 17.6), 2.74 (d, *J* = 17.6)
Malic acid	2.68 (dd, *J* = 16.6, 6.6), 2.78 (dd, *J* = 16.6, 4.7), 4.34 (dd, *J* = 6.6, 4.7)
Lactic acid	1.40 (d, *J*=7.0)
Tartaric acid	4.35 (s)
cis-Resveratrol	6.21 (t, *J*=2.1), 6.31 (d, *J*=2.1), 6.68 (d, *J*=13.3), 6.71 (d, *J*=8.5), 6.76 (d, *J*=13.3), 7.18 (d, *J*=8.5)
Gallic acid	7.03 (s)
Syringic acid	3.89(s), 7.31(s)

Vanillic acid	3.90 (s), 6.77 (d, *J*=8.2), 7.22 (m)
p-Benzoic acid	6.83 (d, *J*=8.7), 7.94 (d, *J*=8.6)
p-Coumaric acid	6.38 (d, *J*=16.0), 6.84 (d, *J*= 8.8), 7.50 (d, *J*=8.8), 7.59 (d, *J*=16.0)
Caffeic acid	6.24 (d, *J*=16.0), 6.87 (d, *J*=8.4), 7.02 (dd, *J*=8.4, 2.0), 7.12 (d, *J*=2.0), 7.52 (d, *J*=16.0)
(+)- Catechin	2.52 (dd, *J*=16.1, 8.2), 2.83 (dd, *J*=16.0, 5.4), 4.04 (m), 4.55 (d, *J*=7.5), 5.89 (d, *J*=2.2), 6.75 (d, *J*=8.0), 6.80 (dd, *J*=8.5, 2.0), 6.88 (d, *J*=8.5), 6.9 (d, *J*=2.0)
(-)- Epicatechin	2.72 (dd, *J*=16.8, 2.6), 2.89 (dd, *J*=16.9, 4.6), 4.26 (m), 6.03 (d, *J*=2.0), 6.06 (d, *J*=2.0), 6.88 (brs), 6.96 (d, *J*=2.2)
Quercetin	6.27 (d, *J*= 2.0), 6.49 (d, *J*=2.0), 6.99 (d, *J*=8.6), 7.66 (dd, *J*=8.5, 2.0), 7.71 (d, *J*=2.0)
Kaempferol	6.28 (d, *J*= 2.0), 6.52 (d, *J*=2.0), 6.74 (d, *J*=8.6), 8.04 (d, *J*=8.6),
trans-Caftaric acid	5.77 (s), 6.29 (d, *J*=16.0), 6.88 (d, *J*=8.4), 7.02 (dd, *J*=8.4, 2.0), 7.12 (d, *J*=2.0), 7.52 (d, *J*=16.0)
trans-p-Coutaric acid	5.84 (s), 6.36 (d, *J*=16.0), 6.87 (d, *J*=8.8), 7.51 (d, *J*= 8.8), 7.59 (d, *J*=16.0)
cis-Caftaric acid	5.34 (s), 5.92 (d, *J*=13.0), 6.71 (d, *J*=8.4), 6.81 (d, *J*=13.0), 7.03 (dd, *J*=8.4, 2.0), 7.44 (d, *J*=2.0)
cis-p-Coutaric acid	5.41(s), 5.94 (d, *J*=13.0), 6.73 (d, *J*=9.2), 6.86 (d, *J*=13.0), 6.93 (d, *J*=9.2), 7.61 (d, *J*=9.2),

Phenolics in wine

Phenolics are a complex group of metabolites particularly contributing to the characteristics of grapes and wines. They are omnipresent and well known for their contribution in pigmentation of different organs along with their role against different biotic and abiotic stresses. In wine they usually are associated with appearance, taste, mouth-feel, to a certain extent fragrance, and antimicrobial activity (Kennedy 2008). Wine phenolics may arise from the fruit (skins and seeds), and/or may be the products of yeast metabolism and can be classified into three categories which include flavonoids, stilbenoids, and simple phenolics.

