Pure & Appl. Chem., Vol. 48, pp. 105–126. Pergamon Press, 1976. Printed in Great Britain.

SOME RECENT ASPECTS IN THE STRUCTURE ELUCIDATION OF NATURAL PRODUCTS†

WOLFGANG VOELTER
Chemisches Institut der Universität, D-74 Tübingen, Auf der Morgenstelle, Federal Republic of Germany

Abstract—Although in the past 20 yr a large number of new synthetic approaches have been developed methods of isolation and structural elucidation have changed more dramatically in the last two decades than have those of synthesis. For the purification of natural products more and more different chromatographic techniques like paper, thin-layer, column, and gas–liquid chromatography are being used. Characterizations of natural products by elemental analysis, melting points, optical rotation values, and degradation and derivatization reactions are being increasingly supported by parameters of spectroscopic tools like proton nuclear magnetic resonance, carbon nuclear magnetic resonance, phosphorus nuclear magnetic resonance, fluorine nuclear magnetic resonance, infrared spectroscopy, Raman spectroscopy, absorption spectroscopy, optical rotatory dispersion, circular dichroism, magnetic circular dichroism and computer analysis.

The principle of one of the most efficient methods for the structural elucidation of natural products, pulse Fourier transform ^{13}C NMR spectroscopy, is discussed in detail. The usefulness of this method is demonstrated using examples from the peptide, carbohydrate and terpenoid field. The structural elucidations by the application of different physicochemical tools for the three following most recently isolated natural products are given: Hypothalamus-releasing hormones, terpenoids from *Melia azadirachta* Linn., and substances, isolated from urines of sick children.

Natural product's chemistry can be divided into three different fields: methods of isolation, methods of structure elucidation and synthetic approaches.

Though in the last two decades numerous new types of chemical reactions have been found the methods of isolation and structure elucidation have changed more fundamentally than those of synthesis.

For the isolation and purification of organic compounds the different methods of chromatography (paper,[1–3] thin-layer,[4–6] column,[7–9] and gas chromatography,[10–12] have been developed.

One of the most recent and fruitful developments in the field of chromatography, which surely will find great applicability for the separation of natural products, is that of high performance liquid chromatography. Some examples, mainly investigated in our laboratory, are given below.

If a natural product's chemist is convinced that he has isolated a compound of high purity he can begin with its structure elucidation. Only two decades ago this work was mainly based on chemical reactions like those of degradation and derivatization.

Only few physico-chemical parameters were available at that time characterizing organic molecules: melting points, solubilities, values of elemental analysis, molecular weights, or specific rotations. Since then different physicochemical instruments like those of absorption spectroscopy (ABS), optical rotatory dispersion (ORD), circular dichroism (CD), mass spectrometry (MS), and nuclear magnetic resonance (NMR) were developed which allow recording of spectra routinely and thus the basis of correlation between spectral characteristics and structure for organic molecules was laid. The enormous amount of data accumulated in the last years in the ABS, ORD, CD, MS and NMR laboratories will force us in future time to look for ways of storing these parameters on computer files.

In the following the principal fundamental newer developments in some of these fields are discussed and finally some practical examples investigated in our laboratory are given.

2. HIGH PERFORMANCE LIQUID CHROMATOGRAPHY[13,14]

The isolation of a natural product can be considered already as part of its structure elucidation and therefore it is a necessity for a natural product's chemist to know about the theory, the limits and applicability of the different chromatographic methods.

As already mentioned before, high performance liquid chromatography seems to have become the most efficient and suitable method for the separation of complex mixtures of natural products. This idea caused us to develop a system for the separation of carbohydrates respectively amino acids and peptides, as our research group works on the synthesis of peptides and carbohydrates.

(a) *Principle of a high resolution carbohydrate analyzer*

More than 20 yr ago the first column chromatographic separations of carbohydrates were reported. Since then three different principles were applied:

1. adsorption chromatography[15–17]
2. partition chromatography[18–20] and
3. ion exchange chromatography.[21–24]

The efficiency of the latter method prompted us to construct an apparatus (Fig. 1) applying this principle. Borate complexes of carbohydrates specifically interact with anion exchange resins and are therefore separable on columns filled with this material.[26]

A Milton Roy–Dosapro micro pump pumps borate buffer from a gradient-generating system (Grad., Fig. 1) respectively borate buffer reservoirs (A, B, Fig. 1) via a prewashing column on a Biotronik (Frankfurt, Germany) glass-jacketed high pressure liquid chromatographic column, filled with an anion-exchange DA-X 4 resin from Durrum Chemical Corporation, Palo Alto, U.S.A. The DA-X 4 resins are 4% cross-linked polystyrenes of 10–20 μm particle size. To the effluent of the column orcinol sulfuric acid reagent is added and the mixture passes a 20 m × 0.7 mm Teflon coil which is placed in a heated (98–100°C) water bath. Under these conditions

†Communication 120 about structures of natural products and organic molecules.

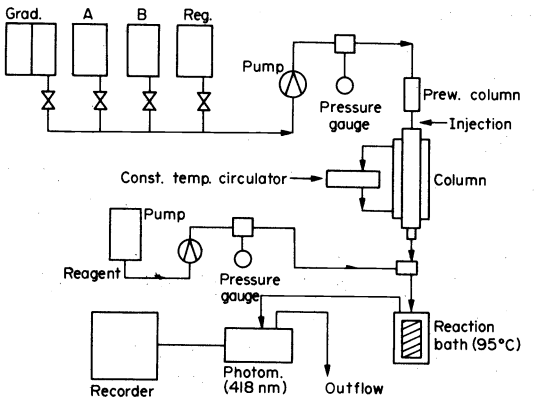
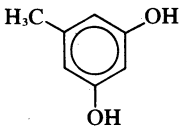
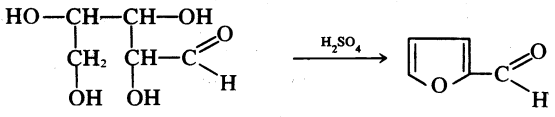


Fig. 1. Flow scheme of a carbohydrate analyzer. Grad.: Gradient generator; A, B: buffer reservoirs; Reg.: regenerating buffer reservoir.[25] Borate complexes of carbohydrates specifically interact with anion exchange resins and are therefore separable on columns filled with this material.[26]

sugars from furan derivatives which react with orcinol



give a red-yellow dye which has an absorption maximum at $\lambda = 420$ nm:



(Pentose) (Furfural)

A Biotronik BT 6620 spectrophotometer with an interference filter at 420 nm is used for photometric detection (path length of the optical cell: 1 cm; extinction ranges 0.1, 0.2, 0.5, 1.0 and 2.0 AU).

