JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

J. Agric. Food Chem. **2003,** *51,* 5715–5722 **5715**

Classification of Edible Oils by Employing ^{31}P and ^{1}H NMR Spectroscopy in Combination with Multivariate Statistical Analysis. A Proposal for the Detection of Seed Oil Adulteration in Virgin Olive Oils

GEORGIA VIGLI, ANGELOS PHILIPPIDIS, APOSTOLOS SPYROS, AND PHOTIS DAIS*

NMR Laboratory, Department of Chemistry, University of Crete, 714 09 Iraklion, Crete, Greece

A combination of ^{1}H NMR and ^{31}P NMR spectroscopy and multivariate statistical analysis was used to classify 192 samples from 13 types of vegetable oils, namely, hazelnut, sunflower, corn, soybean, sesame, walnut, rapeseed, almond, palm, groundnut, safflower, coconut, and virgin olive oils from various regions of Greece. 1,2-Diglycerides, 1,3-diglycerides, the ratio of 1,2-diglycerides to total diglycerides, acidity, iodine value, and fatty acid composition determined upon analysis of the respective ^{1}H NMR and ^{31}P NMR spectra were selected as variables to establish a classification/prediction model by employing discriminant analysis. This model, obtained from the training set of 128 samples, resulted in a significant discrimination among the different classes of oils, whereas 100% of correct validated assignments for 64 samples were obtained. Different artificial mixtures of olive−hazelnut, olive−corn, olive−sunflower, and olive−soybean oils were prepared and analyzed by ^{1}H NMR and ^{31}P NMR spectroscopy. Subsequent discriminant analysis of the data allowed detection of adulteration as low as 5% w/w, provided that fresh virgin olive oil samples were used, as reflected by their high 1,2-diglycerides to total diglycerides ratio ($D \geq 0.90$).

KEYWORDS: ^{1}H NMR and ^{31}P NMR spectroscopy; vegetable oils; discriminant analysis; adulteration of olive oil

INTRODUCTION

Recently, there has been an increasing interest in the classification of edible oils as an effective means to examine authentication and to detect possible adulteration of virgin olive oils with seed oils or low-quality olive oils (*1*, *2*). Such studies are clustered according to the type of analytical parameters (variables) used and the kind of chemometric procedures applied. Classical methods based on gas chromatography (GC) and high-performance liquid chromatography (HPLC) have been successfully used to classify and authenticate the edible oils (*1*−*5*). These methods are more efficient and offer better results when a large number of compositional parameters are analyzed by statistical procedures. Other studies have demonstrated that nonconventional methods, such as GC combined with mass spectrometry (*6*, *7*), infrared, and Raman spectroscopy (*8*−*10*), are able to offer good results in combination with chemometric procedures.

In the past decade, high-resolution nuclear magnetic resonance (NMR) spectroscopy has been applied effectively in the analysis of virgin olive oil to evaluate quality and authentication (*11*, *12*). However, studies that use NMR spectroscopic data of olive oil composition in combination with multivariate analysis of the chemical parameters are rather limited. Fauhl et al. (*13*) have separated virgin olive, sunflower, and hazelnut oils by selecting the peak intensities of the unsaturated fatty acids and polyunsaturated fatty acids, the ratio of linolenic to linoleic acid, and the ratio of linolenic acid to all other fatty acids. Mannina et al. (*14*) described a combined approach by using ^{1}H NMR and GC for the separation of virgin olive oil and hazelnut oil. Mavromoustakos et al. (*15*) reported a method based on ^{13}C NMR spectroscopy to detect the adulteration of virgin olive oil by cottonseed, sunflower, corn, and soybean oils. Vlachov (*16*) detected olive oil adulteration with soybean oil by employing the ^{13}C NMR pulsed sequence DEPT. Finally, Zamora et al. (*17*) classified virgin olive oils from different cultivars and regions of Europe and North Africa, refined olive oils, pomace oils, hazelnut, rapeseed, corn, grape seed, sunflower, and soybean oils on the basis of ^{13}C NMR chemical shifts and linear discriminant analysis. The oil samples used in this study were chromatographically fractionated to suppress the abundant triglycerides and enhance the concentration of minor components that define authenticity and quality (*18*).

In a recent publication (*19*) we introduced a facile NMR method to determine the mono- and diglyceride composition in olive oils. This method is based on the derivatization of the labile hydrogens of the hydroxyl groups of the diglycerides with 2-chloro-4,4,5,5-tetramethyldioxaphospholane (**1**) according to the reaction shown in **Figure 1** and the use of the ^{31}P chemical shifts of the phosphitylated compounds **2** to identify the labile

* Author to whom correspondence should be addressed (e-mail dais@chemistry.uoc.gr).

10.1021/jf030100z CCC: $25.00
Published on Web 08/12/2003

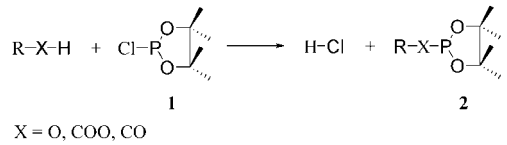

Figure 1. Reaction of hydroxyl groups of diglycerides with 2-chloro-4,4,5,5-tetramethyldioxaphospholane (**1**).

centers. Compound **1** reacts rapidly and quantitatively under mild conditions with the hydroxyl groups. It should be noted that reagent **1** also reacts quantitatively with other functional groups, such as carboxylic acids and aldehydes, replacing the carboxylic and aldehydic hydrogens according to the reaction in **Figure 1**.

This analytical approach was applied to a large number of virgin olive oils from various regions of Greece. as well as to several commercial virgin olive oils, refined olive oils, and pomace oils, in an attempt to examine virgin olive oil quality and freshness (*20*). It has been suggested (*11*, *20*−*22*) that the ratio of 1,2-diglycerides to the total diglycerides [D = 1,2-diglycerides/(1,2-diglycerides + 1,3-diglycerides)] is a useful index to detect the quality and freshness of virgin olive oils. Moreover, the same index can distinguish fresh virgin olive oils from refined oils, because the D ratio of the latter oils is much smaller (∼0.33) than that of the former. This is because the isomerization of 1,2-diglycerides (1,2-DGs) to 1,3-diglycerides (1,3-DGs) that usually occurs during prolonged olive oil storage is very rapid upon olive refinement.

The diglyceride concentration in virgin olive oils and other compositional parameters determined by ^{31}P NMR and ^{1}H NMR spectroscopy, respectively, were used in the present study to classify a large number of vegetable oils including virgin olive oils and to detect possible adulteration of virgin olive oils with other types of oils to limits below which the addition of seed oils becomes unprofitable. This was achieved by employing multivariate statistical procedures.

