CHAPTER 3

NMR Spectroscopy of Natural Substances

B. Diehl

Contents

Abstract

The chapter provides an overview of several options in natural substances analysis by means of examples from routine nuclear magnetic resonance (NMR) spectroscopy. The objects are different chemical classes consisting of reference materials up to complex mixtures of plant extracts and lipids. One- and two-dimensional spectra were applied to demonstrate the characterisation and quantification of natural substances.

Keywords: finger print analysis, primary reference standard, multi component analysis, lipids, lecthins, fats, oils

1. Introduction

Analysis of natural substances is a traditional domain of nuclear magnetic resonance (NMR) spectroscopy and the quantitative applications are in progress.[1] Most of all compounds of interest for pharmaceutical approaches are derived from structures containing one or more chiral centres. This is obvious for the chemical structures of amino acids, carbohydrates or terpenoids, e.g. even the huge class of lipids based on the glycerol backbone undergo a chiral principle although glycerol

itself is not chiral (see Part I, Chapter 1, Section 9). The high pharmaceutical potential of natural compounds often is confronted with an insufficient availability of corresponding primary reference standards. The synthesis of analytical standards to quantify these natural substances is very complicated and expensive or even not possible. To produce an analytical standard, the active compounds must be extracted from their plant or animal sources and purified as much as possible. Following these procedures, there it is a need for an analytical method that enables an ab initio determination of the purified active compound to create a primary standard, e.g. chromatography. ^{1}H NMR spectroscopy was proved to fulfil this request. A round robin test for several natural substances[2] was accomplished successfully.

Besides the definition of standards for chromatographic purposes, NMR spectroscopy is a suitable qualitative and quantitative method to analyse complex mixtures based on natural extracts without any standard of the corresponding natural compounds – only by using one well-defined artificial standard. The molar response in NMR spectra gives the condition for this procedure. Both the application models – the creation of primary reference standards and the analysis with unrelated artificial reference material – will be discussed in this chapter. Owing to the enormous number of papers concerning NMR spectroscopy of natural substances, a complete overview goes beyond the scope of this chapter.

This chapter provides three different approaches in natural compound analysis. The first is a fingerprint-like analysis for characterisation, e.g. plant extracts. The second part deals with the qualitative and quantitative analysis of single natural substances, and the third with a quantitative or semi-quantitative analysis of multi-component systems.

2. Characterisation of Natural Substances as a Finger Print Analysis

The characterisation of pharmacologically relevant plant extracts with chromatographic methods is often complicated because of the heterogeneous composition. Carbohydrates, lipids, amino acids, terpenoids and other polymeric compounds cannot be detected using only one chromatographic method for different reasons. On the one hand, the volatility necessary for gas chromatography is not given; on the other hand, the target molecules have no chromophore as required for high-performance liquid chromatography (HPLC) analysis with UV/VIS detection. Chromatography-based methods only provide an information of a minor or a major part of the whole composition. In a ^{1}H NMR spectrum of a plant extract almost all soluble molecules can be determined, because NMR detection is based on a molecular response. A proton of a mono- or oligosaccharide can be detected in the same analysis like a sugar molecule, which is part of a glycolipid or a flavanoid. This fact may be demonstrated by a ^{1}H NMR spectrum of a plant–oil mixture of lemon peel and olive leaves. The observed NMR spectrum is characteristic and can further be used for the identification of this pharmaceutical raw material. In serial analysis, a computer-based automatic spectra analysis is possible by different pattern recognition techniques. The knowledge of all signals is not necessary for this easy procedure.

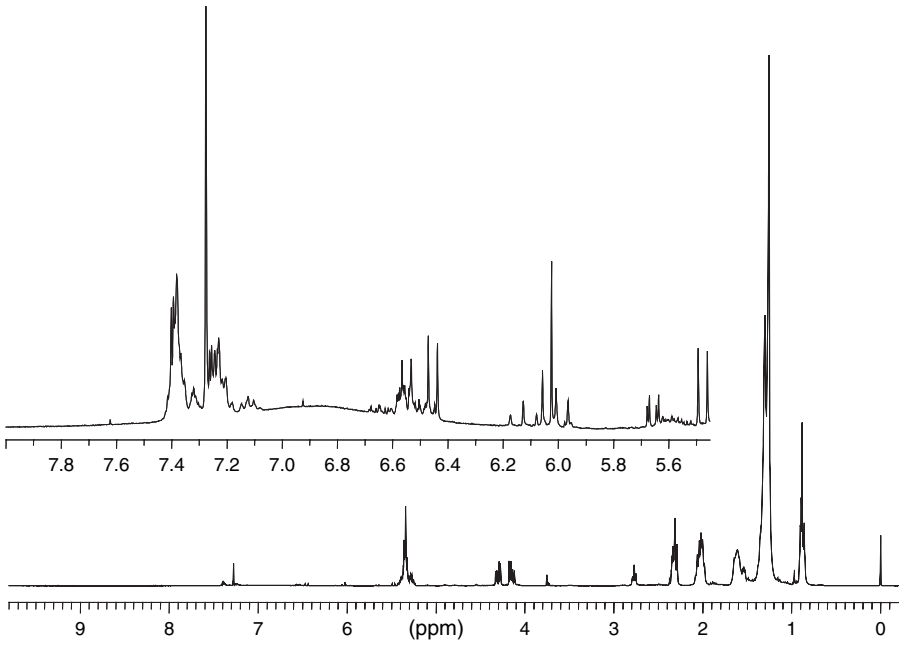

Figure 1 ^{1}H NMR spectrum of *Calophyllum inophyllum* and a detail of the aryl region.

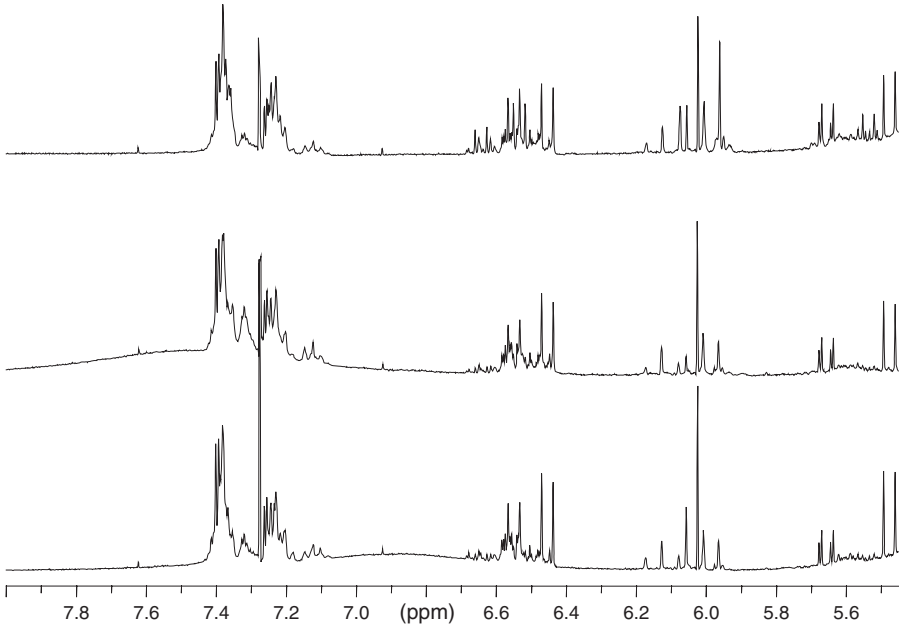

Figure 2 ^{1}H NMR spectra of different oil batches, downfield details in comparison.

In Figure 1 a ^{1}H NMR spectrum of a vegetable oil derived from *Calophyllum inophyllum* is given. The typical signals of triglycerides dominate the spectrum. Furthermore, a typical pattern of calophyllolide and inophyllum in addition to other complex polyphenols can be observed especially in the aromatic region of the spectra. Detailed structure analysis and NMR data are published.[3,4]However, the unassigned ^{1}H NMR data provide a fast tool for testing the origin and the quality of the oil (Figure 2). The simultaneous detection of both the fatty acid composition and the secondary plant compounds using the NMR method provides more information compared to chromatographical methods which of course are used for validation purposes.

