MAGNETIC RESONANCE IN CHEMISTRY
Magn. Reson. Chem. 2000; **38**: 436–443

^{1}H NMR as a tool for the analysis of mixtures of virgin olive oil with oils of different botanical origin

Carsten Fauhl, Fabiano Reniero and Claude Guillou*

DG Joint Research Centre of the European Commission, Institute for Health and Consumer Protection, Food Products and Consumer Goods Unit, I-21020 Ispra, Italy

Received 31 March 1999; revised 17 January 2000; accepted 7 February 2000

ABSTRACT: ^{1}H NMR spectroscopy provides a possible alternative to conventional chromatographic methods for determining the composition of oils. In this study, various oils from olive, hazelnut and sunflower were analysed by ^{1}H NMR spectroscopy. Experimental conditions were chosen in order to have a short experimental time. It was demonstrated that multivariate statistical methods, in particular discriminant analysis, applied to selected predominant peaks in the ^{1}H NMR spectra of oils resulted in a good separation between these three oils of different botanical origin and permitted the detection of their mixtures. Copyright © 2000 John Wiley & Sons, Ltd.

KEYWORDS: NMR; ^{1}H NMR; authentication; edible oils; olive oil; multivariate statistics

INTRODUCTION

Olive oil is a valuable product in terms of its commercial and nutritional value and therefore it is often subject to adulteration with other edible oils, such as sunflower and hazelnut oil, which have a composition of fatty acids close to that of olive oil.

Within the European Community, the control of the authenticity of olive oil is usually carried out by using classical analysis, including chromatographic methods, of several fractions of this product. These analytical methods are described in EC Regulation 656/95. Recently, other analytical approaches have been applied for the characterization of olive oil; for instance, isotopic techniques have been used for the determination of its botanical and geographical origin[1] and for the detection of adulteration with pomace oil.[2] ^{1}H NMR has been used to identify the geographical origin[3] of different Italian olive oils and to classify the variety of olive tree.[4] ^{13}C NMR has also been used for the classification of olive oil.[5]

It has already been shown that the combination of NMR data, or other spectroscopic data such as from Fourier transform infrared (FT-IR) spectroscopy, and multivariate statistics delivers interesting results for authentication purposes.[6–8] The use of discriminant analysis of peak picking results from ^{1}H NMR spectra has, for example, allowed the successful distinction between pure orange juice and adulterated juices.[9,10]

Recently, various methods have been proposed for the detection of the fraudulent addition of hazelnut oil to olive oil. The analysis of filbertone, a flavour compound in hazelnuts, was used to detect its addition to olive oil.[11,12] Spangenberg *et al*. measured the ^{13}C/^{12}C ratios of fatty acids using gas chromatography coupled to isotopic ratio mass spectrometry (GC–IRMS) and found differences between the various botanical origins that were investigated.[13] This approach seems to be promising but must be confirmed at least for hazelnut oil as only one sample of this botanical origin was analysed in their study. Recently Mannina *et al*. described a combined approach using NMR and GC analysis for the determination of hazelnut oil in olive oil.[14]

In the present work, we investigated the possibility of analysing fraudulent mixtures of olive oil with sunflower oil or hazelnut oil by one-dimensional ^{1}H NMR. In fact, ^{1}H NMR delivers qualitative and quantitative information about the composition of many major and minor compounds present in oil.[4] The large amount of information thus provided renders desirable the use of multivariate statistical analysis to facilitate the interpretation of the data and to identify the botanical origin of the oil or to detect adulteration.

Previous studies regarding the detection of adulteration of olive oil with hazelnut oil were based on a restricted number of samples often of commercial origin. We therefore analysed a large number of olive oil and hazelnut oil samples, including samples extracted in the laboratory, in order to have a sound base of data for the application of multivariate statistical analysis.

The final goal of this study was to provide a fast analytical method for the rapid screening of a large number of samples of olive oil, usable by control laboratories (anti-fraud, customs, etc.).

EXPERIMENTAL

Samples

The samples analysed were as follows (see Table 2):

(i) 42 genuine virgin olive oils of different origin (three Greek, seven Spanish, one Tunisian, one

* *Correspondence to*: C. Guillou, DG Joint Research Centre of the European Commission, Institute for Health and Consumer Protection, Food Products and Consumer Goods Unit, I-21020 Ispra, Italy; e-mail: claude.guillou@jrc.it
Contract/grant sponsor: DGXXI.

Hungarian and 30 Italian). The Italian Institute for olive and olive oil studies (Istituto per l'Elaiotecnica) of Pescara provided these samples, which were already being analysed in the frame of other research work regarding the isotopic characterization of this commodity.[1]

(ii) 31 hazelnut oils, of which 14 samples were prepared in laboratory. The remaining 17 were commercial hazelnut oils, purchased in different European countries and in Turkey.

(iii) Seven commercial sunflower oils, purchased in Italy.