MATERIALS AND METHODS

Samples and Reagents. Fifty-six olive oil samples (years of harvesting 2000−2001 and 2001−2002) were collected from various regions of Greece through local agricultural cooperatives. In particular, 45 samples were collected from Crete (Peza, Kolymbari, Heraklion, and Sitia), 4 from Lesvos, 5 from Ilia, and 2 from Halkidiki. All samples considered in this study were virgin olive oils according to the official analytical methods and limits (*23*, *24*). A total of 136 samples of 12 types of refined seed oils were purchased from Henry Lamotte GmbH (Bremen, Germany). All samples from Henry Lamotte were certified and used without further analysis. These oils were as follows (the number of samples analyzed for each type of oil is reported in parentheses): hazelnut (16), sunflower (12), corn (16), soybean (14), sesame (13), walnut (14), rapeseed (5), almond (10), palm (7), groundnut (9), coconut (10), and safflower (10). All samples were stored in brown screw-capped bottles at −20 °C prior to spectrum acquisition. For adulteration studies, fresh virgin olive oils were mixed with hazelnut, sunflower, corn, and soybean oils. Mixtures of 5, 10, 15, 20, 35, and 50% w/w for each seed oil in virgin olive oils were prepared. Fresh virgin olive oil samples of different geographical origin and different samples of seed oils were used for the preparation of the various mixtures. One additional set of mixtures (5, 10, 15, and 20%) of different olive oils and hazelnut oils was prepared for the present analysis. All adulterated virgin olive oils were analyzed twice.

All protonated and deuterated solvents were of analytical grade and were purchased from Sigma-Aldrich Ltd. (Athens, Greece). The derivatizing reagent **1** was synthesized from pinacol and phosphorus trichloride in the presence of triethylamine following the method previously described in the literature (*25*). For synthesis of reagent **1** we utilized hexane solvent instead of benzene, as suggested in the original method, in an attempt to increase the yield of the reaction. This modification resulted in ∼45% yield of the product against 19% obtained with the original method.

Determination of Diglycerides, Acidity, and Total Sterols by ^{13}P NMR Spectroscopy. Samples of the various oils and the mixtures of olive oils with seed oils were prepared as follows: A stock solution (10 mL) composed of pyridine and $CDCl_3$ in a 1.6:1.0 volume ratio containing 0.6 mg of chromium acetylacetonate, $Cr(acac)_3$ (0.165 μM), and 13.5 mg cyclohexanol (13.47 mM) was prepared and protected from moisture with 5A molecular sieves. One hundred and fifty milligrams of the oil sample was placed in a 5 mm NMR tube. The required volumes of stock solution (0.4 mL) and the reagent **1** (15 μL) were added. The reaction mixture was left to react for about 20 min at room temperature. Upon completion of the reaction, the solution was used to obtain the ^{31}P NMR spectra.

^{31}P NMR spectra were obtained on a Bruker AMX500 spectrometer operating at 202.2 MHz for the phosphorus-31 nucleus at room temperature. To neglect NOE effects, the inverse-gated decoupling technique was used. Typical parameters for quantitative studies were as follows: 90° pulse width, 12.5 μs; sweep width, 10 kHz; relaxation delay, 30 s; and memory size, 16K (zero-filled to 32K). Line broadening of 1 Hz was applied, and drift correction was performed prior to Fourier transform. Polynomial fourth-order baseline correction was performed before integration. For each spectrum 32 transients were acquired. All ^{31}P chemical shifts are reported relative to the product of the reaction of **1** with water (moisture contained in all samples), which gives a sharp signal in pyridine/$CDCl_3$ at 132.20 ppm. It should be noted that the presence of the paramagnetic metal center of $Cr(acac)_3$ in the samples lowers the relaxation times of the phosphorus nuclei, thus shortening significantly the duration of the measurements. **Figure 2** shows the 202.2 MHz ^{31}P NMR spectrum of a virgin olive oil sample. The excellent resolution of the ^{31}P chemical shifts permits a reliable detection of the phosphitylated 1,2-diglycerides, 1,3-diglycerides, and total sterols. The concentration of these components was determined upon integration of their corresponding signals (**Figure 2**) with respect to the integral of the signal owing to the phosphitylated internal standard (cyclohexanol). The acidity was obtained from the integral of the signal corresponding to the phosphitylated free fatty acids, which appears at δ 134.79.

Usually monoglycerides cannot be detected by employing this method due to the very low concentration of these minor constituents in olive oil (<0.3%). However, both monoglycerides have been detected in the polar part of olive oils along with certain polyphenols (unpublished results). The applicability of this method to quantitative analysis, as well as its reproducibility and repeatability, has been tested thoroughly in previous studies (*19*, *26*).

^{1}H NMR Spectra. ^{1}H NMR spectra were recorded on a Bruker AMX500 spectrometer operating at 500 MHz for the proton nucleus at room temperature. The phosphitylated oil samples used in ^{31}P NMR experiments were used to obtain ^{1}H NMR spectra with the following acquisition parameters: time domain, 32K; 90° pulse width, 9.3 μs; spectral width, 12 ppm; relaxation delay, 2 s. Sixteen scans and four dummy scans were accumulated for each free induction decay. Baseline correction was performed carefully by applying a polynomial fourth-order function in order to achieve a quantitative evaluation of all signals of interest. The spectra were acquired without spinning the NMR tube in order to avoid artificial signals, such as spinning sidebands of the first or higher order.

Repeatability and Reproducibility of the NMR Measurements. The NMR measurements do not depend very significantly on the spectrometer or the operator provided that the same experimental protocol is followed. However, they do depend on the signal-to-noise ratio (S/N) and possible overlapping of signals; the higher the S/N ratio of the signals in the NMR spectra, the lower the uncertainty of the quantitative measurements. The reproducibility and repeatability of the ^{31}P NMR method have been previously tested thoroughly (*19*, *26*). No significant differences were observed for the repeatability (1.12%) and reproducibility (2.36%) of the present ^{31}P NMR measurements relative to those obtained previously (*19*, *26*). With respect to the repeatability and the reproducibility of ^{1}H NMR experiments, we recorded 10 consecutive spectra, using the same sample of phosphitylated olive oil,