3. Characterisation of Natural Substances According to Single Target Molecules

The quantification of the thermo instable molecule allicin in garlic products is only one routinely performed application[5] of NMR spectroscopy within a number of plant extracts. Analysing mixed extracts of lemon peel and olive leaves demonstrates the NMR tool in a more detailed way (Figure 3). The main target molecules of the plant extracts of citric peel is naringin and of the olive part oleuropein (Figure 4), both members of the family of flavanoids. Identification of these leading compounds is possible by ^{1}H NMR and ^{13}C NMR spectroscopy. NMR data of many different types of flavanoids are available from the literature.[6]

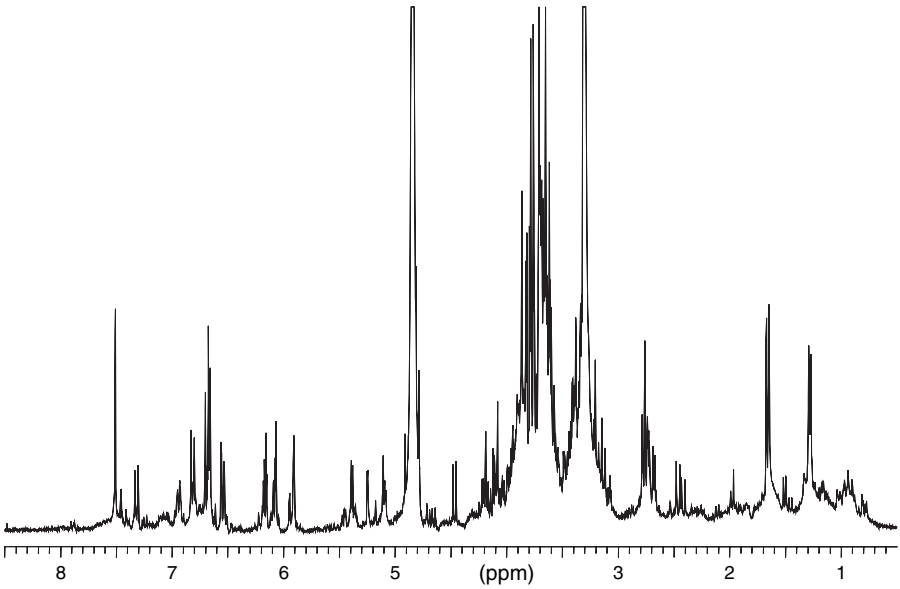

Figure 3 ^{1}H NMR spectrum of a plant extract mixture of lemon peel and olive leaves.

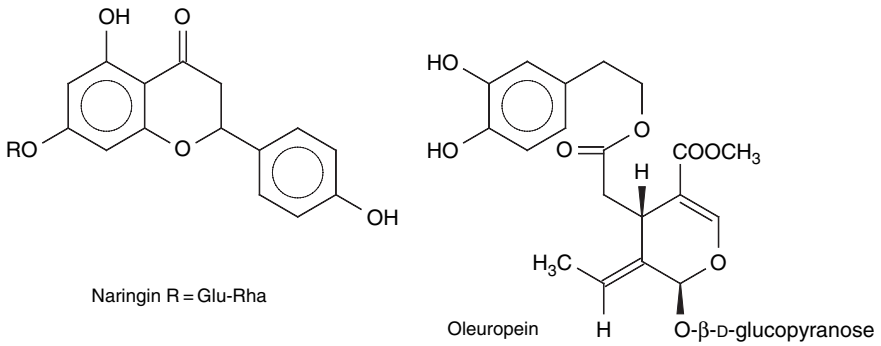

Figure 4 Chemical structures of the leading compounds of lemon peel and olive leaves extracts.

In the expansion of the ^{1}H NMR spectrum (Figure 5) the specific signals can be assigned by comparison with the spectra of reference materials. In case that minimum one characteristic signal of a target molecule does not interfere with signals of other molecules in the mixture, a direct quantification can be done by standard addition or by the use of a suitable internal standard.

These conditions are given within some restrictions in the above-mentioned example. A high field strength of the NMR instrument is a helpful component

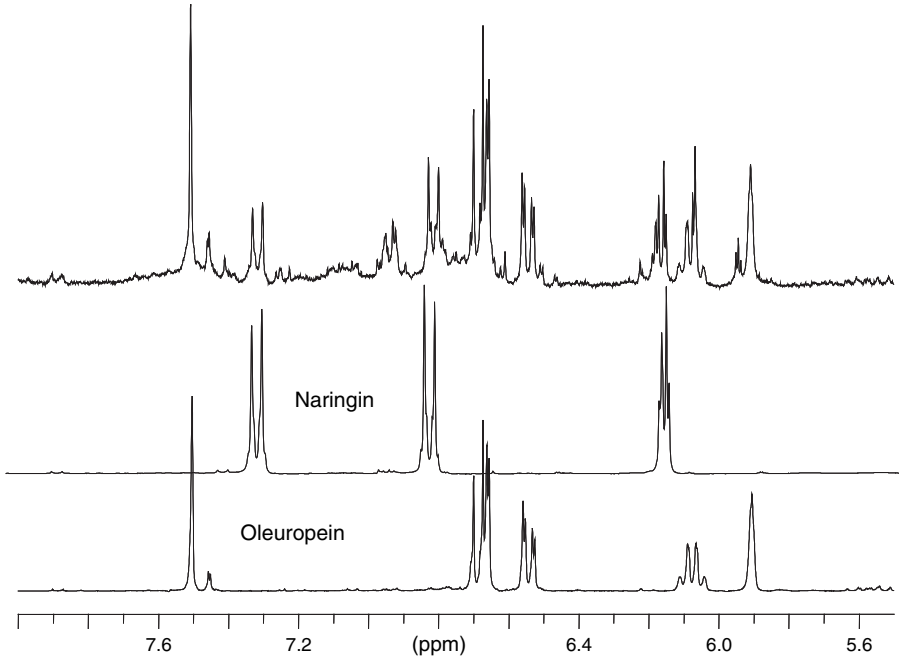

Figure 5 Aromatic region of the ^{1}H NMR spectrum of a plant extract mixture of lemon peel and olive leaves (top) compared with the corresponding spectra of naringin and oleuropein.