Hazelnut oil extraction. The 14 hazelnut oils produced at laboratory level were made from hazelnuts from different countries (two from Turkey, one from USA, five from Italy and six of undeclared origin) purchased on the retail market. The shell was removed and the oil was extracted by mechanical pressing, with a yield of ~40%. For one hazelnut sample (Turkey), extraction was made also with *n*-hexane, with a yield of ~65% (sample 18 in Table 2) and could be compared with its homologue (sample 19) obtained by mechanical pressing. The parameters obtained for these samples are similar. The mode of extraction (mechanical/solvent) therefore seems not to have a very strong effect on the NMR data.

NMR analysis

The first step was the optimization of the oil concentration in the NMR tube, in order to have a short time of analysis with good sensitivity (high signal-to-noise ratio) without losing resolution. In general, oil sample preparation for NMR consists simply of the dilution of the sample with a deuterated solvent. In this case a mixture of deuterated chloroform ($CDCl_3$) and deuterated dimethyl sulphoxide (DMSO-d_6) was chosen. The addition of DMSO is necessary for the solubility of polar minor components in the oil.[4] The highest oil concentration usable with our system was 25% (v/v), which permitted a spectral resolution comparable to that described in the literature[4] where lower concentrations (<1.5%) were used. It was observed that concentrations only slightly above 25% (v/v) led to poorly resolved ^{1}H NMR spectra. The deuterium signal of $CDCl_3$ was used for locking and shimming of the magnetic field. All samples were prepared by mixing 150 µl of oil, 150 µl of DMSO-d_6 and 300 µl of $CDCl_3$.

The longitudinal relaxation times (T_1) of the main signals were determined by the inversion–recovery method.[15] The longest T_1, 1.73 s, was measured for the methyl group signal of the fatty acids, and the shortest T_1, 0.43 s, for the glycerol protons. The following conditions for acquisition of ^{1}H NMR spectra were therefore selected: pulse angle 30°, corresponding to 2.9 µs; 83 332 data points acquired during an acquisition time of 5.5 s; relaxation delay 0.1 s; sweep width 15.15 ppm (7576 Hz); and temperature 300 K. After 16 dummy scans, 400 scans were recorded for each tube. The total acquisition time was 37 min.

The NMR spectrometer used was a Bruker AMX 500 equipped with a 5 mm broadband inverse probe operating at the basic frequency of 500.14 MHz (O1). The spectrometer was equipped with an autosampler and the acquisition of the free induction decay (FID) of the oils was performed with automation.

The processing of the FID was done by Fourier transformation, applying a line broadening factor of 0.3 Hz and zero filling, and the resulting spectra were manually phased. The baseline was also corrected manually by a multi-point correction, setting 12 points between −0.5 and 6.5 ppm in a range free of any signal. Manual processing of the spectra gave better results with respect to the automated processing procedures based on standard NMR spectrometer software.

In order to circumvent any artificial signals, such as spinning side bands of the first or higher order, the spectra were acquired without spinning the NMR tube. The use of an inverse probe gives a high sensitivity for proton spectra and allows a good resolution without spinning. The resolution was estimated using the glycerol signals, whereby a line splitting of 30% of the quartet at 4.3 ppm indicated a reasonable shimming within the automation procedure. Any spectrum with a lower resolution was reacquired. This was the case for less than 5% of the samples analysed.

Statistical analysis

For all the statistical analysis the software package Statistica 5.1[16] was used. A data matrix of 80 cases × 5 variables was used for the statistical analysis.

RESULTS AND DISCUSSION

Edible oils are composed of the esters of glycerol and fatty acids. The main fatty acids in olive oil are oleic and linoleic acid. The ^{1}H NMR spectra of oils are characterized by overlapping signals originating from various fatty acids in different combinations of triglycerides. Owing to the chemical similarity of the different triglyceride esters, the signals resonate close together and build clusters. Figure 1 shows a typical olive oil ^{1}H NMR spectrum composed of 10 main signal clusters. The signal assignment has already been performed by Sacchi *et al.*[4] and is given in Table 1.

Only the chemical groups assigned in Table 1 give separated signals. For this reason, it is not possible to determine all single fatty acid components, but it is possible, for example, to obtain the global unsaturation proportion calculated from signal 1 or 6, which corresponds to the iodine number value.[17] Furthermore, the proportion of *n*-3 linolenic acid can be calculated by the consideration of signal 9 in relation to the ^{13}C satellites of signal 10 (see Fig. 3). The amount of polyunsaturated fatty acids can be determined from signal 4. The content of monounsaturated fatty acids and saturated fatty acids can also be calculated