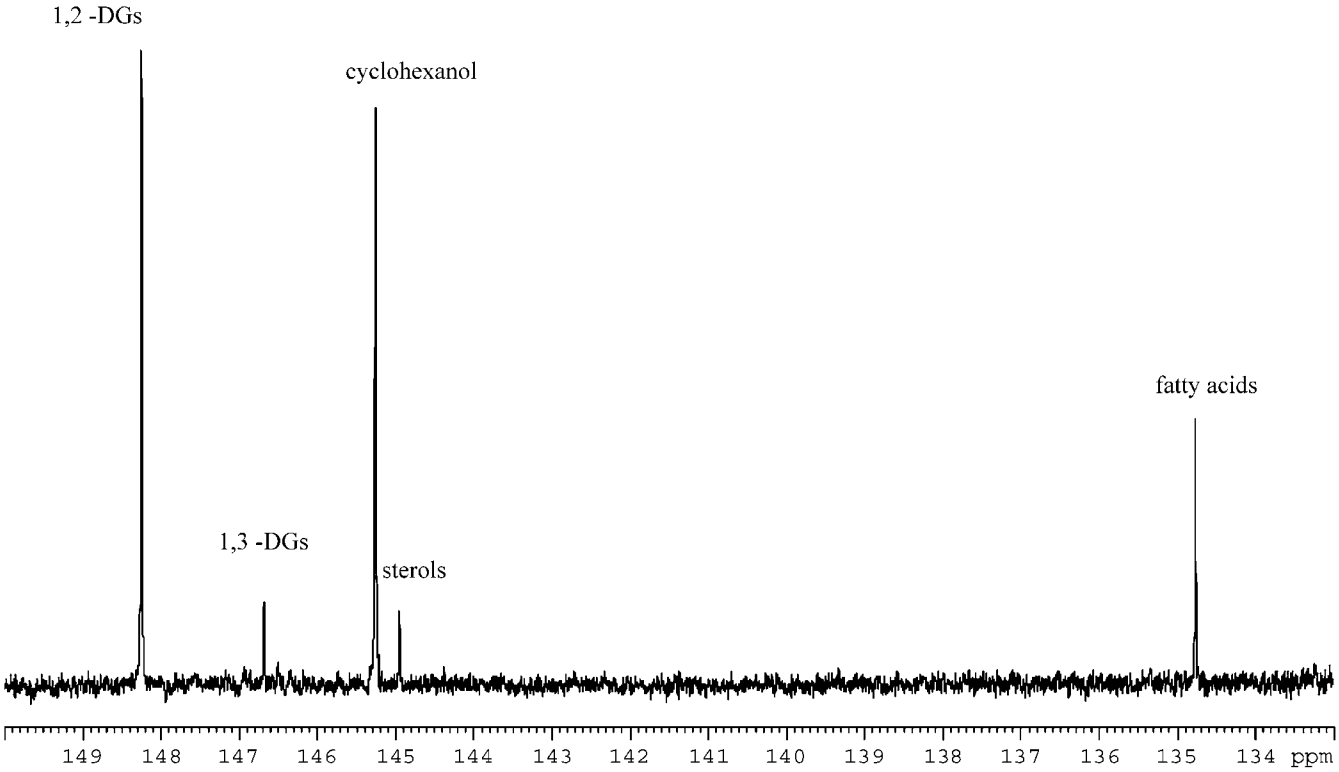

Figure 2. 202.2 MHz ^{31}P NMR spectrum of virgin olive oil. The region where the phosphitylated total sterols, diglycerides, and free fatty acids absorb is illustrated. The phosphitylated cyclohexanol is used as internal standard.

and performed eight measurements on different hazelnut oil samples, using the same experimental protocol for each measurement. The repeatability of the ^{1}H NMR measurements ranged from 0.40 to 6.69%, whereas the reproducibility was 0.35−11.47%, depending on the S/N of the various signals in the spectra. High repeatability and reproducibility were obtained in the ^{1}H NMR spectrum (**Figure 3**) for the olefinic protons (1.45 and 0.35%), signal A (1.45 and 1.54%), signal B (1.30 and 0.65%), signal C (0.77 and 0.44%), and signal F (0.65 and 0.65%). The lowest repeatability (6.69%) and reproducibility (11.47%) were calculated for signal E (linolenic acid) with the lowest S/N.

Statistical Data Analysis. Analysis of the data was performed by classical multivariate procedures including one-way ANOVA and the supervised technique discriminant analysis (DA). In the latter analysis, the total number of samples was randomly divided into two separate sets, a training set (around two-thirds of the samples) and a test set (around one-third of the samples). The first set was employed to establish a classification rule, which afterward allowed the attribution of unknown samples, whereas the second set was used to validate the predictive ability of the optimized statistical model. The canonical discrimination functions were built with the information of the training set, whereas validation of these functions was achieved through the test set. The same statistical treatment of the data was used to detect adulteration. The statistical data processing was performed by using Statistica for Windows 5.1B (StatSoft Inc.).

RESULTS AND DISCUSSION

Determination of Fatty Acid Composition and Iodine Value (IV). Figure 3 shows a typical virgin olive oil ^{1}H NMR spectrum from which the fatty acid composition and the IV can be determined. The spectrum is composed of 10 main signals, which have been assigned previously (*11*, *27*). The chemical shifts of these signals in the solvent pyridine/$CDCl_3$ are summarized in **Table 1**. Most of these signals are assigned to nonequivalent groups of protons that are common to fatty acyl chains. Therefore, the concentration of the acids in the oils can be calculated only by combination of various signal intensities in the ^{1}H NMR spectrum (*11*, *27*). An exception is signal *E* at δ 0.98, which corresponds to the methyl protons of linolenic acid. Therefore, the concentration of this acid is obtained from the relationship (*11*, *27*)

$$[\text{linolenic}] = \frac{E}{E+F} \quad (1)$$

F is the signal intensity of the methyl protons of all acids except linolenic acid. Oleic, linoleic, and total saturated fatty acids (mainly palmitic and stearic acids) (SFA) are calculated from signal intensities *A*, *C*, and *F* and the following set of relationships:

$$[\text{linoleic}] = \frac{3A - 4E}{2(E+F)} \quad (2)$$

$$[\text{oleic}] = \frac{3C}{4(E+F)} - [\text{linoleic}] - [\text{linolenic}] \quad (3)$$

$$[\text{SFA}] = \frac{F}{E+F} - [\text{linoleic}] - [\text{oleic}] - [\text{linolenic}] \quad (4)$$

Signal intensity *B* of the CH_2 protons of all acyl chains is related to signal intensities *E* and *F* by the following relationship:

$$\frac{E+F}{3} = \frac{B}{2} \quad (5)$$

Substituting the relationship 5 into eqs 2 and 3, we obtain two new relationships, which can also be used for the calculation of the molar concentrations of linoleic and oleic acids.

$$[\text{linoleic}] = \frac{3A - 4E}{3B} \quad (6)$$

$$[\text{oleic}] = \frac{C}{2B} - [\text{linoleic}] - [\text{linolenic}] \quad (7)$$