Flavonoids are synthesized from the combination of the shikimic and polyketide pathway and are found primarily in the berry skins and seeds of the grapes. They can be further divided into several sub-groups. The most common flavonoids in wine are flavonols (quercetin, kaempferol, myricetin etc.), flavan-3-ols (catechin and epicatechin), and anthocyanins. Polymerization of polyhydroxy flavan 3-ol units, (+)-catechin and (-)-epicatechin, and their gallate esters produces oligomers and polymers called proanthocyanidins (often referred to as procyanidins). Even though procyanidins

occur primarily as dimers in grapes (Kennedy 2008), they tend to be polymerized and predominate in wine as condensed tannins.

In the presented study, we successfully identified some flavonoids including both flavonols and flavan-3-ols. Flavonols like quercetin, kaempferol, and myricetin were also identified in the aromatic region. The quercetin signals at δ 6.49 of H-6 is correlated in the ^{1}H-^{1}H COSY spectrum with the signal at δ 6.27 of H-8 and a signal at δ 6.99 of H-5′ with one at δ 7.66 of H-6′. Similar correlations for the signals of kaempferol at δ 8.04 of H-2′ and H-6′ with δ 6.74 of H-3′ and H-5′, and at δ 6.52 of H-8 with δ 6.28 of H-6 were also observed. Likewise myricetin signals, δ 6.47 of H-8 with δ 6.25 of H-6, also showed ^{1}H-^{1}H COSY correlations. Compounds like (+)-catechin and (-)-epicatechin were also identified. For both the (+)-catechin and (-)-epicatechin, signals of H-6′ and H-5′ along with signals of H-6 and H-8 showed correlations in ^{1}H-^{1}H COSY spectra.

Stilbenoids is another class of phenolic compounds in grapes (non flavonoids) and attracted great interest because of their potential valuable health effects but occur in relatively low concentrations as compared to other phenolics (Guebaila et al. 2006). The stilbene contents in wines varied significantly due to factors like climate, grape variety, fungal infection (Perrone et al. 2007), UV light, metal ions (Pussa et al. 2006) and enological methods (Gambuti et al. 2004). Resveratrol (a monomeric stilbene) is the major stilbene of grapes, synthesized from phenylalanine and can exist in two isomeric forms. In wines, both isomeric forms were detected due to photochemical isomerization of *trans*- form into *cis*- form during wine making (Sun et al. 2003). Resonances like δ 6.21 (t, J = 2.1 Hz), δ 6.31 (d, J = 2.1 Hz), δ 6.68 (d, J = 13.3 Hz), δ 6.71 (d, J = 8.5 Hz), δ 6.76 (d, J = 13.3 Hz), and δ 7.18 (d, J = 8.5 Hz) are assigned to stilbenoid. The up-field shift of the olefinic protons signals, H-8: from δ 6.79 to δ 6.68, and H-7: from δ 6.89 to δ 6.76, with the reduced coupling constants (from 16.1 Hz to 13.2 Hz), suggests that this compound is the *cis*- isomer of resveratrol. The ^{1}H-^{1}H COSY spectrum showed correlations among the signals of these olefinic protons along with other correlations like H-4 (δ 6.21) with H-2 and H-6 (δ 6.30), and also between H-6′ (δ 7.18) with H-3′ (δ 6.71).

Simple phenolics of grape origin are initially biosynthesized from phenylalanine, whereas those of yeast origin are derived from acetic acid. Simple phenolics are

derivatives of hydroxycinnamic acids (HCA) and hydroxybenzoic acids (HBA). Primarily, they are stored in cell vacuoles of grape cells and are released by crushing. In this study, we have identified major HBAs and HCAs along with their esterified forms. When compared with HCA, wines have lower levels of hydroxybenzoic acids (HBAs). The most common HBAs are gallic acid, gentisic acid, protocatechuic acid, and *p*-hydroxybenzoic acid, which are mainly found in their free form (Pozo-Bayon et al. 2003). In wine the presence of gallic acid has been shown by many studies (Sladkoovsky et al. 2004) along with other compounds like protocatechuic, vanillic, and syringic acids (Baderschneider and Winterhalter 2001). The 2D-*J* resolved and ^{1}H-^{1}H COSY spectra are shown in Fiure 2A and B, respectively.