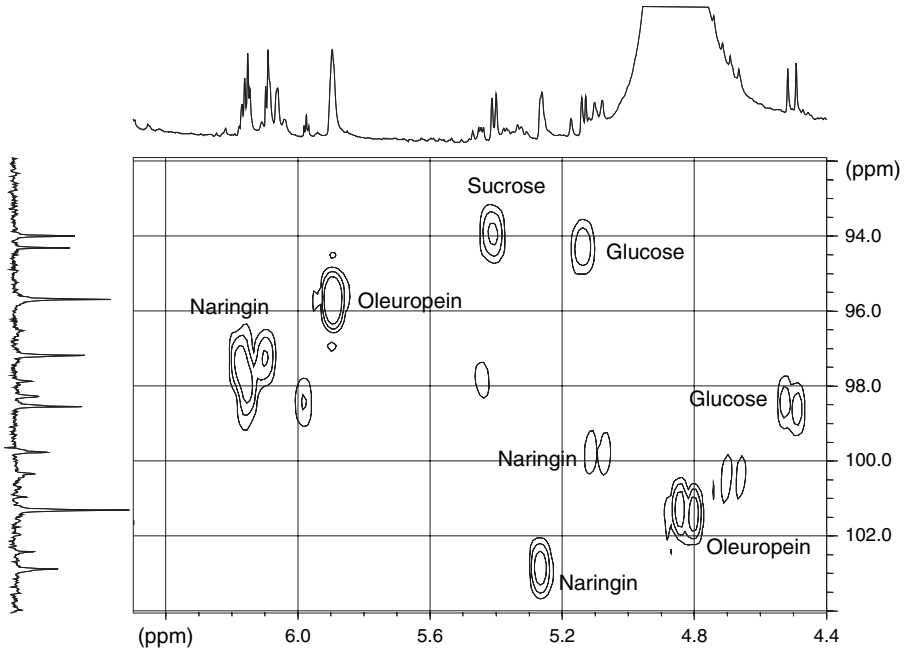

Figure 6 HMQC spectrum of a plant extract mixture of lemon peel and olive leaves

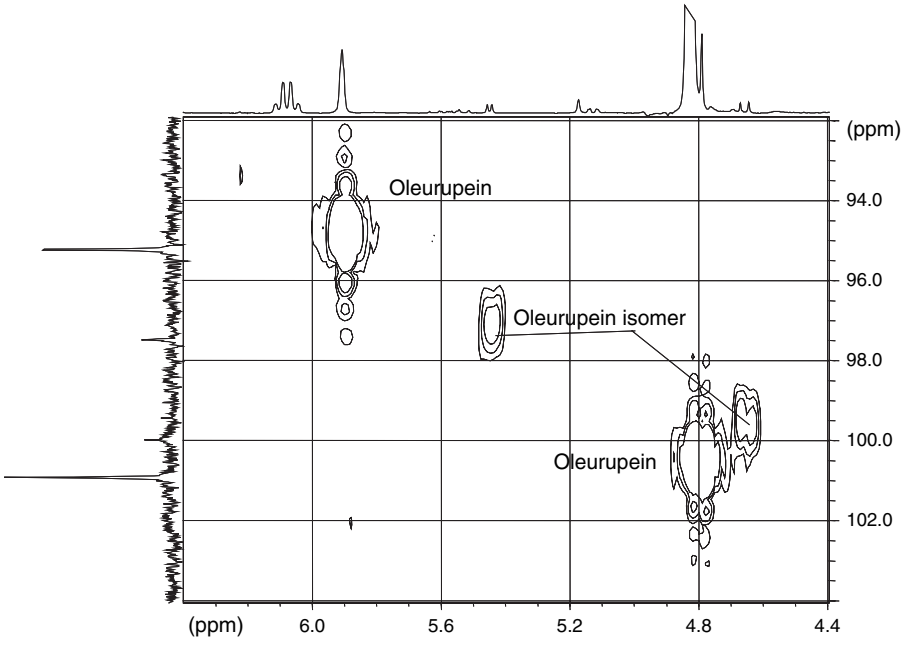

Figure 7 HMQC spectrum of reference item oleuropein (details).

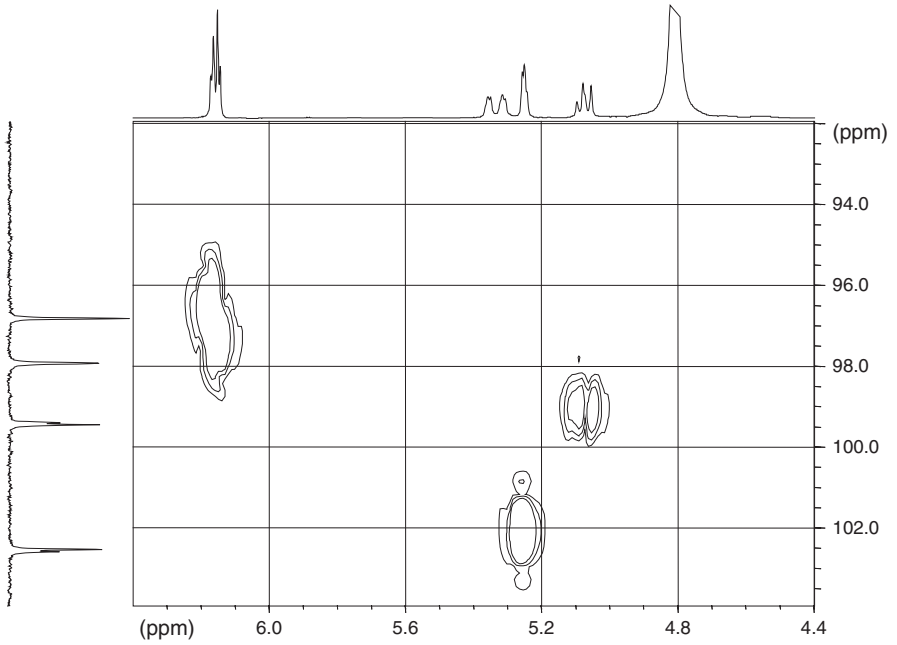

Figure 8 HMQC spectrum of reference item naringin (details).

in this sensible type of analysis. In the near future, modern cryoprobes with a magnitude better signal-to-noise ratio will enable this complex analysis by ^{13}C NMR spectroscopy, which shows a much higher signal dispersion.

In addition, two-dimensional analysis is very helpful in structure elucidation or even only as a fingerprint analysis as demonstrated on the discussed plant extract mixture in Figures 6–8.

4. Characterisation and Definition of Primary Reference Standards

One step further in NMR analysis of natural substances is the characterisation and the definition of primary and secondary reference standards.[7] Corresponding standards of natural compounds are often very expensive or the purities respectively the contents are low. Again the molar response of ^{1}H NMR signals can be used to qualify natural substances as reference standards. Commercially available standards of naringin and oleuropein were characterised by ^{1}H NMR spectroscopy (Figures 9 and 10) and the contents of both were evaluated by ^{1}H NMR using the internal standard method.[8] The detail of the aromatic region may demonstrate the procedure by the integration of the characteristic signals of the test item naringin and the internal standard 4-*N*,*N*-dimethylaminobenzoic acid ethylester (DBEE) (Figure 11). Oleuropein shows an isomer structure of approximately 15% that was

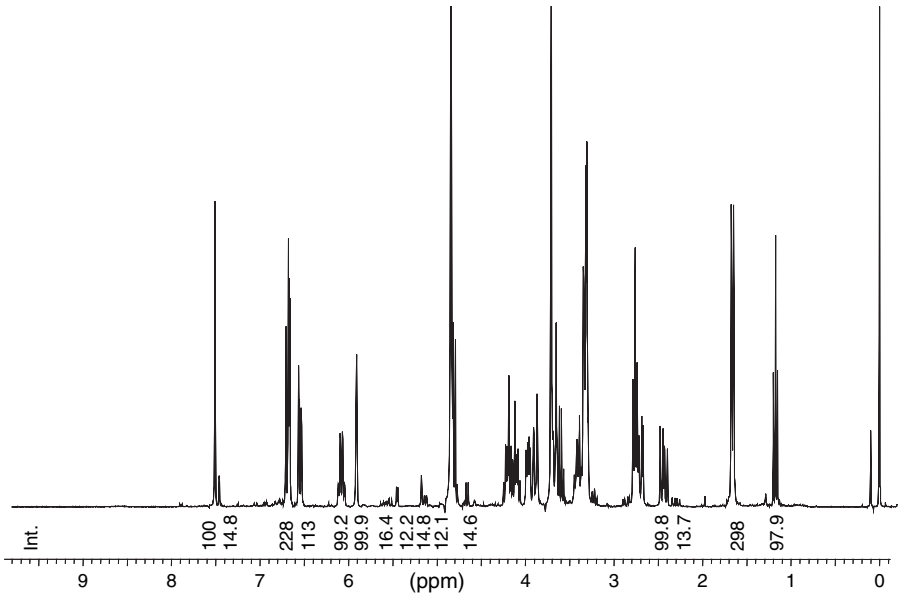

Figure 9 ^{1}H NMR spectrum of reference item oleuropein.