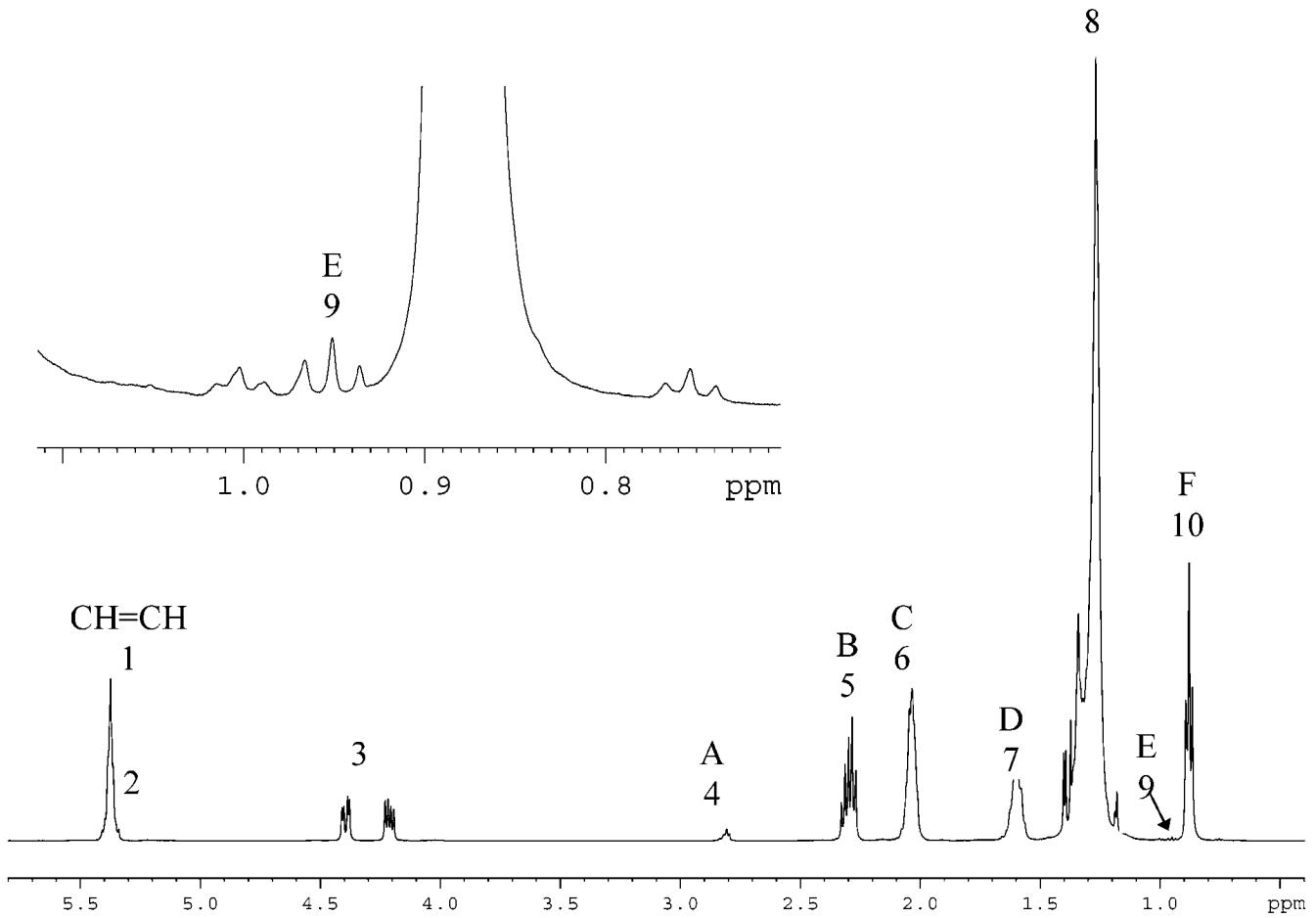

Figure 3. 500 MHz ^{1}H NMR spectrum of virgin olive oil. The inset shows expansion of the spectrum where the methyl protons (signal *E*) of the linolenyl chain appear. The assignments of all signals are summarized in **Table 1**.

Table 1. Chemical Shifts and Assignments of the Signals in the ^{1}H NMR Spectrum of Virgin Olive Oil in Pyridine-d_5/Chloroform-*d* Solvent (**Figure 3**) According to Sacchi et al. (*11*, *27*)

signal	δ	protons	attribution
1	5.40	C***H***═CH	all unsaturated fatty acids
2	5.37	C***H***—OCOR	triglycerides
3	4.22, 4.42	C***H*$_2$**—OCOR	triglycerides
4 (*A*)	2.84	CH═CH–C***H*$_2$**—CH═CH	linolenyl and linoleyl chains
5 (*B*)	2.32	C***H*$_2$**–COOH	all acyl chains
6 (*C*)	2.06	C***H*$_2$**—CH═CH	all unsaturated acyl chains
7 (*D*)	1.63	C***H*$_2$**—CH$_2$COOH	all acyl chains
8	1.30	(C***H*$_2$**)$_n$	all acyl chains
9 (*E*)	0.95	CH═CH—CH$_2$—C***H*$_3$**	linolenyl chain
10 (*F*)	0.88	CH$_2$CH$_2$CH$_2$—C***H*$_3$**	all acyl chains except linolenyl

The validity of each set of relationships, namely, 1−4 and 1, 4, 6, and 7, to calculate the fatty acid concentrations can be checked independently by their ability to reproduce the intensity of signal 1, which belongs to the olefinic protons of all unsaturated acids in the ^{1}H NMR spectrum. The intensity of this signal, which depends on the degree of oil unsaturation, is related to the concentrations of the fatty acids and the signal intensity of the methylene protons of triacylglycerols (TG) in the spectrum (δ 4.42) by the following relationship:

$$[-\mathrm{CH}{=}\mathrm{CH}-] = \frac{2[\text{oleic}] + 4[\text{linoleic}] + 6[\text{linolenic}]}{0.5 \times \left(\frac{E+F}{3} + \frac{B}{2}\right)} + \frac{[\mathrm{TG}]}{2} \quad (8)$$

It should be noted that relationship 2 of the present study is different from relationship 5 of previous reports (*11*, *27*). It appears that the latter relationship is erroneous for two reasons; first, the normalization factor ($E + F$) does not apply to the A protons, and, second, a factor of two-thirds is appropriate for the intensity of the two diallylic protons (signal *A* at δ 2.84), which are compared with the intensity of the methyl protons of linolenic acid (signal *E* at δ 0.95).

The iodine value is given by the relationship

$$\mathrm{IV} = \left(\frac{\frac{[-\mathrm{CH}{=}\mathrm{CH}-]}{2} - \frac{[\mathrm{TG}]}{4}}{\frac{E+F}{3}} \right) \times 86 \quad (9)$$

where the concentrations [−CH═CH−] and [TG] were calculated from the signal integrals at δ 5.40 and 4.42, respectively, and 86 represents the IV of oleic acid.