The aromatic part of the ^{1}H NMR spectra of wine show some signals of hydroxybenzoates like gallic acid, salicylic acid, syringic acid, vanillic acid, and *p*-hydroxybenzoic acid. Resonances for gallic acid, syringic acid, salicylic acid, vanillic acid, and *p*-hydroxybenzoic acid are also observed. ^{1}H-^{1}H COSY correlations for these compounds have been observed for the signals like δ 7.22 with δ 6.77 of vanillic acid, and δ 7.94 with δ 6.83 of *p*-hydroxybenzoic acid.

The concentrations of different HCAs depend on factors like grape variety, growing conditions, and climate. Most common hydroxycinnamic acids reported in wine are *p*-coumaric acid, caffeic acid, sinapic acid, and ferulic acid. The esterified form of HCA with tartaric acid is more common in wine than free HCA. The caftaric acid (caffeic acid conjugated with tartaric acid) is predominant (up to 50%) of the total HCA (Ricardo-Da-Silva et al. 1993). Evidence of the presence of glucosides like 4-*O*-glucosides of ferulic- and coumaric acid (both *cis*- and *trans*- forms), glucosides of caffeic-, coumaric-, ferulic-, and sinapic acid have been shown in some studies (Baderschneider and Winterhalter 2001; Monagas et al. 2005).

Characteristic doublets with a coupling constant of 16.0 Hz in the range of δ 6.20-6.41 and δ 7.51-7.75 represents the H-8′ and H-7′ (olefinic protons) of phenylpropanoids, respectively. The ^{1}H-^{1}H COSY spectra also confirmed the correlation between H-8′ and H-7′ of phenylpropanoids, by the coupling with the carbonyl carbon at δ 168.3 in the HMBC spectra. In ^{1}H NMR spectra of wine samples, these resonances are assigned to four different hydroxycinnamic acids which include caffeic acid, coumaric acid, ferulic acid, and sinapic acid (Table 1). The ^{1}H-^{1}H COSY spectra show correlations among

signals like δ 6.92 with δ 6.82 and δ 7.03 of caffeic acid; δ 7.50 with δ 6.84 of coumaric acid; δ 7.09 with δ 6.79 and δ 7.19 of ferulic acid.

Tartaric esters of hydroxycinnamic acid were also identified using 1D and 2D NMR spectra of wine samples. Three singlets for tartaric acid were observed in the region of δ 5.77-5.84 in ^{1}H NMR spectrum. These singlets are downfield shifted from the typical tartaric acid signal at δ 4.80 due to their bonding to the carboxylic function of cinnamates which is confirmed by their correlation with the signal at δ 167.5-168.5 in the HMBC spectrum. The signals of olefinic protons (H-8′ and H-7′) were also slightly shifted. Based on these assignments, three compounds were identified which include caftaric acid (caffeic acid conjugated with tartaric acid), fertaric acid (ferulic acid conjugated with tartaric acid), and *p*-coutaric acid (coumaric acid conjugated with tartaric acid) (Table 1).

Along with the *trans*- forms, the *cis*- forms of these conjugated cinnamic acids, i.e. *cis*-caftaric acid and *cis*-coutaric acid, were also detected. When compared to their *trans*-configuration, the *cis*- forms showed the up-field shifting of the signals for their H-8′ and H-7′ along with the reduction in the coupling constants from 16.0 Hz to 13.0 Hz. Two clear doublets of 13.0 Hz at δ 5.92 and δ 5.94 were detected for the H-8′ of *cis*-configuration. The ^{1}H-^{1}H COSY spectra also confirmed this by showing the correlation of these signals with their respective H-7′ protons at δ 6.81 and δ 6.86. It was also confirmed by the correlation of this signal with the carbonyl resonance at δ 167.2 in the HMBC spectra.