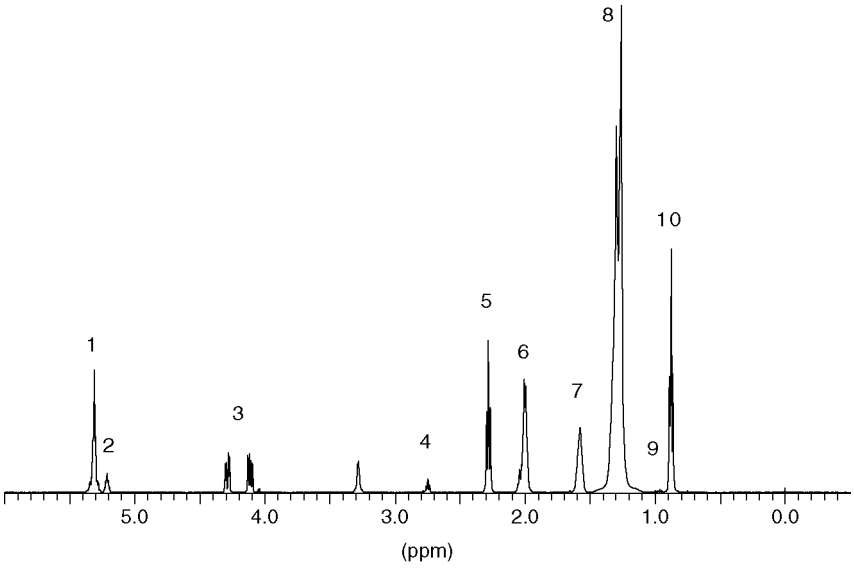

Figure 1. ^{1}H NMR spectrum of olive oil.

Table 1. Assignment of the main signals according to Sacchi *et al*.[4]

Peak	δ(ppm)	Proton	Proton	Compound
1	5.29	C*H*═C*H*	Olefinic	All unsaturated fatty acids
2	5.15	—C*H*—O—CO—R	Glycerol	Triacylglycerols
3	4.19	—CH_2—O—CO—R	Glycerol	Triacylglycerols
4	2.76	—CH═CH—CH_2—CH═CH$_2$—	Diacyl	Linoleic and linolenic acid
5	2.20	—CH_2—COOR	α-Carboxyl	All acyl chains
6	2.02	—CH_2—CH═CH—	α-Olefinic	All unsaturated fatty acids
7	1.6	—CH_2—CH$_2$—COOR	β-Carboxyl	All acyl chains
8	1.2	—(CH_2)$_n$	Methylene groups	All acyl chains
9	0.95	—CH═CH—CH$_2$—CH_3	Methyl groups	Linolenic acid
10	0.85	—CH$_2$—CH$_2$—CH$_2$—CH_3	Methyl groups	All acyl chains except linolenic

using various signals (5, 6, 9, 10) from a one-dimensional ^{1}H NMR spectrum.[18]

In order to obtain the maximum attainable grouping within one type of oil, we used discriminant analysis (DA) for further data evaluation,[19,20] This procedure modifies data in such a way that one obtains the maximum variance between the predefined groups which, in this case, are the three different types of oils. DA works properly only with non-redundant data, which means that correlations between variables have to be avoided, otherwise wrong or unstable solutions may be calculated. This has to be taken into account when using NMR data for DA because many of the signals are internally correlated owing to their chemical nature. For example, signals 2 and 3 (Fig. 1) both originate from the same molecule (glycerol) and are therefore highly correlated. This is also true for the α- and β-carboxyl protons in fatty acids (signals 5 and 7).

From the literature, it was found that there are small differences between the fatty acid compositions of hazelnut oil and olive oil.[21–23] Hazelnut oil is characterized by a slightly higher content of oleic and linoleic acid and a lower content of linolenic acid with respect to olive oil. The relevant information in the NMR spectra can be obtained from peaks 1, 4, 6 and 9, where peaks 1 and 6 carry redundant information about olefinic protons (Table 1). Owing to the better suitability for correct and precise integration, we chose signal 1 for further evaluation and signal 4 for the polyunsaturated acids. As signals 2 and 3 arise from the glycerol protons they do not carry interesting information. From the other signals, 5, 7, 8 and 10, which provide redundant information, only peak 7 was taken into further consideration because of its easy integration, appropriate for quantitative purposes.

The three relative integrals of the peaks 1, 4 and 7 were chosen, which reflect the amount of unsaturated fatty acids, polyunsaturated fatty acids and fatty acids, respectively. Considering that the molar fraction of glycerol is very similar for all oils, one can consider the corresponding NMR signals as an internal reference. For that purpose we selected and normalized to 100 the part of signal 3 at lower field (4.4–4.2 ppm) because it was the only signal free from overlap with other resonances.

In order to evaluate the differences in the content of linolenic acid, two signal height ratios were also taken into account. The first considers signal 4 (2.9–2.6 ppm, Fig. 2), which is split into two parts representing the diallylic moiety of linolenic acid (signal at low field) and of linoleic acid (signal at high field). Because of the overlap of these two signals, it was necessary to use the peak picking procedure for evaluation. The height of the

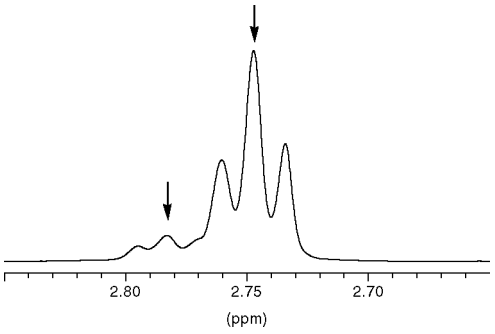

Figure 2. Expansion of signal 4: diallylic protons. The signal at the higher chemical shift arises from diallylic protons in linolenic acid and that at the lower chemical shift from linoleic acid.