Variables of the Oils Used for Statistical Analyses. The mean values and standard deviations of the variables used in this study for the 192 samples of 13 types of vegetable oils as a whole are summarized in **Table 2**. The data for each individual sample are available as Supporting Information. We decided to use compositional parameters and the IV as variables for the statistical analysis instead of NMR signal integrals for the following reasons. First, the former profile gives a first indication of the compositional differences among the various oils, which may be used subsequently for their differentiation. Second, this variables profile resulted in a better discrimination of the oils

Table 2. Mean Values and Standard Deviations of the Variables for 13 Types of Vegetable Oils

oil	1,2-DGs	1,3-DGs	total DGs	*D*	total sterols	acidity	oleic acid	linoleic acid	linolenic acid	SFA	IV[a]
olive	1.71 ± 0.29	0.17 ± 0.09	1.88 ± 0.31	0.91 ± 0.04	0.12 ± 0.03	0.21 ± 0.12	77.13 ± 2.48	7.89 ± 1.36	0.47 ± 0.10	14.51 ± 1.65	80.58 ± 1.93
soybean	0. 33 ± 0.07	0.74 ± 0.20	1.08 ± 0.25	0.33 ± 0.04	0.21 ± 0.04	0.03 ± 0.01	24.63 ± 1.68	50.11 ± 0.94	7.30 ± 0.73	17.98 ± 0.68	127.35 ± 2.27
sunflower	0.50 ± 0.08	1.01 ± 0.17	1.51 ± 0.24	0.33 ± 0.02	0.16 ± 0.04	0.01 ± 0.02	26.06 ± 1.79	60.78 ± 2.65	0.61 ± 0.26	12.37 ± 0.36	129.12 ± 1.82
corn	0.98 ± 0.14	2.16 ± 0.19	3.14 ± 0.31	0.31 ± 0.02	0.33 ± 0.03	0.05 ± 0.03	31.14 ± 1.18	52.51 ± 1.91	1.83 ± 0.72	14.59 ± 0.90	120.49 ± 2.32
hazelnut	0.87 ± 0.15	1.85 ± 0.36	2.72 ± 0.50	0.32 ± 0.02	0.07 ± 0.03	0.12 ± 0.07	77.04 ± 6.85	12.95 ± 6.13	0.29 ± 0.15	9.51 ± 0.82	89.91 ± 4.40
sesame	0.74 ± 0.14	1.55 ± 0.23	2.29 ± 0.36	0.32 ± 0.01	0.17 ± 0.05	0.06 ± 0.02	42.68 ± 0.85	38.74 ± 0.91	0.88 ± 0.21	16.98 ± 2.98	105.81 ± 1.05
groundnut	0.73 ± 0.07	1.62 ± 0.08	2.36 ± 0.15	0.31 ± 0.01	0.15 ± 0.02	0.04 ± 0.05	57.75 ± 0.24	20.65 ± 0.45	1.02 ± 0.22	20.59 ± 0.25	87.48 ± 0.49
walnut	1.20 ± 0.16	2.27 ± 0.26	3.47 ± 0.47	0.35 ± 0.01	0.09 ± 0.02	0.00	16.81 ± 1.33	57.66 ± 1.59	13.69 ± 0.58	11.84 ± 0.52	150.09 ± 1.94
rapeseed	0.30 ± 0.05	0.61 ± 0.11	0.91 ± 0.14	0.33 ± 0.03	0.21 ± 0.04	0.02 ± 0.02	62.82 ± 0.50	16.87 ± 0.29	11.93 ± 0.09	8.39 ± 0.29	112.99 ± 0.88
almond	0.66 ± 0.09	1.16 ± 0.11	1.82 ± 0.18	0.33 ± 0.03	0.17 ± 0.04	0.00	57.51 ± 1.43	25.24 ± 0.61	0.63 ± 0.08	8.78 ± 0.89	94.42 ± 1.32
safflower	0.52 ± 0.04	1.19 ± 0.07	1.71 ± 0.09	0.30 ± 0.02	0.10 ± 0.02	0.04 ± 0.03	13.92 ± 0.14	73.41 ± 0.18	0.71 ± 0.04	11.96 ± 0.18	139.80 ± 0.31
palm	1.32 ± 0.18	5.61 ± 0.69	6.93 ± 0.87	0.19 ± 0.01	0.01 ± 0.01	0.14 ± 0.02	38.38 ± 1.12	8.26 ± 0.25	0.51 ± 0.09	52.84 ± 3.12	47.63 ± 0.48
coconut	1.48 ± 0.51	3.07 ± 1.01	4.55 ± 1.52	0.33 ± 0.01	0.06 ± 0.03	0.04 ± 0.01	7.31 ± 2.30	0.32 ± 0.16	0.37 ± 0.06	92.00 ± 2.23	6.12 ± 0.47

[a] Iodine value.

upon statistical analysis. A third reason is the possibility that the present statistical analysis can be used with compositional indices extracted by employing analytical methods other than NMR.

Inspection of the data in **Table 2** reveals that the parameters that differentiate olive oils from the remaining seed oils are the diglyceride content and acidity. In the former oils the 1,3-DGs are much lower and the *D* ratio is much higher than the corresponding values of the latter oils. This is due to the refinement of the seed oils as mentioned previously. The fatty acid composition is about the same in both olive and hazelnut oils except perhaps for the amount of saturated fatty acid (mainly palmitic acid), which is lower for hazelnut oil. Olive and hazelnut oils are characterized by high amounts of monounsaturated fatty acid (oleic acid), whereas lower concentrations are observed for walnut, safflower, and coconut oils. The amount of polyunsaturated fatty acids (linoleic and linolenic) is high for safflower, soybean, corn, sunflower, and walnut seed oils and low for olive, hazelnut, palm, and coconut oils. The degree of unsaturation is reflected in the iodine values (**Table 2**). High concentrations of saturated fatty acids are observed for palm, coconut, sesame, and walnut oils. Coconut oil contains the highest amount of saturated fatty acids and is characterized by the lowest IV. Finally, high levels of linolenic acid are observed for walnut, rapeseed, and soybean seed oils.