A typical sugar chromatogram of a 16 component standard using a DA-X 4 F resin (11 ± 1 μm particle size) is shown in Fig. 2.

The peaks in Fig. 2 were assigned by single component runs.

The sugar separation system will find valuable application for the investigation of body fluids as is demonstrated by Fig. 3.

(b) *High pressure liquid chromatographic analysis of amino acids and peptides*

The synthesis of hypothalamus peptide hormones makes the separation of amino acid hydrolysates and peptides a day to day task in our laboratory.[27] As most of our synthetic compounds are tested biologically the highest achievable purity of the peptides is a necessity for us. It also becomes obvious that one of the more recently developed synthetic methods, the solid phase peptide synthesis[28], will be an alternative to the classical methods only if the desired reaction product is separable from the products of failed and truncated sequences.

Products of failed sequences occur in minor quantities only in the reaction product of a solid phase peptide synthesis. We were therefore forced to develop a more sensitive detection system for amino acids and peptides than that of Spackman, Stein and Moore.[29]

This newly developed saparation system finds also increasing application in our group in the investigation of biological fluids.

For our separations we modified a high pressure liquid chromatograph (Hewlett-Packard, model 1010 B),[30,31] equipped with an automatic gradient-generating system

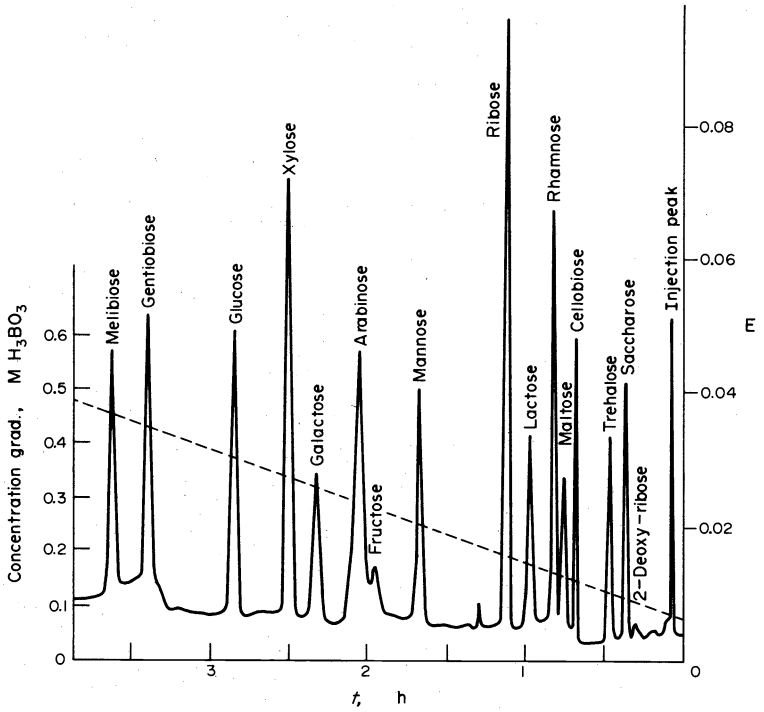


Fig. 2. Sugar chromatogram of a 16 component standard. Gradient: 0.075–0.6 M H_3BO_3/pH 8.0–10.5; column: 0.6 × 19 cm; resin: DA-X 4F; column temperature: 60°C; flow rate: 1 ml/min; each peak represents 16 n Mol/monosaccharide and 8 nMol/disaccharide (≙2.14–2.88 μg per component).[25]

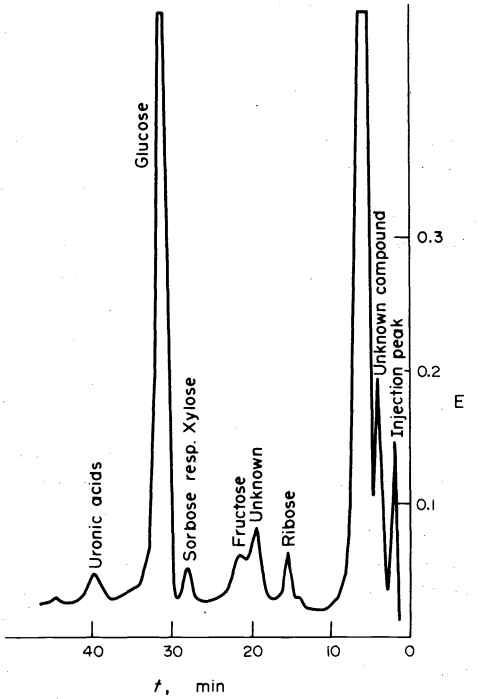


Fig. 3. Human cerebrospinal fluid carbohydrate chromatogram. Sample: 100 µl. of ultrafiltrated human cerebrospinal fluid; one buffer elution: 0.4 M H_3BO_3, pH 9.0; collumn: 0.4× cm; resin: DA-X 4F, 11 ± 1 µm; column temperature: 60°C; flow rate: 0.7 ml/min.[25]