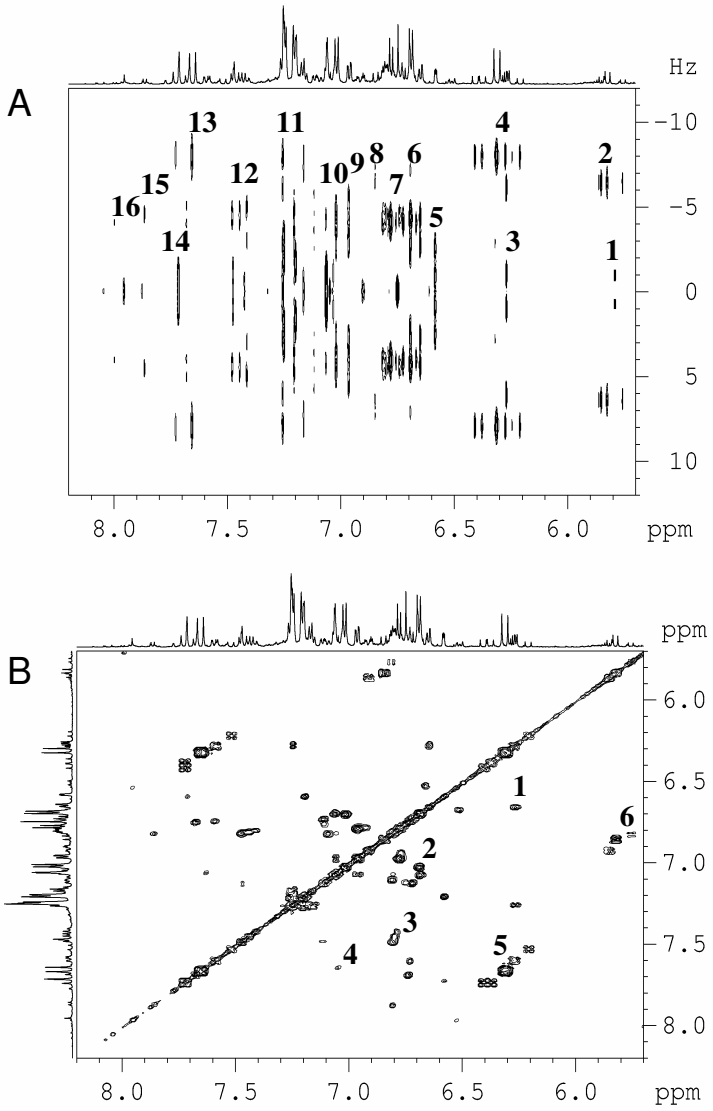

Figure 2. Two dimensional ^{1}H-^{1}H *J*-resolved (A) and ^{1}H-^{1}H COSY (B) spectra of 'Riesling' wine in the range of δ 5.5-δ 8.5. *J*-resolved (A) shows 1: H-6 of flavan-3-ols, 2: H-8 of *cis*-phenylpropanoids, 3: H-6 of flavonols, 4: H-8 of *trans*-phenylpropanoids, 5: H-2 & H-6 of *cis*-resveratrol, 6: H-8 of *cis*-resveratrol, 7: H-5 of phenylpropanoids, 8: H-7 of *cis*-phenylpropanoids, 9 & 10: H-6 of phenylpropanoids, 11: 1H of phenylalanine, 12: H-6 of *p*-coumaric acid, 13: H-7 *trans*-phenylpropanoids, 14: H-2 of flavonols, 15: H-2 & H-6 of *p*-benzoic acid, 16: H-2 of kaempferol. COSY (B) shows correlations between 1: H-6 and H-8 of quercetin, 2: H-5 and H-6 of phenylpropanoids, and H-7 and H-8 of resveratrol, 3: H-5 and H-6 of *p*-coumaric acid, 4: H-5 and H-6 of quercetin, 5: H-7 and H-8 of *trans*-phenylpropanoids, 6: H-7 and H-8 of *cis*-phenylpropanoids.

Multivariate data analyses (MvDA)

The ^{1}H NMR data were further subjected to principal component analysis (PCA) but before that data were mean centered and then scaled using Pareto scaling. Unlike UV scaling, in which standard deviation is used as the scaling factor, Pareto scaling is characterized as a method in between UV scaling and no scaling and uses the square root of standard deviation as the scaling factor. ^{1}H NMR data (δ 0.32-10.0) of both wines were first reduced by PCA and the results are projected in a score plot (data not shown). The PCA failed to highlight any difference based on quality scores and no grouping of samples was observed due to very high variability among the different wine samples.