Classification of Vegetable Oils. From the data in **Table 2** (mean values and standard deviations), it is possible to formulate an immediate judgment on the discrimination ability of a single variable. For instance, the diglyceride content differentiates olive oils from the rest of the vegetable oils, whereas the concentration of saturated fatty acids discriminates coconut oil. Nevertheless, our aim is to discriminate among the 13 groups of oils and to find the necessary variables for this purpose. In other words, we seek the number of variables that should be included in a classification/prediction statistical model without compromising its discrimination ability. There are two ways to do this; principal component analysis (PCA) and one-way ANOVA. The Fisher *F* ratio offered better results selecting variables for which $F > 1$. These are (in parentheses are the *F* values with degrees of freedom of 12 and 179) 1,3-DGs (230.93), 1,2-DGs (81.46), total DGs (101.61), *D* (1231.95), sterols (73.52), acidity (19.99), oleic acid (1401.80), linoleic acid (1732.34), linolenic acid (1665.74), SFA (3303.18), and IV (4068.26) with the probability of the null hypothesis (no statistically significant differences between the means of the groups) to be true at $P < 0.000001$ for all variables. Moreover, the *F* ratios of these variables are higher than the critical values obtained from standard statistical tables at $P = 0.004$, $F_{\text{critical}}(12, \infty) = 2.48$ or upon comparison with $F_{\text{critical}}(12, 4) = 5.91$, four being the number of oil samples minus one for the group of edible oils with the least number of samples (rapeseed with five samples). The strict criterion for the chosen probability (usually $P = 0.05$) was necessitated by the fact that the groups of edible oils contain different numbers of samples. The relatively low Fisher *F* ratio observed for acidity is indicative of its limited discriminatory power, and thus it is excluded from further statistical analysis. The number of variables is further reduced to 9, because total DGs are the sum of 1,2-DGs and 1,3-DGs and, therefore, are completely redundant. Finally, after an investigation of their contribution to the classification/prediction model, the most useful variables in differentiating the different types of oils were 1,2-DGs, 1,3-DGs, *D* ratio, sterols, IV, and oleic, linoleic, and linolenic acid. Nevertheless, very good classification of the edible oils is also obtained upon replacing linolenic acid with saturated acids (SFA) in the assembly of the eight variables.

Table 3. Discriminant Analysis: Raw Coefficient,[a] Wilks' λ Values, and Tolerance Levels for Each Variable Used for the Classification of 13 Types of Vegetable Oils

variable	root 1	root 2	Wilks' λ ($\times 10^{10}$)	tolerance level
1,2-DGs	–1.43	–0.83	2.21	0.46
1,3-DGs	1.30	0.81	5.77	0.48
D	7.70	–13.71	8.15	0.76
total sterols	–2.74	–2.28	5.78	0.90
oleic acid	–0.40	–0.42	2.86	0.18
linoleic acid	–0.47	–0.24	1.86	0.11
linolenic acid	–1.20	1.07	3.63	0.74
IV	–0.33	–0.01	1.77	0.39
constants	65.05	33.38		
eigenvalues	450.34	185.54		

[a] Only the raw coefficients for the two canonical functions (roots 1 and 2) that discriminate the vegetable oils better are summarized.

The use of the two of the eight discriminant functions with the highest discriminating power succeeded in classifying the 192 oil samples into different groups. This is reflected by the largest raw regression coefficients of the canonical function (roots), which in turn reflect the contribution of the respective variable to the discrimination between groups. The coefficients assigned to each variable in the canonical functions, the tolerance value for each variable, and the Wilks' λ criterion for the discriminating power of the selected variables are shown in **Table 3**. The values of Wilks' λ are very small, indicating that the eight selected variables are characterized by an almost perfect discriminatory power. We can visualize the discriminating power of the two canonical functions between groups (types) of oils by plotting the individual scores of the samples with the

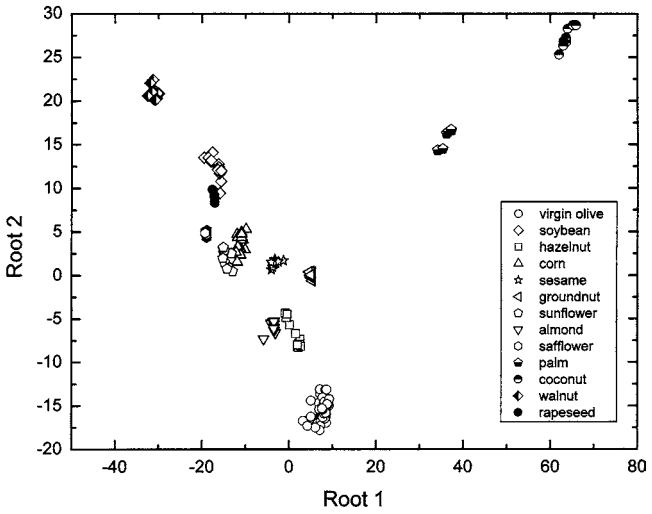

Figure 4. Plot of discriminant functions roots 1 and 2 for 192 samples of 13 types of vegetable oils.

two functions. Such a plot for 13 types of oil samples is seen in **Figure 4**. It is evident that points are nicely grouped within single types of edible oils. A much better separation of the edible oils can be obtained by excluding from the statistical analysis the palm and coconut oils, which are remotely grouped from the remaining groups of oils (not shown).

The reliability of the system used to classify the edible oils has been assessed by its ability to correctly classify unknown samples. In this respect, the data set of all oils was split into the training (128) and test (64) sets, with a 2:1 ratio. The samples of the training and test sets were selected three times at random from the 192 oil samples. Each time, different training and test sets were used. Application of the DA functions obtained for the training set of oil samples in all three runs produced 100% correct assignments for the test set.

Adulteration of Virgin Olive Oil. The aforementioned statistical approach was also applied to the analysis of oil mixtures of virgin olive oil in an attempt to establish the lowest possible detection level of adulteration of virgin olive oil with seed oils. Different mixtures of olive oil and hazelnut oil, olive oil and corn oil, olive oil and sunflower oil, and olive oil and soybean oil were prepared and analyzed as unknowns. These seed oils, especially hazelnut oil, were selected because they are frequent adulterants of olive oil. **Figure 5** shows the plot of the individual scores of the two canonical functions upon applying DA on 38 samples of fresh virgin olive oils ($D \geq 0.90$), 16 samples of hazelnut oils, 12 samples of sunflower oil, 16 samples of corn oil, and 14 samples of soybean oil. **Table 4** contains the coefficients assigned to each variable in the canonical functions, the tolerance value, and the Wilks' λ criteria for each variable. The ellipses in **Figure 5** delimit 95% confidence. To validate this model, we followed the same procedure with the training sets and test sets as mentioned above. All test samples were correctly classified in the appropriate groups.

Table 4. Discriminant Analysis: Raw Coefficient,[a] Wilks' λ Values, and Tolerance Levels for Each Variable Used for the Classification of 5 Types of Vegetable Oils

variable	root 1	root 2	Wilks' λ ($\times 10^7$)	tolerance level
1,2-DGs	.002	0.39	4.04	0.69
1,3-DGs	2.18	−2.56	11.16	0.49
D	43.16	16.70	25.60	0.73
total sterols	−6.39	2.25	8.91	0.79
oleic acid	−0.05	−0.16	5.28	0.08
linoleic acid	−0.13	−0.04	5.67	0.05
linolenic acid	−0.18	0.76	5.29	0.64
IV	−0.26	0.13	4.91	0.26
constants	7.64	−12.16		
eigenvalues	353.17	39.26		

[a] Only the raw coefficients of the two canonical functions (roots 1 and 2) that discriminate the vegetable oils better are summarized.