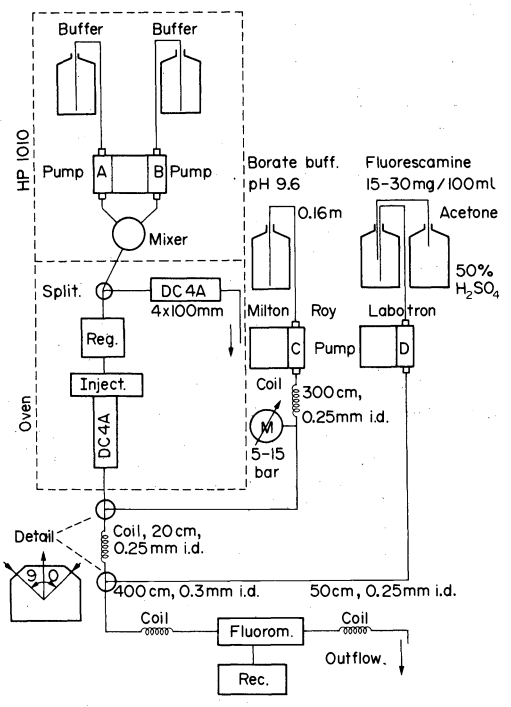


Fig. 4. Functional diagram of the automatic amino acid analyzer.[31]

and use ion-exchange resins of the particle size 8 ± 2 µm and fluorescamine[32-34] (4-phenylspiro[furan-2(3H), 1'-phthalan]-3,3'-dione). Fluorescamine is a non-fluorescent compound which reacts in milliseconds with primary amines at room temperature and pH 9–10 to a fluorimetrically detectable fluorophor. Excess reagent is hydrolyzed to non-fluorescent compounds:

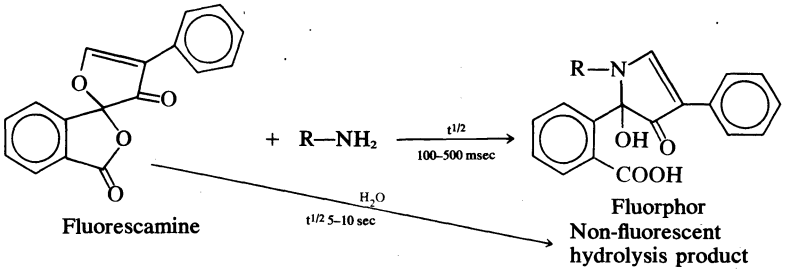


Figure 4 shows the functional diagram of the amino acid analyzer.

The gradient is generated with two citrate buffer solutions of pH 2.6 (0.2 N in sodium) and 6.3 (1.2 N in sodium). The gradient-generating system of the HP 1010 B apparatus allows to achieve constant flow rates of 0.05 ml/min only if the buffer mixture is pumped to a splitting system which is connected with a stainless steel separation column (250 × 3 mm) and a reference column (100 × 4 mm). Both columns are filled with strongly acid 4% cross-linked polystyrene divinylbenzene cation-exchange resin of 8 ± 2 µm particle size (DC-4 A, Durrum, Palo Alto, U.S.A.). To ensure rapid reaction with fluorescamine 0.16 M borate buffer of pH 9.6 is pumped (Dosapro micro-pump, Milton Roy, Philadelphia, U.S.A.) via a coil of 3 m length and 0.25 mm internal diameter to the column effluent. The fluorescamine solution (0.01–0.025%) in absolute acetone is then admitted by a Labotron pump to the now alkaline column effluent. The mixture of column effluent, borate buffer and fluorescamine acetone solution passes a reaction coil of 4 m length and 0.3 mm internal diameter and is detected with a Hewlett–Packard 1033 A fluorimeter equipped with a 10 µl. flow-through cell.

Figure 5 shows a typical chromatogram of a 16 component amino acid mixture with diagrams of the buffer flow rates through pump A and B. The overall flow rate starts with 0.75 and ends at 1.25 ml/min.

To test our system and compare it with conventional apparatus a solid phase synthesized hexapeptide amide with natural sequence parts of physalaemin (H-Asn-Lys-Phe-Tyr-Gly-Met-NH_2) was hydrolyzed with 6 N HCl and analyzed with a Beckman Unichrom amino acid analyzer (using ninhydrin as detection reagent) and our developed analyzer. Figures 6a and b show the result.

The comparison of the two chromatograms shows that, compared to conventional methods, high pressure liquid chromatography in combination with the application of fluorescamine and ion-exchange resins of small particle size improves the sensitivity by a factor of 100–1000 and shortens the separation time by a factor of 3–4.

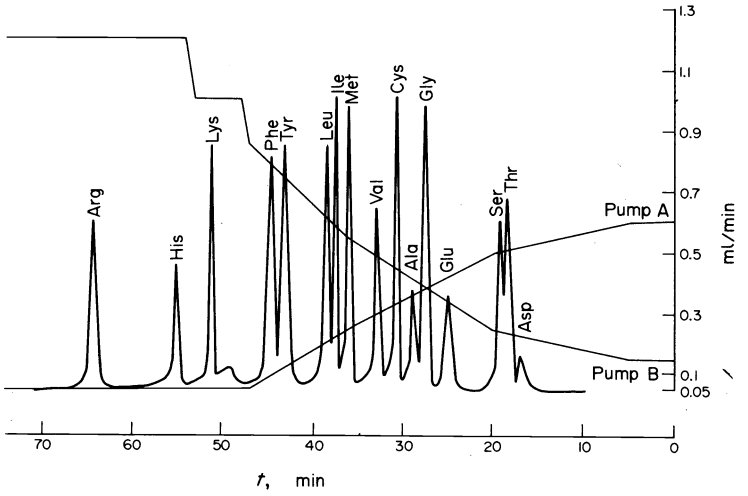


Fig. 5. High pressure liquid chromatogram of a 16 component amino acid mixture with flow-rate diagrams of pumps A and B. Experimental conditions see text.[31]