In order to emphasize the differences among the wine samples, we used PLS on the same NMR data. It is a projection method that, by using a linear multivariate model, can relate the two blocks of variables, i.e. *X* and *Y*. In our case *X*-variables are the values obtained after bucketing the ^{1}H NMR spectra and the *Y*-variables are the quality scores corresponding to each wine sample. The score plot showed (Figure 3A) a pattern of arrangement of wine samples according to their quality scores, from inferior to superior tasting wines arranged from negative to positive side of component 1 (52.9%), respectively.

By examining the loading column plot (Figure 3B), several compounds were found participating in the discrimination of wines. The superior wines were found containing higher levels of amino acids like proline and arginine along with 2,3-butanediol. Organic acids like malic and tartaric acids were also found discriminating for the higher quality scores wine along with some phenolics like quercetin, (+)-catechin, and (-)-epicatechin. For the poor taste wines, resonances for the lactic, acetic, and succinic acid were found higher with threonine and alanine. Among the phenolics caffeic acid, gallic acid, and vanillic acid signals showed higher intensity in low taste wines.

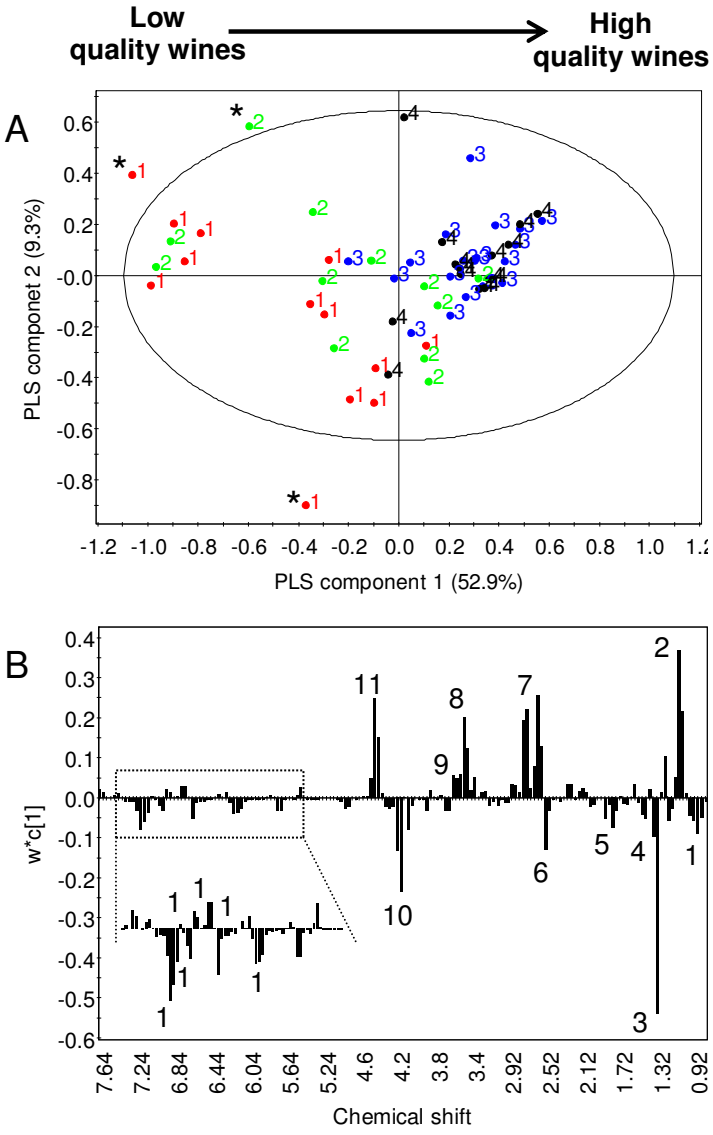

Figure 3. Projections to latent structures (PLS) score (A) and loading line (B) plots. The score plot (A) shows arrangement of wine samples from low to high quality along component-1. Numbers (1 to 4) indicate the class to the wine sample belongs. Samples with '*' are outliers. The column plot shows higher levels of compounds like 2: 2,3-butanediol, 7: malate, 8: proline, 9: arginine, 11: tartarate, 13, quercetin, 14: (+)- catechin, and 15: (-)- epicatechin, where as metabolites like 1: valine and leucine, 3: lactate, 4: alanine, 5: acetate, 6: succinate, 10: threonine, 12: caffeoyl moiety, 16: gallate, 17: vanillate, were found discriminating for low quality wines.