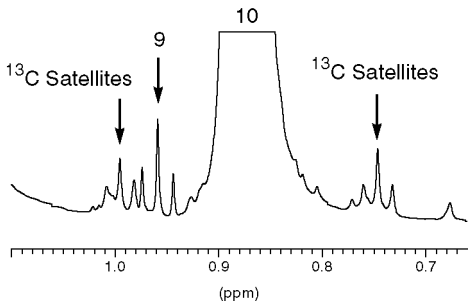

Figure 3. Expansion of signal 10: methyl group of linolenic acid (signal 9), and ^{13}C satellites of signal 10 (methyl groups of other fatty acids).

high-field signal was normalized to 100 and the relative height of the signal at the lower field was then retained for further data evaluation. The resulting ratio (ratio 1) gives the proportion of linolenic acid with respect to that of linoleic acid.

The second signal height ratio was obtained considering the height of the ^{13}C satellite of signal 10 (main satellite, low field, Fig. 3) and the height of signal 9, which corresponds to the methyl group of linolenic acid. This ratio (ratio 2) reflects the ratio of linolenic acid to all other fatty acids.

The three integrals and the two ratios described above constituted the dataset used for further statistical evaluation. Table 2 shows the single results, means and standard deviations obtained.

It can be observed that for olive oil the ratio of linolenic acid to all fatty acids >1.5. Compared with olive oil, hazelnut oil is enriched in unsaturated fatty acids (peaks 1 and 4) and characterized by lower ratios 1 and 2. Sunflower oils are clearly distinguished from the other oils

Table 2. Relative NMR integrals and signal height ratios for each sample of oil[a]

				Integral			Signal height	
No.	Oil	Type	Origin	Peak 1	Peak 4	Peak 7	Ratio 1	Ratio 2
1	Olive	Authentic	Italy	283.5	41.5	304.4	10.8	2.8
2	Olive	Authentic	Italy	278.4	33.3	320.9	12.4	2.1
3	Olive	Authentic	Italy	279.6	32.1	310.8	11.8	2.1
4	Olive	Authentic	Italy	272.5	30.6	302.1	12.7	1.8
5	Olive	Authentic	Italy	266.7	19.6	313.5	13.0	1.5
6	Olive	Authentic	Italy	291.7	35.5	318.1	13.0	2.7
7	Olive	Authentic	Italy	281.6	18.6	324.1	23.5	2.1
8	Olive	Authentic	Italy	293.0	37.5	318.1	9.5	2.1
9	Olive	Authentic	Italy	281.3	29.0	308.5	16.0	2.6
10	Olive	Authentic	Italy	279.4	22.4	310.9	19.2	2.2
11	Olive	Authentic	Italy	276.0	19.2	316.4	23.1	2.4
12	Olive	Authentic	Italy	273.7	18.2	316.9	23.1	2.0
13	Olive	Authentic	Italy	284.1	31.0	313.2	20.1	2.4
14	Olive	Authentic	Italy	279.0	24.1	313.5	16.8	2.2
15	Olive	Authentic	Italy	273.6	34.1	308.2	12.7	2.4
16	Olive	Authentic	Italy	289.4	25.0	314.5	16.9	2.4
17	Olive	Authentic	Italy	278.8	34.4	306.2	9.8	2.1
18	Olive	Authentic	Italy	296.4	51.8	315.8	9.5	2.8
19	Olive	Authentic	Italy	284.0	30.4	318.8	16.0	2.6
20	Olive	Authentic	Italy	280.0	32.3	320.1	13.6	2.6
21	Olive	Authentic	Italy	280.4	22.2	318.3	19.2	2.2
22	Olive	Authentic	Italy	279.8	22.9	314.2	18.6	2.3
23	Olive	Authentic	Italy	281.3	26.0	316.5	11.1	1.8
24	Olive	Authentic	Italy	278.6	22.1	312.6	17.3	2.0
25	Olive	Authentic	Italy	281.5	24.2	313.5	16.4	2.3
26	Olive	Authentic	Italy	280.3	26.5	309.8	16.5	2.2
27	Olive	Authentic	Italy	284.7	26.6	313.8	17.5	2.5
28	Olive	Authentic	Italy	280.8	25.9	312.2	17.9	2.4
29	Olive	Authentic	Italy	284.7	25.8	316.3	18.9	2.7
30	Olive	Authentic	Italy	269.2	21.1	312.4	19.6	2.2
31	Olive	Authentic	Spain	284.0	26.0	318.0	20.1	2.2