The same statistical approach was applied to the analysis of the oil mixtures (adulterated virgin olive oils), which have been analyzed by ^{1}H NMR and ^{31}P NMR spectroscopy and considered as unknown samples. The results were obtained using the model obtained by the training set and are depicted in **Figure 5**. The adulterated virgin olive oils (open symbols) lie between the group of virgin olive oils and the respective group of the seed oils, depending on the amount of seed oils in the mixtures. Even mixtures containing the lowest amount of seed oils in olive

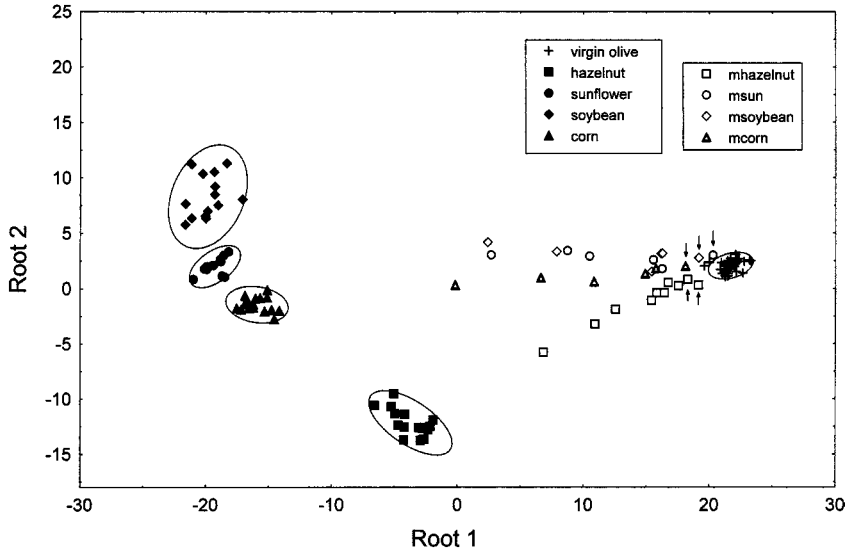

Figure 5. Plot of discriminant functions roots 1 and 2 for five types of edible oils. Virgin olive oil, hazelnut oil, corn oil, and sunflower oil are shown by crosses and solid symbols. Four sets of mixtures of 5, 10, 15, 20, 35, and 50% w/w of virgin olive oils with hazelnut (mhazelnut), sunflower (msun), soybean (msoybean), and corn (mcorn) oils and one set of mixtures of 5, 10, 15, and 20% of virgin olive oils with hazelnut oils (mhazelnut) are denoted by open symbols. Arrows indicate mixtures of 5% w/w of seed oils in virgin olive oils.

Table 5. Squared Mahalanobis Distances of Adulterated Olive Oil Samples from Olive Oil Centroid

adulterant	5% (w/w)	10% (w/w)	15% (w/w)	20% (w/w)	35% (w/w)	50% (w/w)
hazelnut[a]	16.97	48.41	39.24	43.98	151.77	299.86
hazelnut[b]	15.30	24.20	53.42	106.77		
sunflower	8.83	44.25	40.19	153.84	177.21	379.73
corn	14.31	44.69	51.86	123.52	233.16	487.16
soybean	13.53	35.03	31.89	43.55	194.64	397.80

[a] First set of mixtures (5–50%) of olive oils and hazelnut oils. [b] Second set of mixtures (5–20%) of olive oils and hazelnut oils.

oil (5%) fall outside the ellipses, except perhaps the 5% mixture of olive oil and sunflower oil (**Figure 5**).

Another means to examine the validity of the present DA method to detect adulteration is the Mahalanobis distance, which measures the distance between each point and the group centroid. The calculated Mahalanobis distances for mixtures (**Table 5**) are to be compared with the Mahalanobis distances of virgin olive oils. Apart from 3 olive oil samples of 38 samples with Mahalanobis distances of 22.49, 19.14, and 17.89, which fall outside the 95% olive oil ellipse, all other olive oil samples are characterized by distances lower than those calculated for the 5% adulterated olive oils, except perhaps the mixture of olive oil with sunflower. These data clearly show that this methodology and the selected variables are able to detect adulteration as low as 5%.

The selection of 38 samples of virgin olive oil of the 56 originally used for the classification of edible oils was dictated by the freshness of olive oils as reflected by their *D* values. Previous studies (*20*−*22*) have demonstrated that fresh virgin olive oils extracted from olives of normal ripeness are characterized by high values of the *D* ratio, which subsequently decrease upon storage due to the isomerization of 1,2-diglycerides to 1,3-diglycerides. Although some other factors (olive variety, climatic conditions, fruit ripeness, and extraction methods) may play a role in determining the diglyceride content of olive oil, virgin olive oils freshly extracted from olives of normal ripeness should have high *D* ratios close to 1. In this respect, the virgin olive oil samples with $D \geq 0.90$ were chosen for adulteration studies. Inclusion of olive oil samples with *D* ratios <0.90 in the present analysis decreases the limit of the adulterant detection. For instance, applying the same statistical analysis as above for the adulterated olive oil mixtures and including 43 virgin olive oils with $D \geq 0.88$, adulterated mixtures of 5−10% fall outside the 95% ellipse of virgin olive oils, whereas for 48 virgin olive oils with $D \geq 0.85$, the detection limit is 10−15%. This finding emphasizes the importance of the diglyceride indices in attempts to detect olive oil adulteration with edible oils.

This study has demonstrated that ^{31}P and ^{1}H NMR spectroscopy are powerful tools to classify oils of diverse botanical origin and to detect adulteration of fresh virgin olive oils with other seed oils at very low concentrations. It appears that the diglyceride content is a meaningful variable for successful classification of edible oils and detection of virgin olive oil adulteration. Compared with conventional methods, NMR spectroscopy offers a number of advantages, among which speed appears to be the most important. NMR spectroscopy allows, within a certain range, the determination of a large number of oil constituents in one or two experiments. This feature makes it very useful for fast screening of large numbers of samples and for setting up a comprehensive data bank of authentic oils. Additional information from other minor oil components, such as phenolic acids (polyphenols), can also be included in this procedure. These constituents can be detected from the polar part of the oils by employing ^{31}P NMR spectroscopy. We are currently investigating this possibility.

Supporting Information Available: Detailed table of the compositional data for all edible oil samples and their mixtures used in this study. This material is available free of charge via the Internet at http://pubs.acs.org.