Finally for the identification of quality markers for high quality wine, we used another method known as bidirectional orthogonal projections to latent structures (O2PLS). This method removes the structured noise by extracting linear relationships from *X* and *Y* data blocks (Trygg and Wold 2002; 2003). Analysis like PLS regression can cause systematic variation of any data block due to structured noise present in the data blocks. The structures noise can be removed from *X* and *Y*, in a bidirectional way, by O2PLS regression and results in the decomposition of systematic variation into two model parts: the predictive or parallel part and the orthogonal part. The regression has been validated using cross validation-analysis of variance (CV-ANOVA), with a p-value < 0.001. As indicated by Figure 4A (the score plot of O2PLS regression) the wine samples are arranged according to their quality grade (almost same as PLS). For the identification of markers for the high quality wines, we use S-plot (Figure 4B) and compounds like 2,3-butanediol, malate, proline, and GABA were found as markers for good taste. For low quality wines, signals of lactate and alanine were identified.

Proline is the most prominent amino acid in wine. The higher proline content is mainly because of the inability of yeast to use it as a nitrogen source and may be important to give wine the apparent mouth feel or body. This is because a salivary protein, which has a strong affinity for polyphenols, is rich in proline (Bennick 1982). 2,3-butanediol is the major dialcohol in wine, produced from pyruvate or acetoin, and contributes in wine sensory attributes with a slightly bitter taste and viscous body. Higher levels of 2,3-butanediol indicate reduction of acetoin which has a negative impact on wine taste with low taste and odor threshold (Cambon et al. 2006; Romano et al. 1996).

The levels of organic acids like malate, lactate, tartarate, succinate, and citrate, mainly depend upon the grape cultivar and/or fermentation process. The malate, tartarate, and citrate, are the three major acids of grapes. The citrate and lactate is formed by malolactic fermentation while succinate is formed by alcoholic fermentation during the wine making process. Increased levels of malic or lactic acid have been associated with increased perception of sourness and astringency (pH dependent) and may be also with a milkier taste of wine (Son et al. 2008). Though the high molecular weight polyphenolics mainly account for color and astringency of wine but the presence of these organic acids apparently augments this (Hufnagel and Hofmann 2008).