(*continued overleaf*)

Table 2. (*continued*)

No.	Oil	Type	Origin	Integral			Signal height	
				Peak 1	Peak 4	Peak 7	Ratio 1	Ratio 2
32	Olive	Authentic	Spain	274.8	23.0	316.1	20.2	2.3
33	Olive	Authentic	Spain	278.1	15.5	316.7	22.9	1.7
34	Olive	Authentic	Spain	273.3	14.6	312.0	27.6	1.9
35	Olive	Authentic	Spain	279.9	17.7	315.7	19.7	1.8
36	Olive	Authentic	Spain	288.0	27.0	320.0	13.5	1.9
37	Olive	Authentic	Spain	274.6	15.1	314.0	21.8	1.7
38	Olive	Authentic	Greece	279.7	22.3	313.7	16.7	1.9
39	Olive	Authentic	Greece	280.7	24.8	323.8	20.8	2.1
40	Olive	Authentic	Greece	285.5	25.3	325.5	17.1	2.1
41	Olive	Authentic	Hungary	273.2	21.3	323.9	24.9	2.2
42	Olive	Authentic	Tunisia	294.6	63.1	308.6	10.4	4.0
		Average		280.7	27.1	314.8	17.0	2.2
		Standard deviation		6.4	9.3	5.2	4.5	0.4
			Declared origin					
1	Hazelnut	Commercial	France	320.3	47.2	299.6	4.3	1.1
2	Hazelnut	Commercial	France	325.8	51.4	307.2	5.1	1.4
3	Hazelnut	Commercial	France	311.6	27.8	304.4	1.8	0.4
4	Hazelnut	Commercial	France	319.5	37.5	307.3	4.0	0.9
5	Hazelnut	Commercial	France	318.0	36.1	307.0	4.1	0.5
6	Hazelnut	Commercial	France	328.0	38.1	312.0	0.8	0.4
7	Hazelnut	Commercial	France	328.0	46.1	308.8	2.1	0.6
8	Hazelnut	Commercial	Unknown	339.1	57.4	325.5	3.5	1.1
9	Hazelnut	Commercial	Unknown	317.7	57.7	301.2	2.5	0.7
10	Hazelnut	Commercial	Unknown	307.3	33.5	299.3	3.1	0.5
11	Hazelnut	Commercial	Unknown	321.0	47.0	296.2	2.5	0.5
12	Hazelnut	Commercial	Unknown	302.8	28.1	295.8	3.4	0.8
13	Hazelnut	Commercial	Unknown	322.6	47.9	300.1	1.9	0.4
14	Hazelnut	Commercial	Unknown	317.5	43.9	299.0	2.6	0.6
15	Hazelnut	Commercial	Unknown	325.8	52.1	299.6	11.1	3.3
16	Hazelnut	Commercial	Unknown	315.2	42.0	297.8	2.7	0.6
17	Hazelnut	Commercial	Turkey	313.1	41.3	295.8	2.1	0.4
18	Hazelnut	Solvent extraction	Turkey	311.0	26.6	313.8	2.4	0.3
19	Hazelnut	Pressed	Turkey	329.0	27.8	333.0	2.2	0.3
20	Hazelnut	Pressed	USA	320.0	34.7	300.2	1.8	0.4
21	Hazelnut	Pressed	Italy	304.0	22.9	306.5	1.5	0.4
22	Hazelnut	Pressed	Italy	302.0	23.4	306.0	1.8	0.4
23	Hazelnut	Pressed	Italy	300.0	16.5	304.0	1.8	0.3
24	Hazelnut	Pressed	Italy	316.0	24.9	309.7	3.1	0.4
25	Hazelnut	Pressed	Italy	300.0	24.0	305.0	1.9	0.4
26	Hazelnut	Pressed	Unknown	301.0	22.0	303.0	1.8	0.4
27	Hazelnut	Pressed	Unknown	302.0	20.0	308.0	1.3	0.4
28	Hazelnut	Pressed	Unknown	320.0	40.2	308.6	1.8	0.4
29	Hazelnut	Pressed	Unknown	304.0	22.6	305.0	1.9	0.4
30	Hazelnut	Pressed	Unknown	317.0	35.3	306.0	1.7	0.5
31	Hazelnut	Pressed	Unknown	313.0	28.2	304.0	1.9	0.4
		Average		315.2	35.6	305.5	2.7	0.6
		Standard deviation		10.1	11.6	7.9	1.8	0.6
1	Sunflower	Commercial	Italy	456.2	199.4	294.1	Not detected	Not detected
2	Sunflower	Commercial	Italy	446.6	187.2	297.8	Not detected	Not detected
3	Sunflower	Commercial	Italy	440.4	181.8	298.2	Not detected	Not detected
4	Sunflower	Commercial	Italy	435.7	178.7	296.5	Not detected	Not detected
5	Sunflower	Commercial	Italy	447.1	190.3	298.9	Not detected	Not detected
6	Sunflower	Commercial	Italy	444.0	186.7	297.9	Not detected	Not detected
7	Sunflower	Commercial	Italy	451.8	194.3	298.1	Not detected	Not detected
		Average		446.0	188.3	297.4	—	—
		Standard deviation		6.8	7.1	1.6	—	—