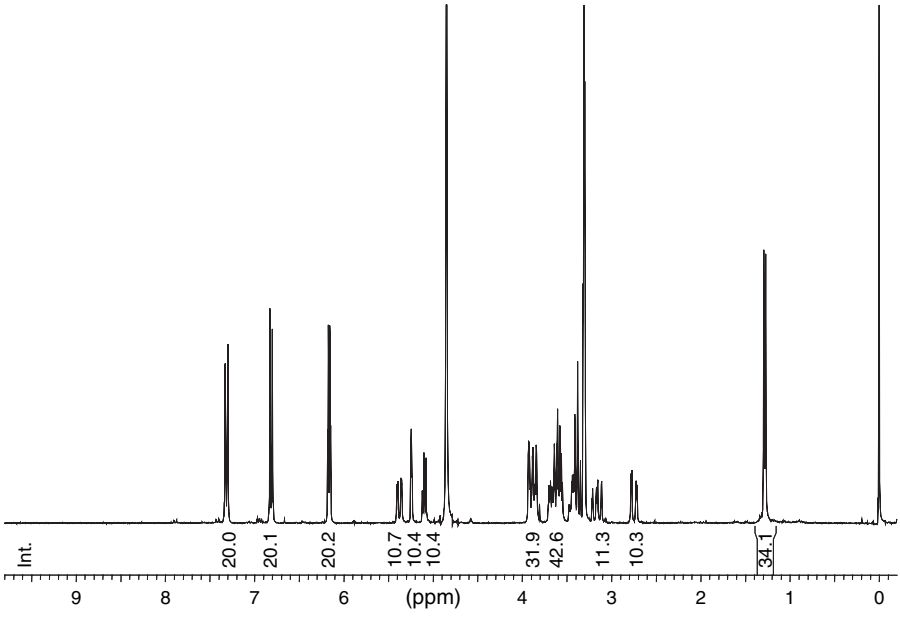

Figure 10 ^{1}H NMR spectrum of reference item of naringin.

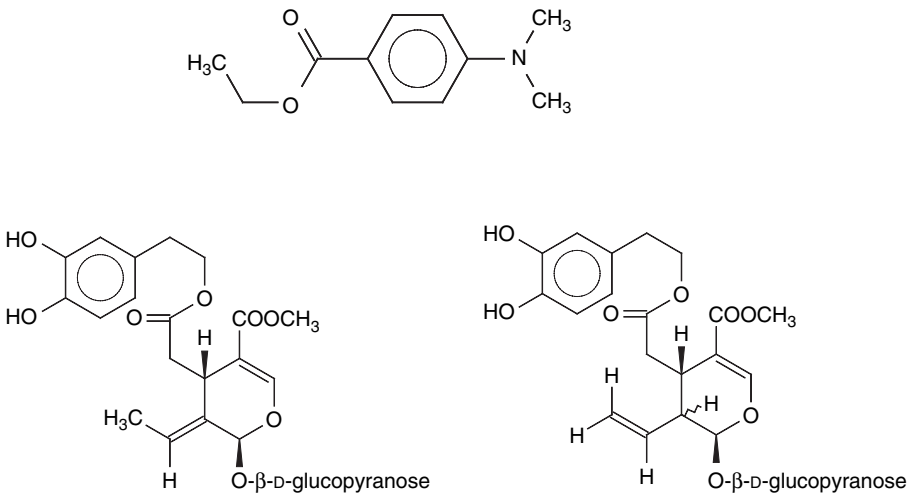

Figure 11 Chemical structure of DBEE used as internal standard (top) and the oleuropein isomers.

confirmed by HPLC/MS (mass spectrometry). The quantification was done in the same analysis.

Natamycin is an antimycotic fungicide. It is commercially available in a formulation with glucose. Amounts of glucose and natamycin, both very different in their chemical structures, can be quantified after dissolving in a mixture of DMSO-d_6 and methanol-d_4 (MeOD). The latter is used for deuterium exchange to avoid integration disturbances by sugar OH signals. A comparison of natamycin and glucose integral areas enables the evaluation of relative amounts of both compounds, which can be sufficient as a quality test (Figure 13).

For exact quantification a suitable internal standard, e.g. nicotinic amide (NA), must be used in a single or multiple analysis. In the present study the downfield signals (11H) of the fungicide are used for quantification as a sum of several proton signals. Additionally, other integral areas may be used, e.g. signals at $\delta = 2.8$ ppm representing 1 proton and $\delta = 0.9$–1.3 ppm representing 10 protons. Simultaneous signal integration of monomeric carbohydrates is complicated due to the equilibrium of different anomeric molecules. For glucose, the sum of one proton of each type α-D-glucose ($\delta = 5.0$ ppm) and β-D-glucose ($\delta = 4.35$) must be used for quantification (Figure 13).

In case that the signal of β-D-glucose interferes with the water signal, alternatively the double doublet at $\delta = 3.0$ ppm can be used for quantification. Furthermore, the water signal can be shifted by changing the solvent ratio between DMSO-d_6 and MeOD, as DMSO-d_6 is shifting the water signal upfield.

To enhance the confidence in this quick analysis, a defined amount of a glucose standard can be added (standard addition method) for further analysis under the same conditions. The increase of the signals of one or both anomeric glucose types enables an absolute quantification of glucose independent of the use of an internal standard. A combination of the internal standard and the glucose standard addition

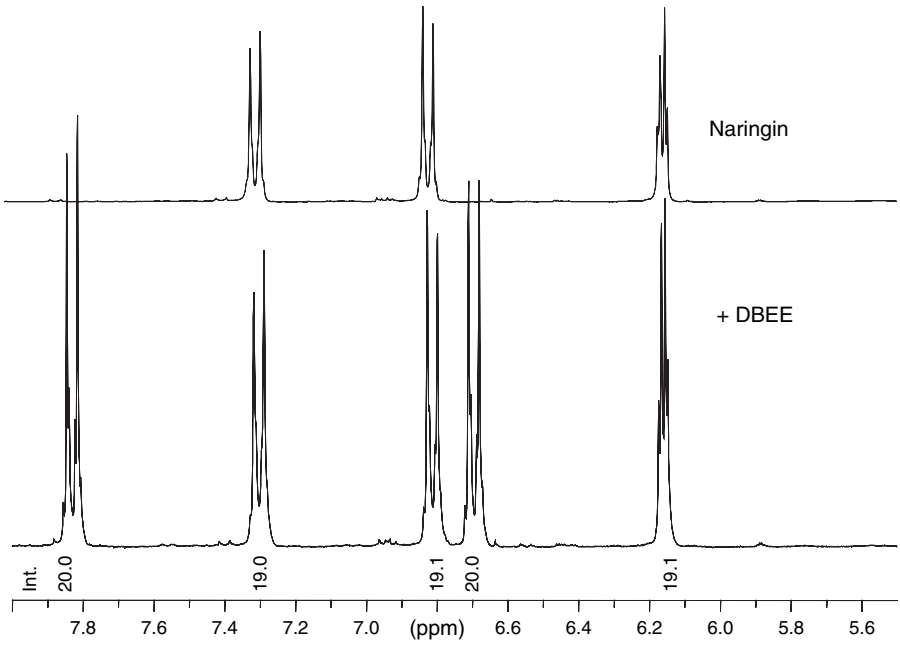

Figure 12 ^{1}H NMR spectrum of reference item of naringin with (bottom) and without (top) internal standard DBEE.

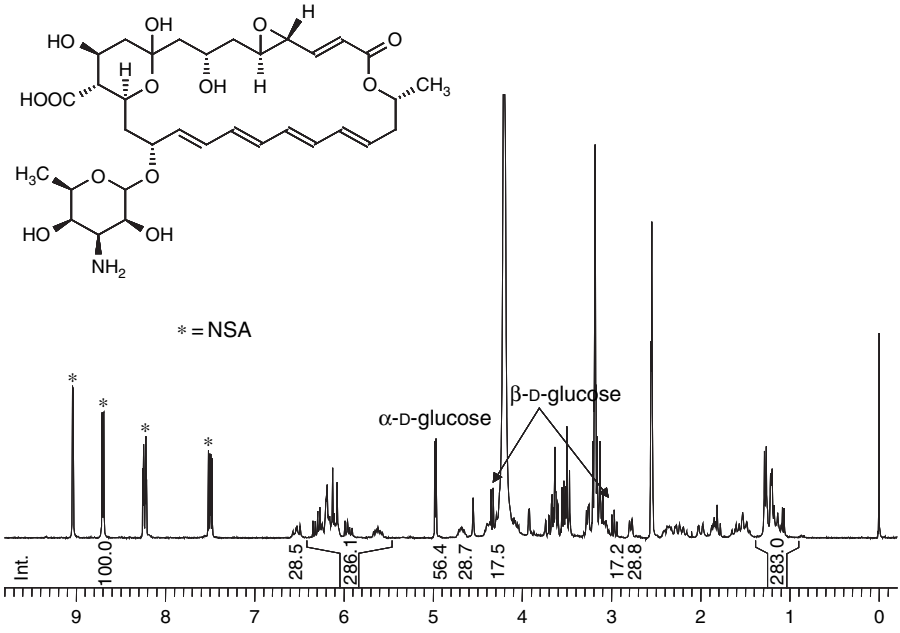

Figure 13 Natamycin glucose formulation in DMSO-d_6/MeOD, internal standard NA.

method allows a complete validation of the qNMR (quantitative nuclear magnetic resonance) method for this preservative formulation.