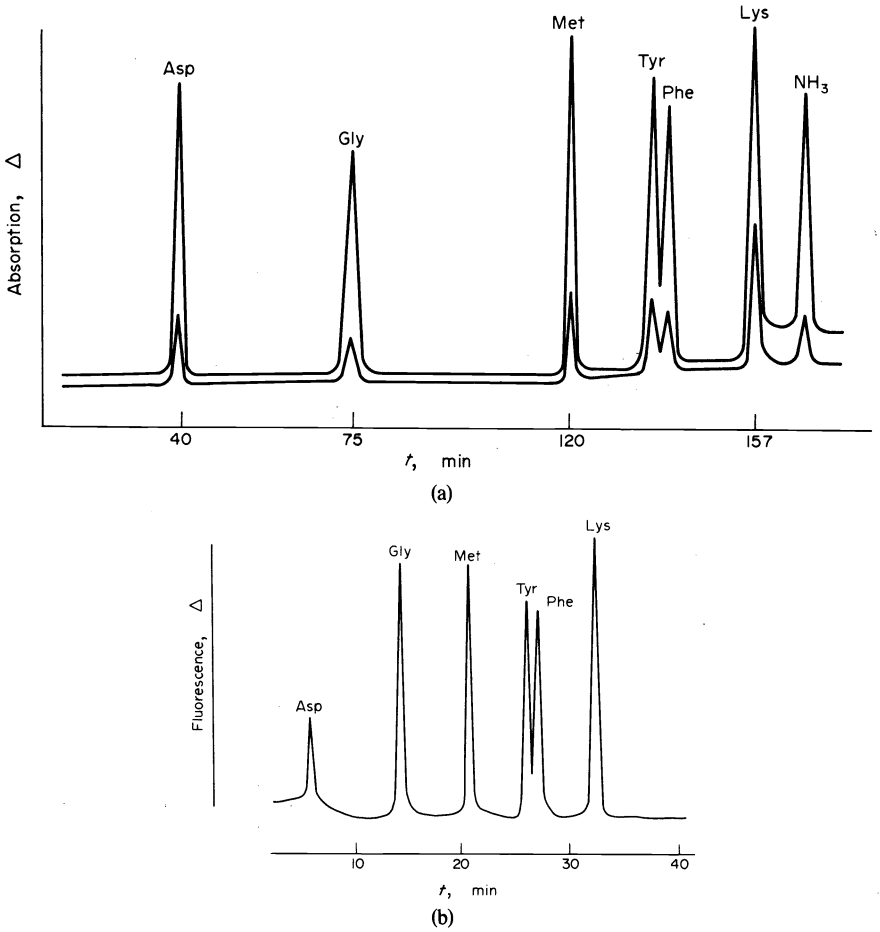


Fig. 6. Chromatograms of a hydrolyzed hexapeptide amide with natural sequence parts of physalaemin. (A) Injection of 100 nm of the hydrolyzed hexapeptide amide in a Beckman Unichrom amino acid analyzer using ninhydrin for detection. (B) Injection of 500 pm of the same solution as in A into the high pressure liquid chromatograph using fluorescamine for detection.

3. MASS SPECTROMETRY

Without any doubt mass spectrometry has become one of the most valuable and efficient tools in the structure elucidation of natural products and an enormous amount of work has been published in the past which is surveyed in a series of excellent monographs.[35-39] Mass spectra can be recorded from very low sample quantities (10^{-6}–10^{-10} g) and a great deal of structural information is received from them.[35-39]

Sufficient volatility of a compound, however, is neces-

sary if the most common method, the electron bombardment mass spectrometry, is applied. Polar functional groups and high molecular weights of natural products often prevent recording their mass spectra with this method. The volatility, however, can be often enlarged by simple chemical modifications of the polar groups like methylation, trimethylsilylation or trifluoroacetylation.

In connection with our work on hypothalamus-releasing hormones[40] trimethylsilyl derivatives of the thyrotropin-releasing hormone (TRH) and its analogs proved to be very suitable for mass spectroscopic investigation.[41]

The interpretation of the most important ions is given in Table 1.

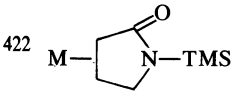
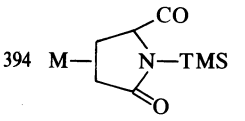
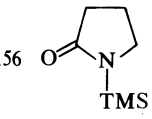
Non-derivatized tripeptides are in most cases too low volatile and therefore not suitable for mass spectroscopic investigation by electron bombardment. With increasing chain length even derivatized peptides decompose before vaporization and no fragmentation pattern of the molecular ion is obtainable after electron bombardment. The polar functional groups of intermediates in peptide synthesis often bear protecting residues of large molecular weights and therefore no volatile derivatives can be synthesized from them for mass spectroscopic investigation.

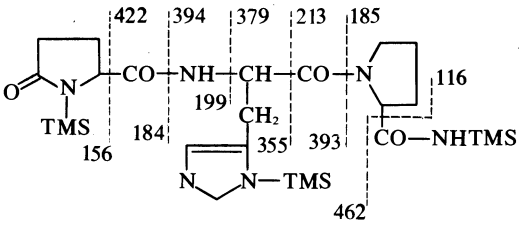
As an example the synthesis of a TRH derivative which has been shown to have low thyrotropin-releasing but strong antidepressive activity is given in Fig. 8.

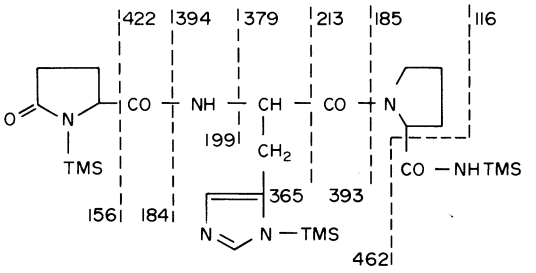
From the intermediate benzyloxycarbonyl - N^{γ} - 4,4' - dimethoxybenzhydryl - glutamyl - phenylalanyl - prolinamide of the Phe^2-TRH synthesis, shown in Fig. 8, for instance no volatile derivative can be made; besides its high molecular weight prevents its vaporization so that no molecular ion is found upon electron bombardment.