[a] For the final calculation of the statistical model hazelnut oil, samples 2, 8, 9 and 15 were eliminated, because their authenticity cannot be guaranteed.

by a very high content of linoleic acid (peaks 1 and 4). Furthermore, the samples of sunflower oils that were investigated in this study were also characterized by the absence of linolenic acid. However, the latter must not be considered as a general rule, because it is known that sunflower oil may contain up to 2% of linolenic acid.[23]

In order to prove the significance of the selected variables, an analysis of variance (ANOVA) was performed, to prove that the null hypothesis (no statistical significant differences between the variances of the groups) is not valid. Table 3 summarizes the results when considering the three botanical origins and also when comparing only the variability between olive oil and hazelnut oil. In both cases all of the selected variables were significantly different for each group and therefore contain relevant information for discrimination purposes.

For the further statistical analysis using DA, the type of oil was chosen as grouping variable. Figure 4 shows the graph of the resulting discriminant functions Root 1 and Root 2. The oil types are clearly separated into three clusters. Root 1 mainly separates olive and sunflower oil, whereas hazelnut oil is separated from olive oil by both discriminant functions. The main contributions to Root 1 are by far the ratio 2 and the integral of peak 1, while Root 2 is mainly made up of the integral of peak 4 and 1 and ratio 1. Therefore, the above remark about the possible occurrence of a low proportion of linolenic acid in sunflower oil should not greatly affect the discrimination of the three types of oils considered.

To prove the reliability of the system, some oils were not included in the calculation and were considered as unknowns in further calculations. In order to obtain a

Table 3. Analysis of variance (ANOVA) of the selected intensity NMR data[a]

Variable	Peak 1	Peak 4	Peak 7	Ratio 1	Ratio 2
All 3 oils:					
F (2, 77)	1284.8	780.4	34.9	177.3	145.6
F critical ($p = 0.95$)	3.1	3.1	3.1	3.1	3.1
P	<0.000001	<0.000001	<0.000001	<0.000001	<0.000001
Olive and hazelnut oil:					
F (1, 72)	321.0	12.0	37.0	270.1	202.4
F critical ($p = 0.95$)	4.0	4.0	4.0	4.0	4.0
P	<0.000001	<0.000894	<0.000001	<0.000001	<0.000001

[a] The F values, with the degrees of freedom for the ANOVA in parentheses, are the test values for each variable to be compared with the table value (F critical) for the chosen probability of 95% ($p = 0.95$). P is the probability for the conformation of the null hypothesis.

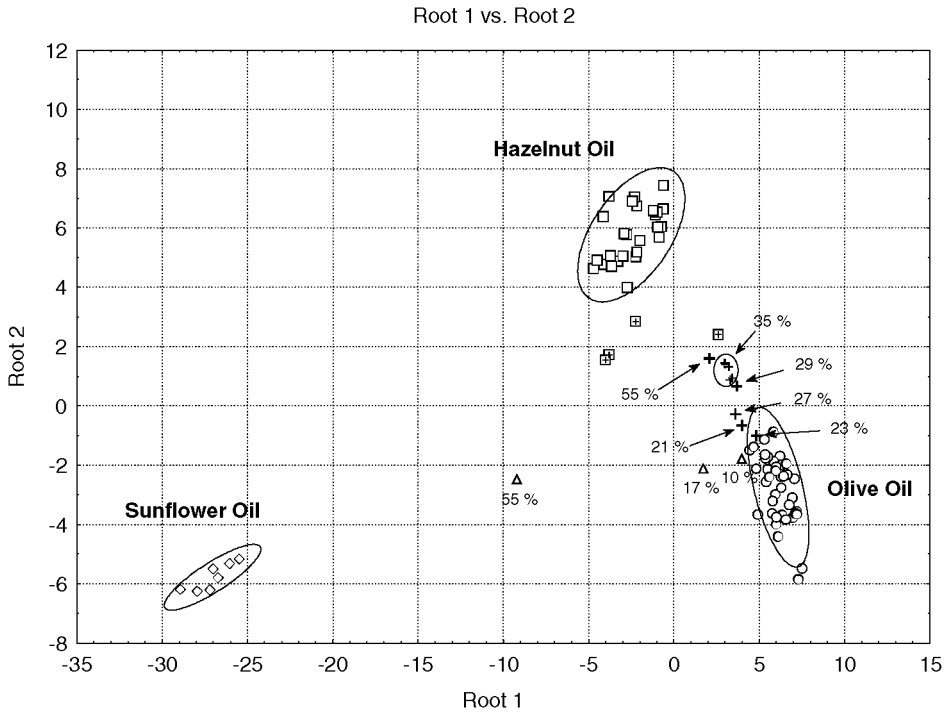

Figure 4. Plot of discriminant functions Root 2 over Root 1. Circles, olive oils; squares, hazelnut oils; diamonds, sunflower oils; triangles, mixtures of olive oil and sunflower oil, percentage indicates the proportion (%, m/m) of sunflower oil; crosses, mixtures of olive oil and hazelnut oil, percentage indicates the proportion (%, m/m) of hazelnut oil. The ellipses display the 95% confidence range for each group of oil. One mixture of olive oil and hazelnut oil (35%) was repeated five times (crosses in small ellipse) over a period of 4 months. This demonstrates very good repeatability and gives an idea of the overall reproducibility of the proposed method.