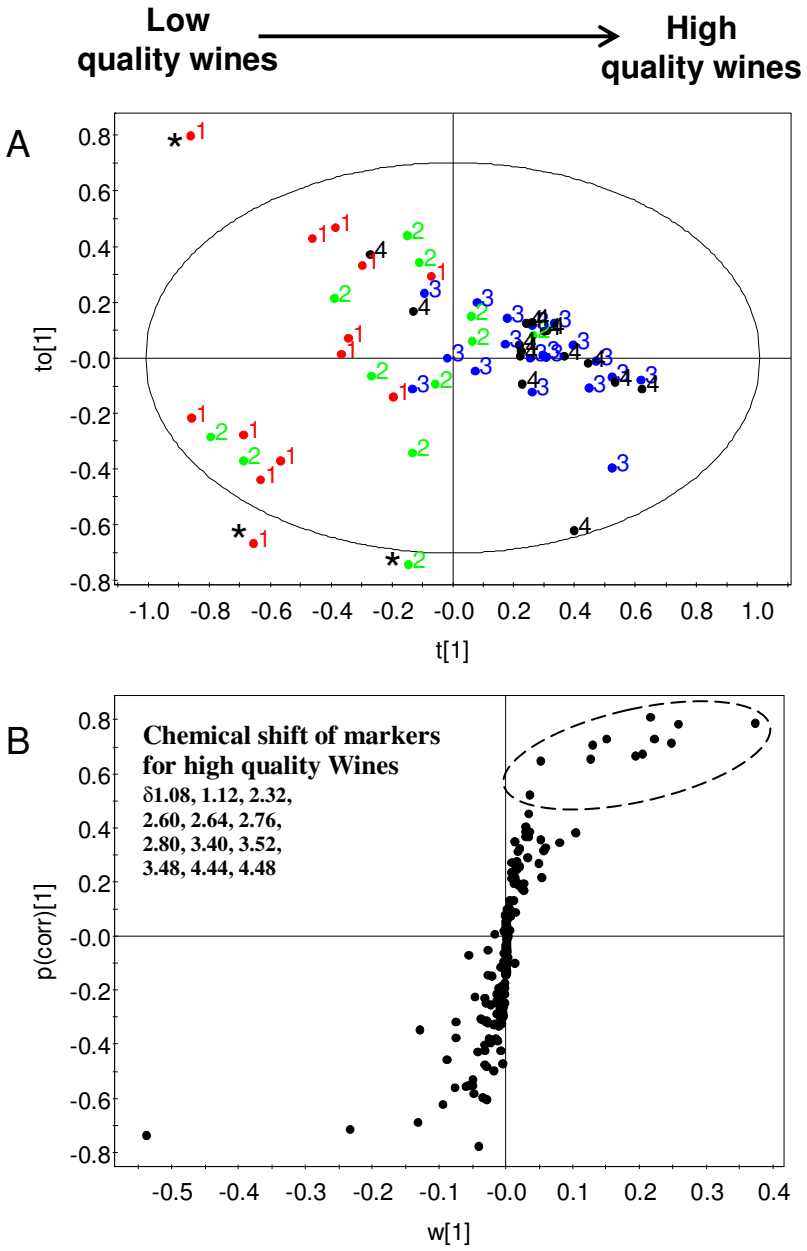

Figure 4. Bidirectional orthogonal projections to latent structures (O2PLS) score (A) and loading S-plot (B). The score plot (A) shows arrangement of wine samples from low to high quality. Samples with '*' are outliers. The S-plot (B) shows markers (buckets) for the high quality wine.

Wine phenolics and amino acids content can be used for the classification of wine (Bevin et al. 2008; Bennick 1982; Anastasiadi et al. 2009; Kosir and Kidric 2001). This study also resulted in characteristic phenolics and amino acids profiles of 'Riesling' and 'Mueller-Thurgau' wines, dominated by different sets of metabolites. As we have used two types of wine from two different vintages, another multivariate analysis method, i.e. bidirectional orthogonal projections to latent structures (O2PLS), was used to highlight differences among the samples based on grape cultivar and vintage. The samples were classified into four classes and each class was represented by samples from the same year of same cultivar. By examining the score plot (Figure 5), it can be easily observed that not only 'Riesling' was separated from 'Mueller-Thurgau', but also the samples from 2006 were distinct from the samples of 2007, in case of both the wine types. Figure 1 also highlights the different metabolic profiles of 'Riesling' and 'Mueller-Thurgau' wines. The loading plot reveals that among the phenolics compounds like (+)-catechin, (-)-epicatechin, caftarate, and coutarate are present in higher levels in 'Riesling' while resonances for quercetin, kaempferol, resveratrol, and gallate were more intense in 'Mueller-Thurgau'. The amino and organic acid profiles of both wines were also found quite distinctive. Compounds like leucine, alanine, threonine, glutamine, and GABA, were more concentrated in 'Mueller-Thurgau' with some organic acids like lactate, acetate, and succinate. 'Riesling' was characterized by higher resonances for the amino acids like valine, proline, 2,3 butanediol, along with malate and citrate.

The vintage effect on metabolic profile of grapes and ultimately on wine has been extensively studied (Pereira et al. 2005; Lee et al. 2009) and the importance of the vintage is widely accepted. The amino acids and polyphenols contents are highly affected by the climatic conditions of a grape production area. It has been shown that a hot and dry climate results in a higher proline and phenolic contents in wine (Lee et al. 2009). Our findings also support the effects of vintage on the wine metabolome as samples from different vintages show different metabolic constituents. In the case of both 'Riesling' and 'Mueller-Thurgau', the vintage effect was also evident as indicated by the Figure 5. By examining the loading plot, resonances responsible for the differentiation were identified and annotated. For the vintage of 2006, both these wines have higher levels of metabolites like leucine, phenylalanine, malate, acetate, citrate,

tartarate, choline, GABA, along with higher levels of phenolics. The 2007 vintage was characterized mostly by higher levels of amino acids like valine, proline, alanine, arginine, threonine, together with organic acids like succinate and lactate.