5. Multi-Component Analysis of Complex Natural Substance Mixtures and Phytopharmaceutics

5.1. Extracts from medical plants

Multi-component analysis by NMR spectroscopy in phytopharmaceutics is a further development. The analysis of extracts from *Aloe vera* (*barbadensis* Miller) will demonstrate the procedure. As many others *A. vera* extracts consist of a mixture of natural substances classes, e.g. amino acids, fruit acids (derived from the citric acid cycle), mono and polysaccharides, lipids and flavanoids.

First the NMR analysis is used to test the origin of an extract and to distinguish between different plant sources. The origin test is based on the presence and the amount of the biopolymer aloverose, a partly acetylated polymannose (see Chapter 2, Figure 5). Besides the original test, it is necessary to find out whether a product was exclusively taken from the inner gel of an *A. vera* leaf or whether it is a product from the whole leaf (WL) – including substances from the green rind. The composition of fruit acids deriving from the citric acid cycle can be used as marker to distinguish between these raw materials. Only in the green part of an *A. vera* leaf citric and isocitric acid are produced (Figures 14–16), the latter is partly

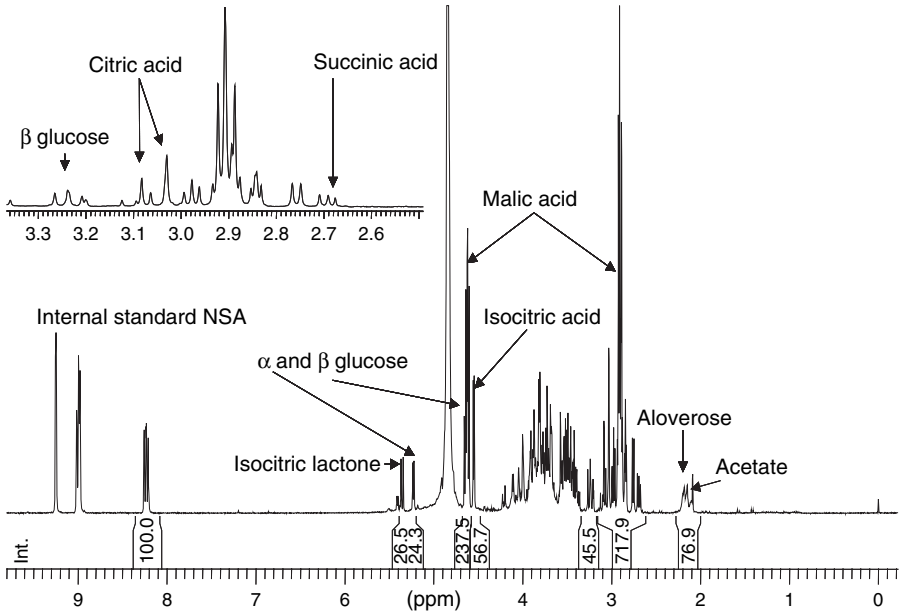

Figure 14 ^{1}H NMR spectrum of *A. vera* whole leaf + NSA as internal standard.

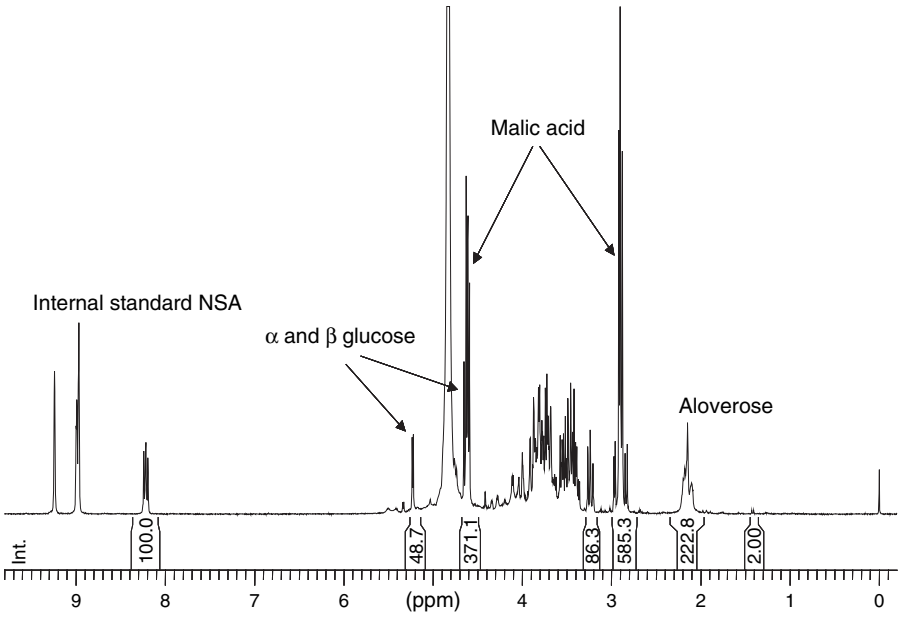

Figure 15 ^{1}H NMR spectrum of fresh *A. vera* gel + NSA as internal standard.

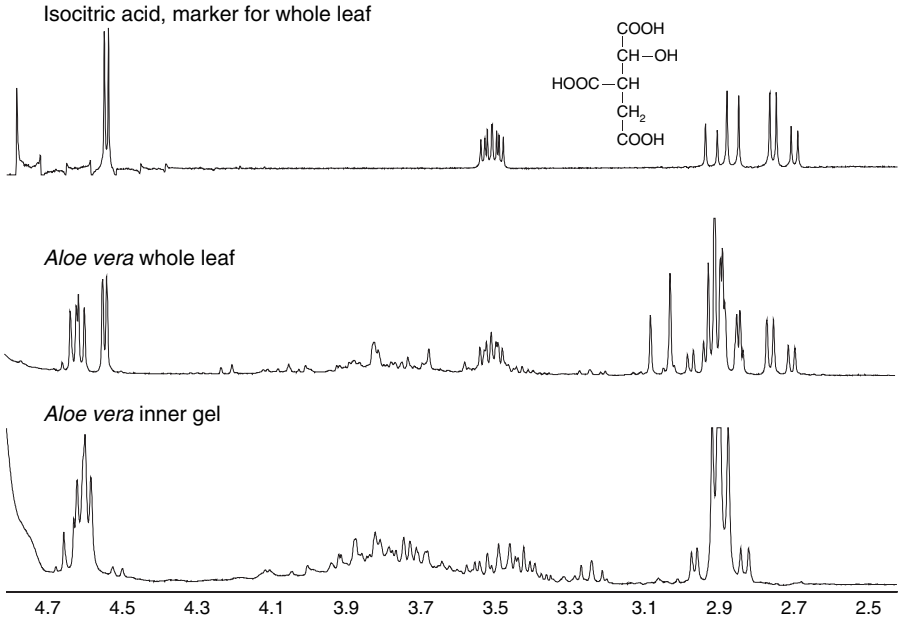

Figure 16 Isocitric acid (top), WL *A. vera* (middle) and *A. vera* gel (bottom).

converted to the corresponding lactone. If both the molecules can be detected in an *A. vera* product, the sample is therefore either a WL material or it is contaminated with parts of the rind.

Using the internal standard method, the characterisation of *A. vera* samples can be done quantitatively or at least semi-quantitatively. Malic acid as well as glucose and the polymer aloverose are main constituents of fresh *A. vera* gel and are quantified routinely by ^{1}H NMR. Additional signals originated from WL material are quantified in the same way (Figure 14).

For validation purposes the following procedure was performed. An *A. vera* leaf was split into two parts. From one part the inner filet was taken. A juice was prepared by using a blender (GEL). The second part of the leaf was equally processed without removing the rind to give a WL product. Both samples were analysed separately as well as in mixtures of different ratios by ^{1}H NMR. The increasing signals of the WL markers (isocitric acid and its lactone) correspond to the amount of *A. vera* WL material used.