In contrast to conventional mass spectrometry a more recently developed method, field desorption mass spectrometry,[43] allows study of polar molecules of much higher molecular weight and is therefore most attractive

Table 1. Interpretation of the most important ions of the trimethylsilyl derivative of TRH upon electron impact

m/e	Interpretation
578	M^+(3 TMS)
563	M–CH_3
560	M–H_2O
545	M–(CH_3 + H_2O)
506	M^+(2 TMS)
488	506–H_2O
463	T–TMS–N=C=O
422	
394	
365	393–CO
156	





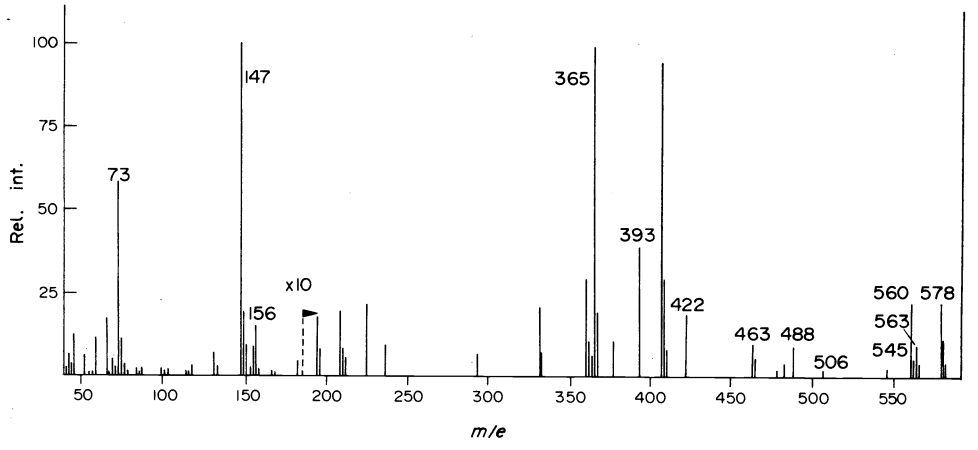


Fig. 7. Mass spectrum of the trimethylsilyl derivative of TRH (ionizing energy: 70 eV; accelerating voltage: 3.5 kV).[41]

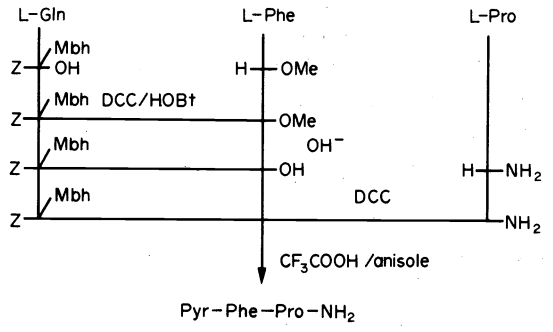


Fig. 8. Synthesis of pyroglutamyl-phenylalanyl-prolinamide. (L-Gln; L-glutamine; L-Phe: L-phenylalanine; L-Pro: L-proline; Z: benzyloxycarbonyl; Mbh: dimethoxybenzhydryl; Me: methyl; DCC: dicyclohexylcarbodiimide; Pyr: pyroglutamic acid).[42]

for natural product chemistry. The next two figures (Figs. 9 and 10) show two examples which could not be investigated by conventional electron bombardment but by field desorption mass spectrometry.

4. NUCLEAR MAGNETIC RESONANCE

Without any doubt besides mass spectrometry nuclear magnetic resonance spectroscopy gives the most valuable information about the structure of a natural product.

The Larmor frequency of a nucleus with a magnetic moment depends characteristically on the chemical environment of the nucleus and is expressed in terms of its chemical shift δ:

$$\delta = \frac{\Delta\nu \cdot 10^6}{\nu}$$

$\Delta\nu$[Hz]: frequency difference between standard signal and compound signal

ν[MHz = 10^6 Hz]: frequency used.

Only nuclei having a magnetic moment are investigated by nuclear magnetic resonance spectroscopy. Very important isotopes of organic chemistry, like ^{16}O or ^{12}C have no magnetic moment and therefore cannot be investigated with nuclear magnetic resonance spectroscopy. ^{1}H, ^{13}C, ^{14}N and ^{31}P are therefore the most important nuclei in natural product's chemistry investigated by NMR.[51-54] Most knowledge is available in the field of proton nuclear magnetic resonance; a series of monographs covers this area.[45-50]

One of the most recent and most exciting developments in the field of NMR is that of the pulse Fourier transform ^{13}C NMR spectroscopy which allows routine measurement of carbon-13 spectra.[51-54]

Nuclei possessing an angular momentum may undergo nuclear magnetic resonance in a static magnetic field H_0. P_0, the component of p in the direction of H_0 is related to the spin quantum number (m):