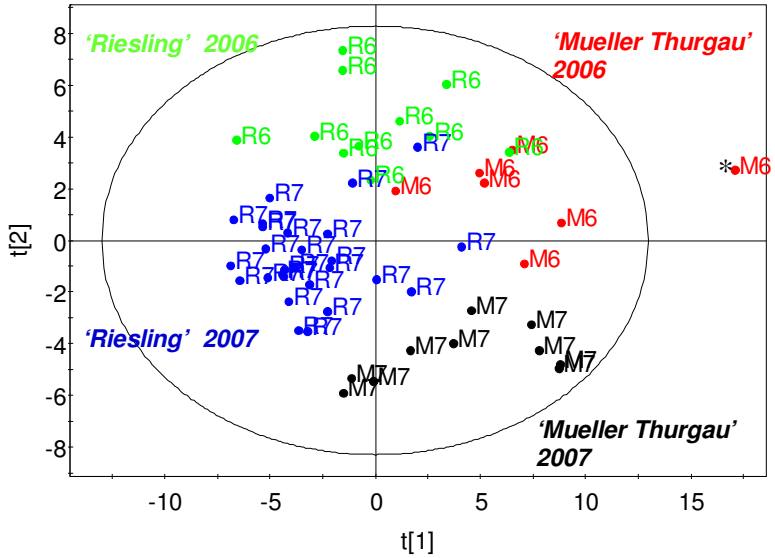

Figure 5. Bidirectional Orthogonal projections to latent structures (O2PLS) score plot. The score plot clearly indicates differentiation among the samples based on wine types and vintage. The samples M6 with '*' is an outlier.

In general, the enormous potential of NMR for the chemical analyses of complex food items has attracted the interests of many research groups around the world. One- and two-dimensional NMR experiments have been already used for wine (Son et al. 2009b) and grape berries (Pereira et al. 2005) classification and characterization but very few reports are available for the quantification of wine components using intensities from ^{1}H NMR spectra (Larsen et al. 2006; Viggiani and Morelli 2008). Most of the papers regarding the NMR analysis of wine either include some chromatography to identify the phenolic compounds or simply ignore the phenolics identification mainly because of the complexity of the spectra. In this study, ^{1}H NMR has been combined with different 2D NMR techniques, like *J*-resolved, COSY, HMBC, and HSQC, and resulted in successful identification of nearly all the major classes of phenolics, along with amino acids, carbohydrates, and organic acids. This approach has proved quite powerful in

identification of cinnamic acids, benzoic acid, stilbenoids, flavonols, and flavane-3-ols, along with differentiating in *trans*- and *cis*- forms.

Wine is considered as one of the most complex foods regarding the metabolic contents. The chemistry of wine is known to be highly influenced by many factors which include grape variety, climate, cultural practices, yeast strain, and the fermentation process as well (Rodríguez-Delgado et al. 2002). Several wine characteristics, like mouth feel, aroma, and color, are directly affected by the grape cultivars and many compounds can serve as a marker for wine discrimination (Stines et al. 2000). The present study is an attempt to underline those metabolic differences and link them with sensory analysis. The NMR spectroscopy in conjugation with multivariate data analysis was not only proven to be effective in identification of different compounds in wine but also in highlighting the differences among the wines belonging to different wine types and vintages and which have different quality grades.

Conclusion

In conclusion, ^{1}H NMR based metabolic profiling was shown to be effective in terms of identifying diverse metabolites, including amino and organic acids, carbohydrates, hydroxycinnamates, hydroxybenzoates, and flavonoids, in the complex mixtures present in wine. The extraction method used was found short, simple, and less time consuming with high reproducibility. In combination with ^{1}H NMR, PLS and O2PLS were found very effective in highlighting the differences among the wines based on quality scores. By using loadings of the score plot, many metabolites responsible for that discrimination have been identified. The effect of vintages on metabolic profiles of wine has also been studied (with OPLS) and clear differences were observed among the samples of different vintages, showing the effects of the climate. It should be mentioned that in terms of quality the interaction among the various components of wine are amazingly complex and knowledge regarding individual components might not be that much useful as such, increasing our insight in their interactions should be of interest. Metabolomics can not only offer to visualize the complexity of these specific interactions and provide better understanding of wines but more importantly how to make good quality wines.