reliable system for the assessment of unknown samples, it is a prerequisite to have a lot of training samples for each distinct group. It has been observed that when DA is applied to only a few samples, the so-called 'lasso effect' occurs.[24] This will artificially show a good separation between the different groups. As a result, unknown samples could be absolutely wrongly classified. In order to prevent this and to ensure that the statistical model is stable, the system has to be checked using known test samples as 'unknown samples' to be identified. If these samples are correctly classified, the system may be used for real samples. At this stage it can be concluded that the system is stable and that it contains the relevant information for the required discrimination.

Three times a different and randomly selected set of oils composed of 14 olive oils, 10 hazelnut oils and three sunflower oils was removed from the data. With the remaining data (28 olive oils, 21 hazelnut oils and four sunflower oils), the model (training system) was calculated again. The excluded oils were then introduced to the system as unknowns (test set). Four hazelnut oils of commercial origin were found to be outside the hazelnut group. These oils were excluded from the calculation of the final model since their authenticity could not be guaranteed. In Fig. 4 these four hazelnut oils were included as unknowns denoted by crossed squares. For all remaining cases the classification was appropriate in all three runs. This clearly demonstrates the reliability of the system.

For practical use of the proposed method, one has also to take into account the long-term stability of the ^{1}H NMR measurement and the uncertainty of the quantitative data obtained. An estimation of the laboratory internal reproducibility for these determinations was made by repetition of the measurement for some samples of oils over a period of 4 months. An example of one of these repeated measurements is shown in Fig. 4: the analysis of one mixture of olive oil and hazelnut oil (35% hazelnut oil; crosses in small ellipse) was repeated five times, including sample preparation. As can be observed, the variability due to the measurement is small and much lower than the natural variability within the oil groups.

This statistical approach was also applied to the analysis of oil mixtures. Different mixtures of olive and hazelnut oil (crosses) and olive and sunflower oil (triangles) were prepared and analysed as unknowns. The results were calculated using the model obtained by the training set and are also presented in Fig. 4. The percentages of the adulterating oil (%, m/m) of each mixture are also given.

The ellipses represent the 95% probability area of each group, which means that an oil has a 95% probability of being assigned to the appropriate group. The triangles are mixtures of olive and sunflower oil. Even the mixture with only 10% of sunflower oil is found to be outside the 95% olive oil ellipse. One mixture of olive and hazelnut oil (23%) was found in the olive oil ellipse.

Adulteration mixtures with about 10% sunflower oil and those with about 25% hazelnut oil would fall outside the 95% ellipse of olive oil and would therefore be clearly recognized as suspicious. DA is normally not designed for quantitative evaluation, but a semi-quantitative indication of the composition of mixtures can be estimated from the plot for a suspicious sample. The system can be considered almost linear because the positions found for the mixtures are in agreement with what could be predicted, from the actual positions in the diagram, for the original oils which were used to prepare the various mixtures.

A precise quantitative determination of adulterated mixtures is not possible but an estimation of the proportion of added foreign oil can be obtained from a plot of the graphic results of the multivariate analysis. The method we propose cannot detect adulteration at a very low level, but it represents a concept for a robust and fast screening technique. In this study we have shown that by considering only the main constituents (Table 2) it is possible to achieve a good discrimination between oils of different botanical origin. Furthermore, the selected acquisition parameters allow one to obtain an exhaustive amount of quantitative information also for signals of minor compounds. It was found that these compounds are of little interest for the purpose of recognition of botanical origin since their variability is very large even within a single category of vegetable oil. However, for other purposes, as shown by Sacchi *et al.*,[3] minor compounds also provide interesting information concerning the geographical and varietal origin of olive oils.

With interest restricted to a few main signals, one could consider shortening the analytical time by reducing the number of scans in order to achieve a sufficient signal-to-noise ratio (S/N), which would allow the precise quantitative determination of all corresponding components. The S/N of the smallest signal that we considered for further statistical evaluation of the data (signal 9) is generally about 80, so that we can roughly estimate that according to the equation

$$\mathrm{S/N} = k\sqrt{NS}$$

a minimum of 150 scans should provide an S/N of about 50, which should be acceptable for precise quantification, thus reducing the total acquisition time to about 15 min, which is much shorter than that (>4 h) previously proposed in the study of Mannina *et al.*[14]

CONCLUSION

It has been demonstrated that ^{1}H NMR spectroscopy can be used for the determination of the botanical origin of oils and for the detection of mixtures. Compared with conventional chromatographic methods, ^{1}H NMR spectroscopy is faster and requires only a simple sample preparation. In contrast with conventional methods, which most often focus on the analysis of one class of specific components, NMR spectroscopy enables one to record, within a certain range, more or less all the constituents of a mixture in a single experiment. This feature makes it very interesting for the fast screening of large numbers of samples and for the building up of a comprehensive database of authentic products. Further development of the proposed method would require the use of advanced software allowing complete automation for the processing of NMR data.