To complete the quality control of *A. vera* products, besides their main constituents, chemical and bacterial degradation products are also detected as well as additives and adulterations, e.g. preservatives or maltodextrin. The detection is independent of the chemical structures (mono or polymer) and does not need an UV-active chromophore as been necessary in HPLC/UV analyses. Also, volatility, a condition for gas chromatographical analysis, is not required (Table 1).

5.2. Lipids

NMR spectroscopy has become a universal method in lipid analysis. It allows structure elucidation, qualitative and quantitative analysis of defined molecules and even complex mixtures. Not all nuclei are accessible to the NMR experiment.

Table 1 NMR chemical shifts of compounds in *A. vera* products

Compound	δ (ppm)[a,b]	No. of protons
Formic acid	**8.4 (CH)**	1
Fumaric acid	**6.5 (2 × CH)**	2
Benzoic acid	8.0, 7.6, **7.5 (CH)**	2
Sorbic acid	7.0, 6.2, 5.8, **1.8 (CH_3)**	3
Glucose	**5.3(α-CH)**, 4.6 and **3.2 (β-CH)**, 3.2–4.0	1[c]
Malic acid	**4.3 (CH)**, 2.4–2.8 **(CH)**	1
Succinic acid	**2.5 (2 × CH_2)**	4
Citric acid	**2.4–2.8 (2 × CH_2)**	4[d]
Acetic acid	**1.9 (CH_3)**	3
Lactic acid	4.1, **1.3 (CH_3)**	3
Isocitric acid	**4.55 (CH)**, 3.5 2.7–3.0	1
Isocitric lactone	**5.35 (CH)**, 3.5 2.9–3.2	1

[a]Chemical shifts are pH dependent and may change slightly.
[b]Bolded signals normally are used for calculation.
[c]Mix of anomeric forms. The integral value of the α-type has to be added to the β-type.
[d]Integration is complicated by interference with malic acid signals.

But those nuclei that are important in lipid chemistry like ^{1}H, ^{13}C and ^{31}P are recordable as a matter of routine with modern instruments equipped with a multi-nuclear probe.

Compared to common chromatographic analysis, e.g. gas chromatography (GC) and HPLC, NMR spectroscopy is a non-destructive ab initio method. Analysis of oils and lecithin of different origin demonstrates the usefulness of multi-nuclear high-resolution NMR spectroscopy.[9] Besides the composition of the main components, e.g. fatty acid and phospholipid composition, secondary substances such as sterols or terpenoids can be determined within the same analysis.

5.3. Lecithin

The 100% amount of the ^{31}P isotope in nature, the selectivity and the sensitivity make ^{31}P NMR spectroscopy the reference method of the ILPS (International Lecithin and Phospholipid Society). Artificial standards, e.g. triphenyl phosphate (TPP) or distearoyl phosphatidyl glycerol (DSPG), are used to quantify all phospholipids from complex mixtures of any origin. The quantification of phospholipids is a valid method. Method precision (reproducibility) and instrument precision for main components show a standard deviation $<1\%$, components near the limit of quantification below 5%.

In the case of liposome lyophilisate preparation, the extraction recovery of the phospholipids is 100%. The analysis method is robust. It is almost insensitive to variation in sample preparation and parameters of measurement. Prepared samples can be stored up to 5 days at room temperature (typically <24 h) until they are measured without negative influence on the results. The internal standard method had been fully validated concerning selectivity, recovery, reproducibility, instrument precision and robustness. A lyophilised drug formulation on the basis of egg yolk phosphatidyl choline (PC) was used. The results of all experiments are summarised in Table 2.[10]

For further validation studies, a round robin test involving more than 23 NMR labs had been done for approving the ^{31}P NMR as an official method.[11] The measurements were performed on different field strength between 200 and 700 MHz and by using 10 different samples. Five phospholipid mixtures for pharmaceutical approaches were tested – native soya bean lecithin, purified

Table 2 Survey on validation experiments

Test	*n*	PC by weight %
Recovery	7	33.3 ± 0.2
Instrument precision	6	33.4 ± 0.3
Variation of pH value	7	33.3 ± 0.2
Variation of solvent mix	5	33.2 ± 0.2
Variation of sample amount	7	33.4 ± 0.2

n = number of tests.

Table 3 Abbreviation of phospholipids

PC	Phosphatidyl choline	APE	*N*-Acylphosphatidyl ethanolamine
PI	Phosphatidyl inositol	DPG	Diphosphatidyl glycerol, cardiolipin
2-LPC	2-Lysophosphatidyl choline	PG	Phosphatidyl glycerol
PS	Phosphatidyl serine	PA	Phosphatidic acid
PE	Phosphatidyl ethanolamine	LPA	Lysophosphatidic acid
LPE	Lysophosphatidyl ethanolamine	SPH	Sphingomyelin

phosphatidyl serine, synthetic distearyl phosphatidyl choline, de-oiled egg yolk lecithin and a crude reaction mix of an enzymatically prepared phosphatidyl serine (based on soya lecithin). The results demonstrate the robustness of the NMR analysis, the deviation (mean of all five experiments) from the mean value is given in Figure 17. Only 3 of 29 experiments failed due to technical problems or by the use of irregular parameters.

Figure 18 shows the phospholipid profile of native soya bean lecithin. The NMR data evaluation of this kind of ^{31}P NMR spectra is comparable to usual chromatographic methods, but comprising the advantages of optimised selectivity, linearity and dynamics.

Figure 19 shows a medical application, the ^{31}P NMR of human blood plasma.

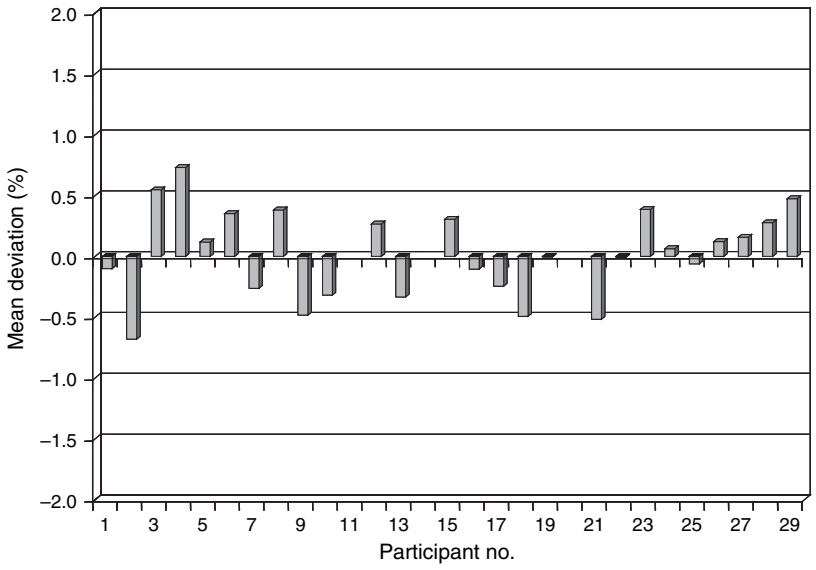

Figure 17 Results of the round robin test on ^{31}P NMR.