$$p_0 = \frac{m \cdot h}{2\pi}; \quad m = \pm n\,1/2; \quad 0, 1, 2 \qquad (1)$$

$m = I, I-1, \ldots\ldots\ldots, -I.$

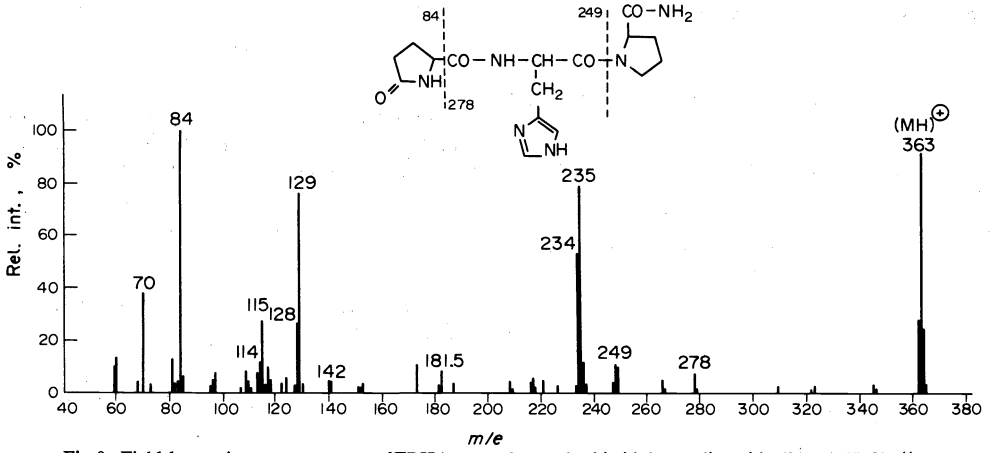


Fig. 9. Field desorption mass spectrum of TRH (L-pyroglutamyl-L-histidyl-L-prolinamide, (21 mA, H_2O)).[44]

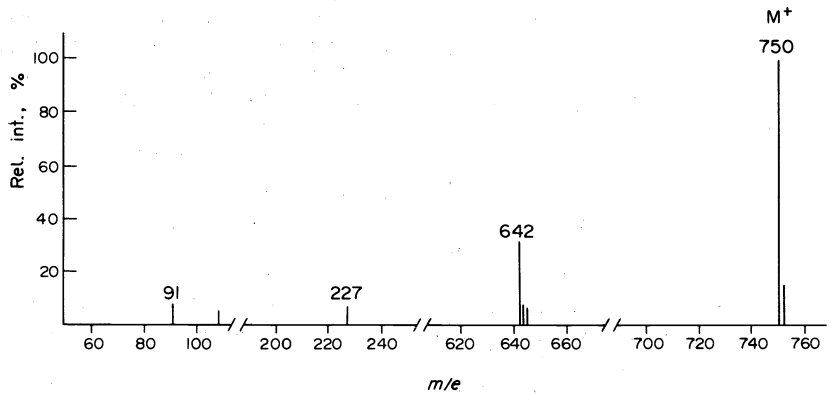


Fig. 10. Field desorption mass spectrum of Z-Gln(Mbh)-Phe-Pro-NH_2 (17 mA).[44]

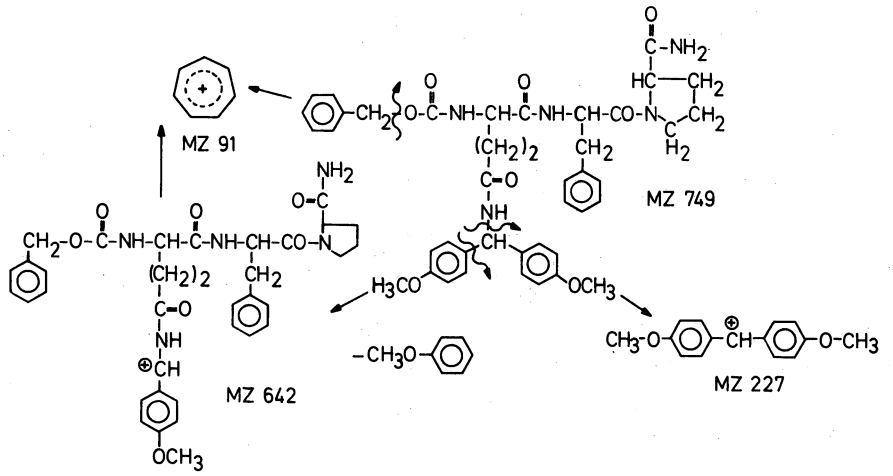


Fig. 11. Fragmentation pattern of the field desorption mass spectrum of Z-Gln(Mbh)-Phe-Pro-NH_2.[44]

(I: total spin quantum number)
Nuclei with a magnetic moment μ interact with magnetic fields:

$$\mu = \gamma \cdot p \tag{2}$$

(γ: gyromagnetic ratio)
From eqns 1 and 2 follows:

$$\mu \cos\theta = \mu_0 = \gamma I \frac{h}{2\pi} \tag{3}$$

(μ_0: component of μ in the direction of H_0)
The energy E of a magnetic moment μ in a field H_0 follows from eqn 4:

$$E = -\mu_0 H_0 = -\gamma \frac{h}{2\pi} I H_0. \tag{4}$$

In a magnetic field for a nucleus with the total spin quantum number I, $(2I+1)$ energy levels are possible. For the important nuclei of natural products ^1H, ^{13}C and ^{31}P I is 1/2 and therefore two energy levels, symbolized with $E_{+1/2}$ and $E_{-1/2}$ are possible

$$E_{+1/2} = -\mu_0 H_0 = -\gamma \frac{h}{4\pi} H_0 \tag{5}$$

$$E_{-1/2} = \mu_0 H_0 = \gamma \frac{h}{4\pi} H_0 \tag{6}$$

$$\Delta E = E_{-1/2} - E_{+1/2} = 2\mu_0 H_0 = \frac{\gamma h}{2\pi} H_0. \tag{7}$$

As

$$E = h \cdot \nu_0, \tag{8}$$

the Larmor precession frequency ν_0 for nuclei with $I = 1/2$ follows for eqn 9:

$$\nu_0 = \frac{\gamma}{2\pi} H_0. \tag{9}$$

For a field strength of 21.3 kG the Larmor frequency for ^{13}C nuclei is in the range between 22–23 MHz.

To observe NMR signals a sample is placed in a static magnetic field H_0 and is irradiated with an alternating field H_1 with the radio-frequency ν_1 (frequency sweep). If ν_1 becomes equal to the Larmor frequency of the sample nuclei an absorption signal is registered by the recorder.