Additional information from ^{13}C NMR could also easily be included and we are now investigating this possibility in order to extend the use of NMR as a tool for the characterization and authentication of edible oils. For practical use in control laboratories, it would be advisable to evaluate the reproducibility of the proposed method within several laboratories. For routine application of this methodological approach it would also be desirable to establish oil reference materials to calibrate the response for each laboratory.

Acknowledgements

The technical assistance of Veronica Holland is gratefully acknowledged. Carsten Fauhl thanks the European Commission for a postdoctoral grant. The authors thank DGXXI for additional funding of this project. Thanks are due to Professor G. J. Martin for constructive suggestions for the improvement of this work, to G. Remaud (Eurofins) for providing four samples of commercial hazelnut oil and Dr F. Angerosa (Istituto per l'Elaiotecnica) for providing some authentic hazelnut oil samples.

REFERENCES

1. Angerosa F, Bréas O, Contento S, Guillou C, Reniero F, Sada E. *J. Agric. Food Chem.* 1999; **47**: 1013.
2. Angerosa F, Camera L, Cumitini S, Gleixner G, Reniero F. *J. Agric. Food Chem.* 1997; **45**: 3044.
3. Sacchi R, Mannina L, Fiordiponti P, Barone P, Paolillo L, Patumi M, Segre A. *J. Agric. Food Chem.* 1998; **46**: 3947.
4. Sacchi R, Patumi M, Fondanazza G, Barone P, Fiordiponti P, Mannina L, Rossi E, Segre AL. *J. Am. Oil Chem. Soc.* 1996; **73**: 747.
5. Zamora R, Navarro JL, Hidalgo F. *J. Am. Oil Chem. Soc.* 1994; **71**: 361.
6. Defernez M, Kemsley EK, Wilson RH. *J. Agric. Food Chem.* 1995; **43**: 109.
7. Forveille L, Vercauteren J, Rutlegde D. *Food Chem.* 1996; **57**: 441.
8. Lai YW, Kemsley EK, Wilson RH. *J. Agric. Food Chem.* 1994; **42**: 1154.
9. Vogels JTWE, Terwel L, Tas AC, Van den Berg F, Dukel F, Van der Greef J. *J. Agric. Food Chem.* 1996; **44**: 175.
10. Belton PS, Colquhoun IJ, Kemsley EK, Delgadilo I, Roma P, Dennis MJ, Sharmann M, Holmes E, Nicholson JK, Spraul M. *Food Chem.* 1998; **61**: 207.
11. Blanch GP, Caja M, Ruiz del Castillo ML, Herraiz M. *J. Agric. Food Chem.* 1998; **46**: 3153.
12. Ruiz del Castillo ML, Caja M, Herraiz M, Blanch GP. *J. Agric. Food Chem.* 1998; **46**: 5128.
13. Spangenberg J, Macko S, Hunziker J. *J. Agric. Food Chem.* 1998; **46**: 4179.
14. Mannina L, Patumi M, Fiordiponti P, Emanuele MC, Segre AL. *Ital. J. Food Sci.* 1999; **11**: 139.
15. Friebolin H. *Ein- und Zweidimensionale NMR Spektroskopie*. VCH: Weinheim, 1988.
16. *STATISTICA for Windows* [Computer program manual]. StatSoft: Tulsa, OK, 1997.
17. Johnson LF, Shoolery JN. *Anal. Chem.* 1962; **34**: 1136.
18. Sacchi R, Addeo F, Paolillo L. *Magn. Reson. Chem.* 1997; **35**: 133.
19. Manly BFJ. *Multivariate Statistical Methods: a Primer* (2nd edn). Chapman and Hall: London, 1994.
20. Otto M. *Chemometrie: Statistik und Computereinsatz in der Analytik*. VCH: Weinheim, 1997.
21. Angerosa F, Mascolo A, Cucurachi A. *Ann. Ist. Sper. Elaiot.* 1981–83; **9**: 3.
22. Parcerisa J, Boatella J, Codony R, Rafecas M, Castellote AI, García J, López A, Romero A. *J. Agric. Food Chem.* 1995; **43**: 13.
23. Scherz H, Senser F (eds). *Food Composition and Nutrition Tables* (5th edn). Medpharm Scientific: Stuttgart, 1994.
24. Kaufmann A. *J. AOAC Int.* 1997; **80**: 665.