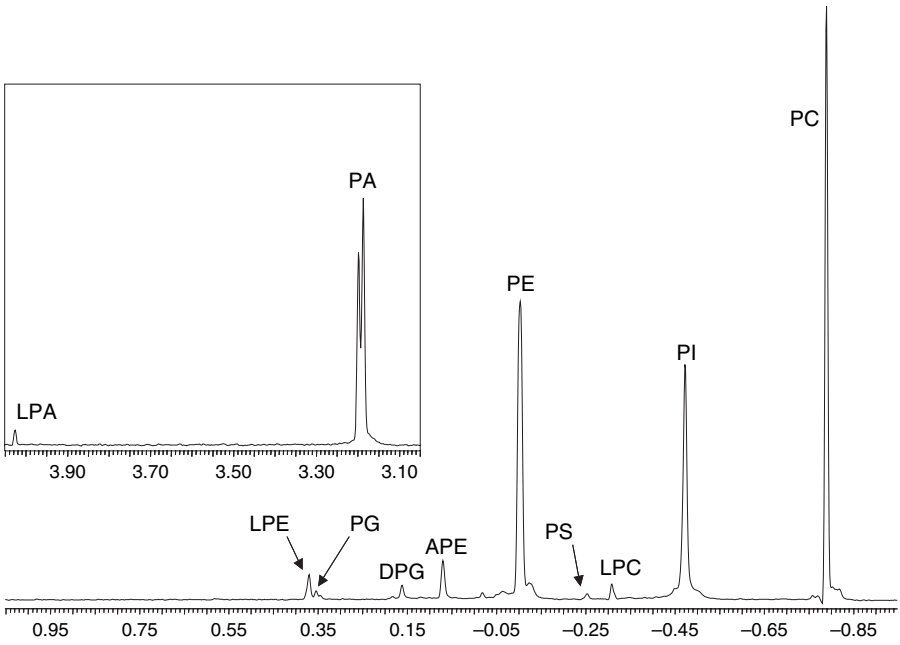

Figure 18 ^{31}P NMR of soya bean lecithin.

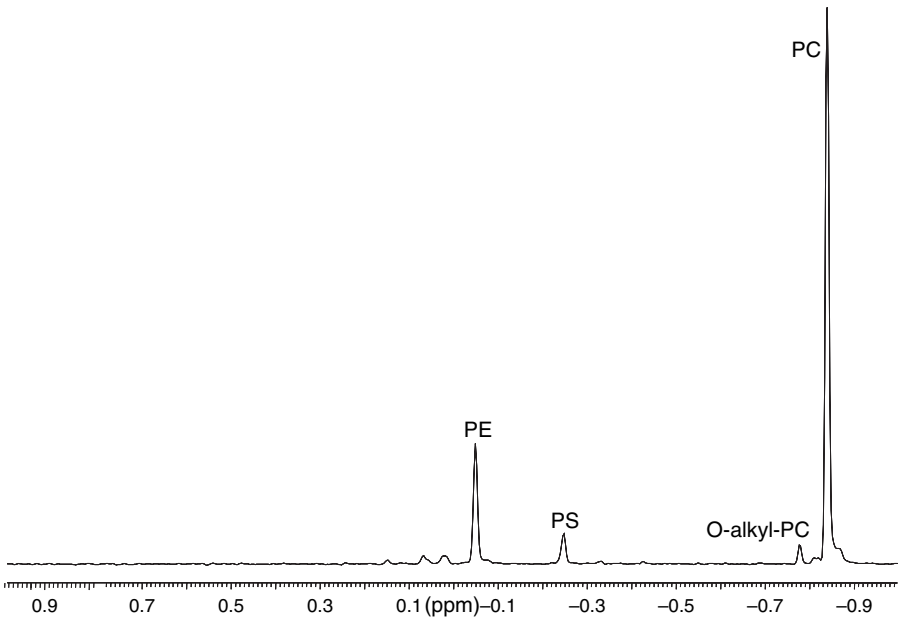

Figure 19 ^{31}P NMR human blood plasma.

5.4. Fats and oils

The determination of fatty acid composition in fats and oils can be done routinely by NMR spectroscopy. ^{1}H NMR enables analysis of double bonds as a sum parameter comparable to the iodine number,[12] a more detailed analysis is useful for the distinction of oils of different origin. So the relative amount of all ω-3 fatty acids can be calculated by using the integrals of the terminal methyl groups (Figure 20). Secondary constituents of plant and animal origin such as cholesterol or other sterols can be quantified using an internal standard,[8] lecithin, tocopherol, sesamol just to mention as well as degradation and oxidation products are other targets of ^{1}H NMR.

The higher spectral dispersion of ^{13}C NMR enables a more detailed look to the glyceride and fatty acid composition of fats and oils, which naturally are complex mixtures.[13] The carbonyl signal region is sensitive for position analysis of the different fatty acid types. Distinction of SN1/3 and SN2 are possible even for phospholipids and glycolipids.[14] The carbonyl signal is sensitive to the double-bond position in the responding fatty acid. A signal separation depending on the double-bond distance up to 11 atoms in the chain is given. This fact allows the individual fatty acid composition especially for lipids of marine origin with its high amounts of polyunsaturated fatty acids. For example, the ^{13}C NMR spectrum of the carbonyl region of a fish oil is shown in Figure 21.

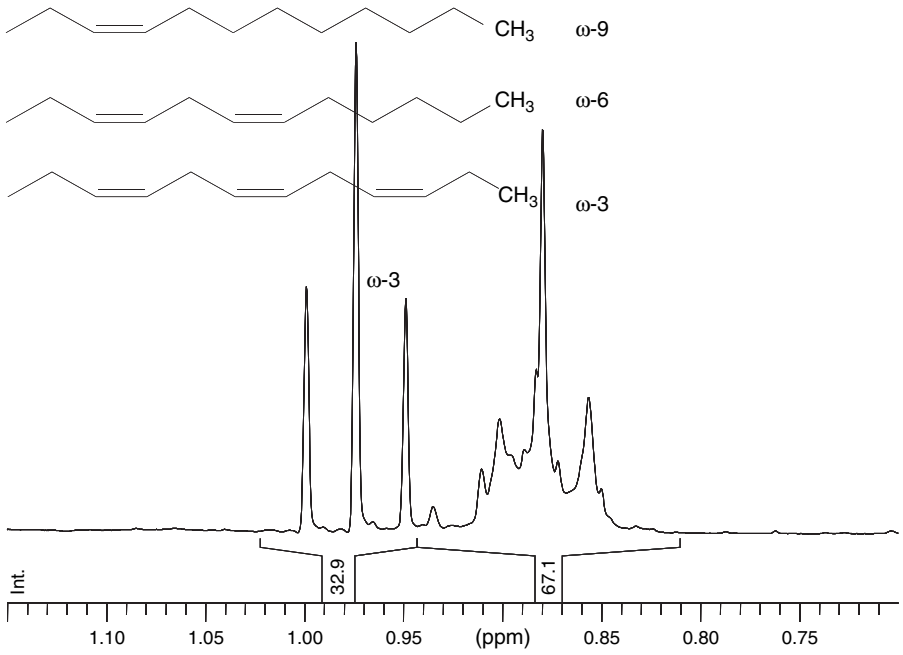

Figure 20 ^{1}H NMR spectrum, methyl region of a vegetable oil.

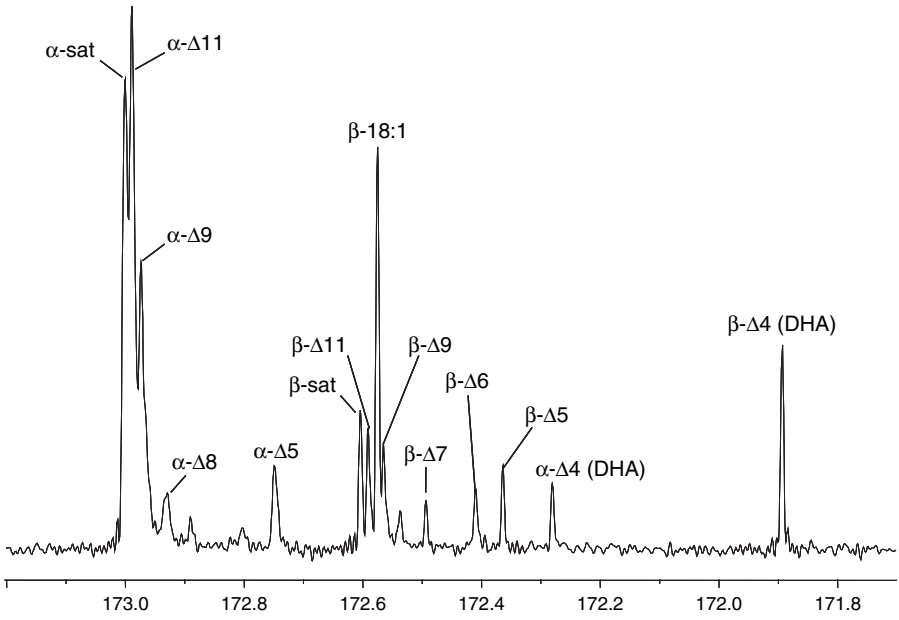

Figure 21 ^{13}C NMR spectrum, carbonyl region of fish oil.