LITERATURE CITED

(1) Boskou, D. Olive oil adulteration. In *Olive Oil Chemistry and Technology*; Boskou, D., Ed.; AOCS Press: Champaign, IL, 1996; Chapter 8, pp 121−133.

(2) Blanch, G.; del Mar Caja, M.; Ruiz del Castillo, M.-L.; Herraiz, M. Comparison of different methods for the evaluation of the authenticity of olive oil and hazelnut oil. *J. Agric. Food Chem.* **1998**, *46*, 3153−3157.

(3) Rezanka, T.; Rezankova, H. Characterization of fatty acids and triacylglycerols in vegetable oils by gas chromatography and statistical analysis. *Anal. Chim. Acta* **1999**, *398*, 253−261.

(4) Salivaras, M.; McCurdy, A. R. Adulteration of olive oil detection through reversed phase high performance liquid chromatography. In *Food Flavors, Ingredients and Composition*; Charalambous, G., Ed.; Elsevier Science Publishers: Amsterdam, The Netherlands, 1993; pp 279−300.

(5) Tsimidou, M.; Macrae, R. Authentication of virgin olive oil using principal component analysis of triglycerides and fatty acid profiles; Part 2—Detection of adulteration with other vegetable oils. *Food Chem.* **1987**, *25*, 251−259.

(6) Goodacre, R.; Kell, D. B. Rapid assessment of adulteration of virgin olive oils by other seed oils using pyrolysis-mass spectrometry and artificial neural networks. *J. Agric. Food Chem.* **1993**, *63*, 297−305.

(7) Spitzer, V. Structure analysis of fatty acids by gas chromatography—Low resolution electron impact mass spectrometry of their 4,4-dimethyloxazoline derivatives—A review. *Prog. Lipid Res.* **1997**, *35*, 387−408.

(8) Ozen, B. F.; Mauer, L. J. Detection of hazelnut oil adulteration using FT-IR spectroscopy. *J. Agric. Food Chem.* **2002**, *50*, 3898−3901.

(9) Downey, G.; McIntyre, P.; Davies, A. N. Detecting sunflower oil adulteration in extra virgin olive oils from the eastern Mediterranean by visible and near-infrared spectroscopy. *J. Agric. Food Chem.* **2002**, *50*, 5520−5525.

(10) Baeten, V.; Hourant, P.; Morales, M.; Aparicio, R. Oil and fat classifications by FT-Raman spectroscopy. *J. Agric. Food Chem.* **1998**, *46*, 2636−2646.

(11) Sacchi, R.; Addeo, F.; Paolillo, L. ^{1}H and ^{13}C NMR of virgin oil. An overview. *Magn. Reson. Chem.* **1997**, *35*, 5133−5145.

(12) Vlahov, G. Application of NMR to the study of olive oils. *Prog. Nucl. Magn. Reson. Spectrosc.* **1999**, *35*, 341−357.

(13) Fauhl, C.; Reniero, F.; Guillou. ^{13}C NMR as a tool for the analysis of mixtures of virgin olive oils of different botanical origin. *Magn. Reson. Chem.* **2000**, *38*, 436−443.

(14) Mannina, L.; Patumi, M.; Fiordiponti, P.; Emanuele, M. C.; Segre, A. L. Olive and hazelnut oils: A study by high-field ^{1}H NMR and gas chromatography. *Ital. J. Food Sci.* **1999**, *11*, 139−149.

(15) Mavromoustakos, T.; Zervou, M.; Bonas, G.; Kolocouris, A.; Petrakis, P. A novel analytical method to detect adulteration of virgin olive oil by other oils. *J. Am. Oil Chem. Soc.* **2000**, *77*, 405−411.

(16) Vlahov, G. Quantitative ^{13}C NMR method using the DEPT pulse sequence for detection of olive oil adulteration with soybean oil. *Magn. Reson. Chem.* **1997**, *35*, S8−S12.

(17) Zamora, R.; Gomez, G.; Hidalgo, F. J. Classification of vegetable oils by high-resolution ^{13}C NMR spectroscopy using chromatografically obtained oil fractions. *J. Am. Oil Chem. Soc.* **2002**, *79*, 267−272.

(18) Zamora, R.; Gomez, G.; Dobarganes, M. C.; Hidalgo, F. J. Oil fractionation as a preliminary step in the characterization of vegetable oils by high-resolution ^{13}C NMR spectroscopy. *J. Am. Oil Chem. Soc.* **2002**, *79*, 261−266.

(19) Spyros, A.; Dais, P. Application of ^{31}P NMR spectroscopy in food analysis. I. Quantitative determination of the mono- and diglyceride composition of olive oils. *J. Agric. Food Chem*. **2000**, *48*, 802−805.

(20) Fronimaki, P.; Spyros, A.; Christophoridou, S.; Dais, P. Determination of the diglyceride content in Greek virgin olive oils and some commercial olive oils by employing P31 NMR spectroscopy. *J. Agric. Food Chem*. **2002**, *50*, 2207−2213.

(21) Perez-Camino, M. C.; Modera, W.; Cert, A. Effects of olive oil fruit quality and oil storage practices on the diacylglycerols content of virgin olive oils. *J. Agric. Food Chem.* **2001**, *49*, 699−704.

(22) Mariani, C.; Fedeli, E. Determinazione delle forme gliceridiche presenti negli oli alimentary. Nota I: caso dell′ olio di oliva. *Riv. Ital. Sostanze Grasse* **1985**, *62*, 3−7.

(23) European Communities. Regulation 2568/91. *Off. J. Eur. Communities* **1991**, L 248.

(24) International Olive Oil Council. COI/T.15/NC n. 2/rev. 4 (June 6), Madrid, 1996.

(25) Zwierzak A. Cyclic organophosphorus compounds. I. Synthesis and infrared spectral studies of cyclic hydrogen phosphites and thiophosphites. *Can J. Chem*. **1967**, *45*, 2501−2512.

(26) Christoforidou, S.; Spyros, A.; Dais, P. ^{31}P Nuclear Magnetic Resonance Spectroscopy of Polyphenol-Containing Olive Oil Model Compounds, *Phosphorous, Sulfur Silicon Relat. Elem.* **2001**, *170*, 139−157.

(27) Sacchi, R.; Patumi, M.; Fontanazza, G; Barone, P.; Fiordiponti, P.; Mannina, L.; Rossi, E.; Segre, A. A high field ^{1}H nuclear magnetic resonance study of the minor components in virgin olive oils. *J. Am. Oil Chem. Soc.* **1996**, *73*, 747−758.

Received for review February 12, 2003. Revised manuscript received June 23, 2003. Accepted June 23, 2003. We thank the Greek Ministry of Education for financial support through the B' EPEAEK Graduate Program.

JF030100Z