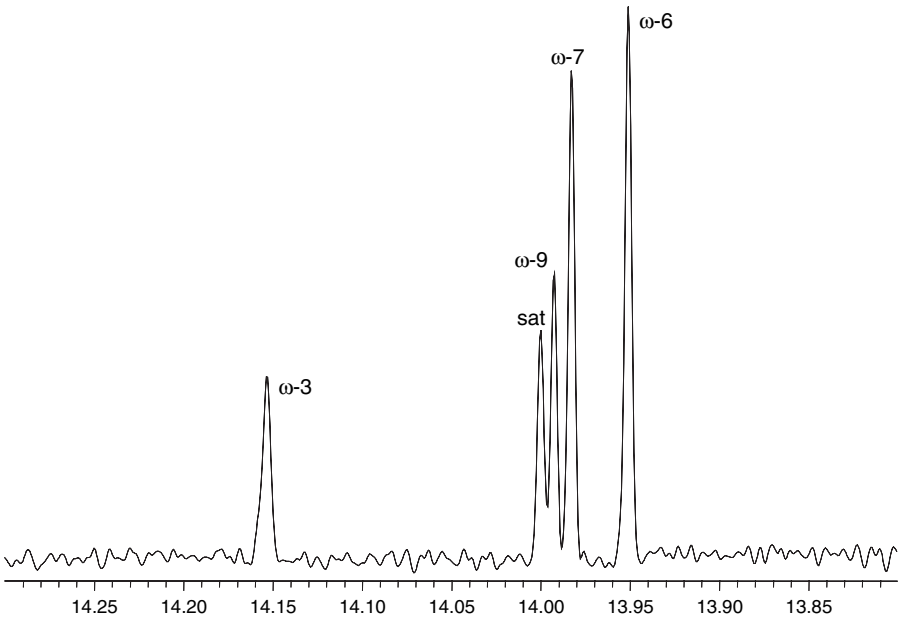

Figure 22 ^{13}C NMR spectrum, methyl region of fatty acid methyl ester mixture.

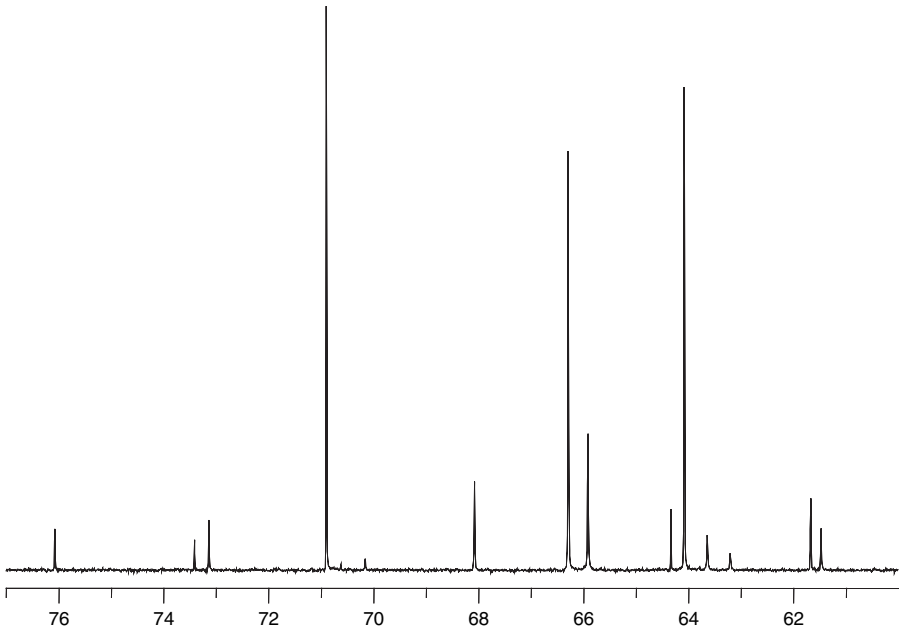

Figure 23 ^{13}C NMR glycerol region of lipids.

The individual position of the fatty acids within the glycerol backbone is reflected in the double-bond signals, too. Using the signals of the methyl groups, which are representing the chain ends, a quantitative analysis of the fatty acid profile concerning the ω-*n* and saturated types is possible (Figure 22).

In this chapter, the glycerol part of a mixture of glycerides is shown in Figure 23. The signal assignment is given in Table 4. The amount of a special glyceride (e.g. 2-acylmonoglucoside) used in an emulsifier mixture is quantified using a defined triglyceride (e.g. trinonadecanoate) as an internal standard. The chemical similarity of the atoms used for quantification avoids the problems normally caused by differences in ^{13}C NMR relaxation time and the nuclear Overhauser effect (NOE).

Table 4 Signal assignment of glyceride C-atoms

	2 M	Gly	1,2 D	1 M	T	1,3 D	1 M
ppm	76.1	73.5	73.1	70.9	70.2	68.1	66.4
	1,3 D	**Gly**	**1 M**	**1,2 D**	**T**	**2 M**	**1,2 D**
ppm	65.9	64.3	64.1	63.6	63.2	61.7	61.5

Gly = glycerol, T = triacylglycerol, M = monoacylglycerol, D = diacylglycerol.

6. Experimental Data

All NMR spectra given in this chapter were performed using a Bruker 300 MHz Avance 1 spectrometer equipped with a BBI probe. Data evaluation was performed using WIN-NMR.

REFERENCES

1. G.F. Pauli, B.U. Jaki, D.C. Lankin, Quantitative ^{1}H-NMR: Developement and Potential of a Method for Natural Products Analysis, J. Nat. Prod. 68 (2005), 133–149.
2. AiF Project, data not published.
3. M.C. Yimdjo, A.G. Azebaze, A.E. Nkengfack, A.M. Meyer, B. Bodo, Z.T. Fomum, Phytochemistry 65 (2004), 2789–2795.
4. Y.C. Shen, M.-C. Hung, L.T. Wang, C.Y. Chen, Inocalophyllins A, B and Their Methyl Esters from the Seeds of *Calophyllum inophyllum* Chem. Pharm. Bull. 51(7)(2003), 802–806.
5. U. Holzgrabe, I. Wawer, B. Diehl, NMR Spectroscopy in Drug Developement and Analysis, Wiley-VCH, Weinheim, (1999) 49–50.
6. P.K. Agrawal, (Ed.), Studies in Organic Chemistry 39, Carbon–13 NMR of Flavonoids, Elsevier, Amsterdam (1989).
7. G.F. Pauli, qNMR – A versatile Concept for the Validation of Natural Product Reference Compounds Phyt. Anal. 12 (2001), 28–42.
8. Spectral Service, internal studies.
9. B.W.K. Diehl, Lipid Analysis of Oils and Fats. Chapter 4: In Multinuclear High Resolution NMR Spectroscopy, edited by Prof. R.J. Hamilton, Blackie Academic & Professional, London, (1997), 87–135, ISBN 07514 0414 4.
10. B.W.K. Diehl, W. Ockels, Quantitative Analysis of Phospholipids. In Proceeding of the 6th International Colloquium Phospholipids. Characterization, Metabolism and Novel Biological Applications, edited by G. Cevc & F. Paltauf, Champaign: AOCS Press (1995), 29–32.
11. Round robin test on ^{31}P-NMR spectroscopy on phospholipids, in Press.
12. K. Manz, Schneider: Bestimmung der Jodzahl in epoxidiertem Sojabohnenöl mit automatisierter ^{1}H-NMR Spektroskopie. GIT Fachzeitschrift für Laboratorien, 3 (1995), 197–199.
13. B.W.K. Diehl, W. Ockels, Fatty Acid Distribution by ^{13}C-NMR-Spectroscopy, Fat Scientific Technol. 97(3) (1995), 115–118.
14. B.W.K. Diehl, H. Herling, E.H. Heinz, I. Riedl, ^{13}C-NMR analysis of the positional distribution of fatty acids in plant glycolipids. Chem. Phys. Lip. 77 (1995) 147–153.