For nuclei with $I = 1/2$ the distribution of the spin populations (N_+/N_-) may be expressed:

$$\frac{N_+}{N_-} = e^{-(E/kt)} \approx 1 + \frac{\Delta E}{kT} = 1 + \frac{2\mu_0 H_0}{kT}. \tag{10}$$

After resonance the nuclear spins relax from the upper level to the equilibrium with their surroundings ('lattice') in a time called spin-lattice relaxation time T_1(10^{-4}–10^4 sec) and according to the Heisenberg equation

$$\Delta E \cdot \Delta t = h \cdot \Delta\nu_{1/2} \cdot T_1 \gtrsim h \tag{11}$$

$$\Delta\nu_{1/2} \gtrsim \frac{1}{T_1}. \tag{12}$$

The life-times of spin states may be further shortened by dipole–dipole interactions and exchange of energy quanta ($\Delta E = 2\mu_0 H_0$; spin–spin relaxation time T_2):

$$\Delta\nu_{1/2} = \text{const.} \frac{1}{T_2}. \tag{13}$$

The contribution of the field inhomogeneity to $\Delta\nu_{1/2}$ follows from eqn 14:

$$\Delta\nu_{1/2} = \frac{\gamma \Delta H_0}{2\pi} = \text{const.} \frac{\gamma \Delta H_0}{2}. \tag{14}$$

If a sample of identical nuclei with $I = (1/2)$ is brought into a magnetic field H_0 a net macroscopic magnetization M_0 along the z axis arises. If perpendicular to H_0 a radiofrequency field with a proper frequency ν_1 is applied the resultant magnetization vector M is now composed of three components along the axes x, y and z:

$$M = m_x i + M_y j + M_z k \tag{15}$$

$M_z k$: longitudinal magnetization
$M_x i$, $M_y j$: transverse magnetization.

The magnetization vector components are related with the longitudinal relaxation time T_1 and the spin–spin relaxation time T_2 by the following equations:

$$\frac{dM_z}{dt} = -\frac{M_z - M_0}{T_1} \tag{16}$$

$$\frac{dM_x}{dt} = -\frac{M_x}{T_2} \tag{17}$$

$$\frac{dM_y}{dt} = -\frac{M_y}{T_2}. \tag{18}$$

After excitation the transverse magnetization M_{xy} decays exponentially with the transverse magnetization time T_2 (free induction decay). Immediately after the resonance the interference of H_1 and M_{Xy} is a beat decreasing exponentially with T_2. These beat interference signals are also obtained irradiating with short radiofrequency pulses ΔH_1 (Fig. 12).

If the radiofrequency pulse covers all the ^{13}C Larmor frequencies of the irradiated probe the result is a pulse interferogram which contains all the information of a frequency or field sweep NMR spectrum (Fig. 13) and, like any complex wave function, can be converted by a Fourier transformation[55] into a spectrum of Larmor frequencies.

Pulse Fourier NMR spectrometry is the most powerful method up to date; increasing the sensitivity of an NMR

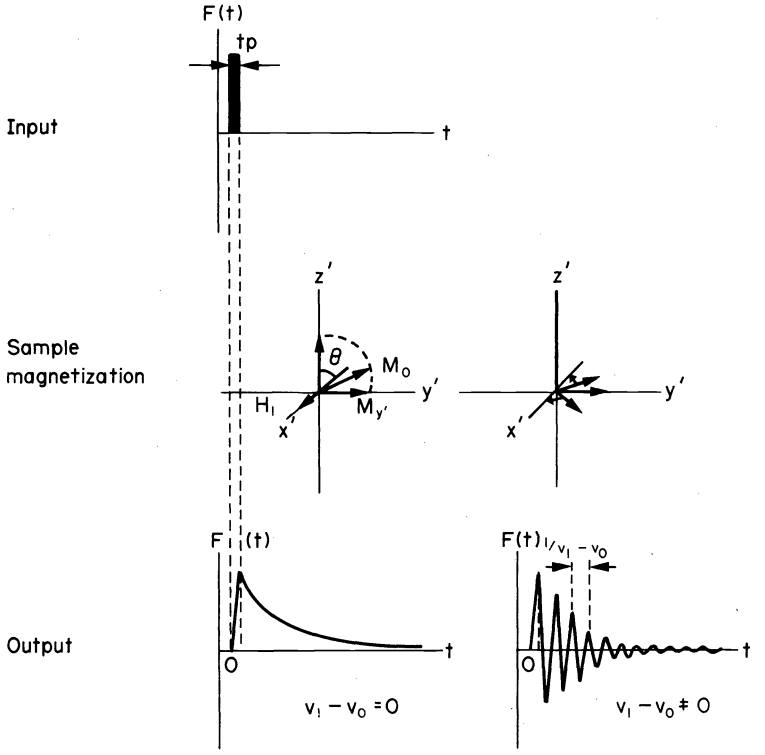


Fig. 12. Free induction decay. Input: Radiofrequency pulse as input signal; sample magnetization: During the radiofrequency pulse (left), free induction decay following the radiofrequency pulse (right) output: Output signal at resonance (left), output signal for off-resonance (right).[54]

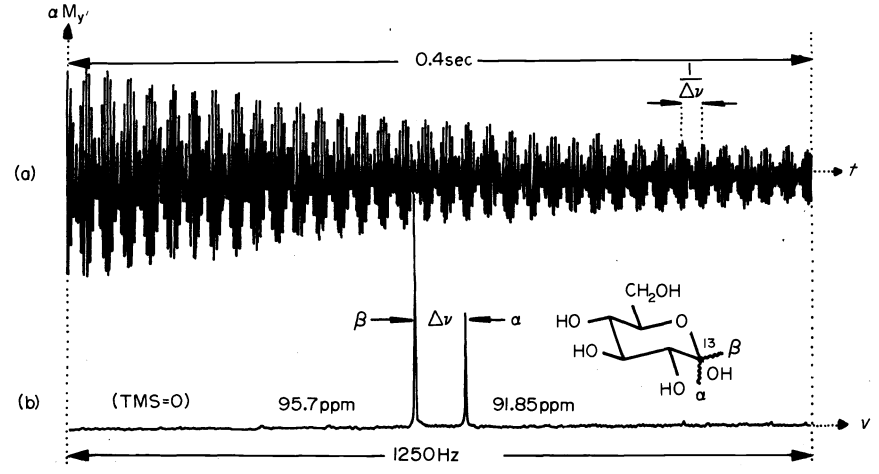


Fig. 13(a). Pulse interferogram of mutarotated 1-^{13}C-glucose (60% ^{13}C). 22.63 MHz, 50 mg/ml D_2O, proton decoupled, temperature: 30°C, accumulation of 32 scans, pulse width: 12 μsec, pulse interval: 0.4 sec. (b). Fourier transform of 13a.[54]

spectrometer for two reasons:

(1) A sufficiently large radiofrequency pulse is practically equivalent to an irradiation from a multichannel transmitter and the signal (S): noise (N) ratio increases with the number of channels applied:[56]

$$\frac{S}{N} \sim \sqrt{m}. \tag{19}$$

(2) If the pulse interferograms are accumulated by an averaging computer the $S:N$ ratio improves with increasing number (n) of accumulations:

$$\frac{S}{N} \sim \sqrt{n}. \tag{20}$$

The ^{13}C–^{1}H couplings are in the range of 100–250 Hz. Overlapping ^{13}C–^{1}H multiplets make the interpretation of complex ^{13}C NMR spectra difficult. However, if the ^{13}C sample is stimulated by the ^{13}C radiofrequency pulse and additionally irradiated with a field covering all the Larmor frequencies of the protons the ^{13}C–^{1}H multiplets collapse to singlets. Proton broad band decoupling besides increases the $S:N$ ratio by the nuclear Overhauser effect[57] which is based on changes of the natural population of the ^{1}H spin levels. This effect is demonstrated by Fig. 14 for the ^{13}C NMR spectra of ^{13}C-enriched formic acid.[58]

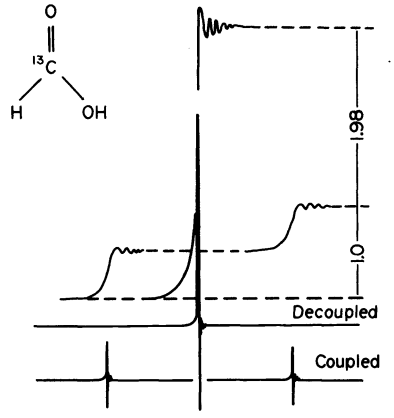


Fig. 14. Coupled and proton broad band decoupled ^{13}C NMR spectrum of formic acid for the demonstration of the nuclear Overhauser effect.[58]

Though proton broad-band decoupling reduces the complexity of ^{13}C NMR spectra a lot of information is lost which could be deduced from ^{13}C–^{1}H couplings.

A valuable aid in the signal assignment of ^{13}C spectra is received from proton off-resonance recording: Besides the radiofrequency pulse, exciting the ^{13}C nuclei, the probe is irradiated with a second frequency which is several hundred Hz apart from the Larmor frequency of the protons. Under these conditions vicinal and long-range couplings collapse and a spectrum of multiplets of first order is observed.

(a) ^{13}C *NMR spectra of terpenes and steroids*[51,52,54]

The ^{13}C chemical shifts of organic molecules are spread over a range of 200 ppm. With increasing number of hydrogen atoms attached to carbon atoms of hydrocarbons the ^{13}C signal shifts generally to higher field. Electron-withdrawing functional groups or heteroatoms cause downfield shifts especially on the resonances of neighbouring carbon atoms. ^{13}C chemical shift ranges of the main types of carbon nuclei are surveyed in Fig. 15.

As a demonstrating example for a ^{13}C NMR spectrum of higher molecular weight terpenes and steroids that of nimbin is given (Fig. 16).

The signal assignments are done on the following basis:

(1) application of general chemical shift rules (see Fig. 15),

(2) determination of the number of hydrogen atoms attached to each carbon atom by proton off-resonance spectroscopy,

(3) spectral comparison with similar and partial structures.

(b) ^{13}C *NMR spectra of carbohydrates and nucleosides*[51,52,54]

Polyols are reduction products of carbohydrates and often used for identification of the parent compound.

The ^{13}C NMR spectra of polyols are much more easy to interpret than those of their corresponding sugars because they can not undergo mutarotation.

Figure 17 shows the ^{13}C NMR spectrum of ribitol.

At highest field of Fig. 17 the signal of the primary carbons is found, at lower field the resonances of the secondary carbons 2, 3, 4 are expected. The carbon atoms 2 and 4 are magnetically equivalent and must cause the signal at 75.3 ppm with almost double intensity compared to the signal at somewhat lower field. Therefore the resonance at lowest field is assigned to C-3.

The following ^{13}C chemical shift (δ, ppm, relative to

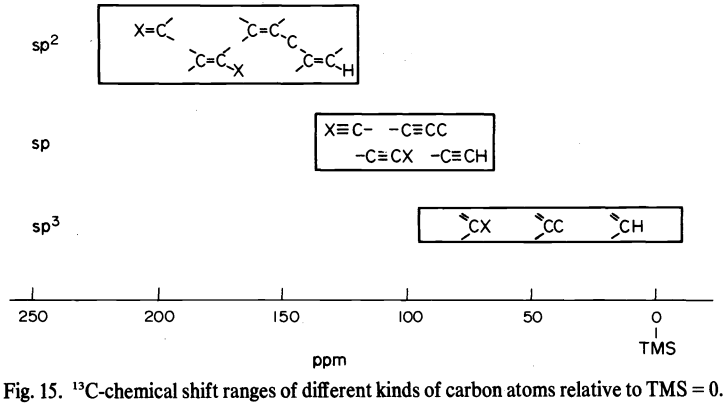


Fig. 15. ^{13}C-chemical shift ranges of different kinds of carbon atoms relative to TMS = 0.

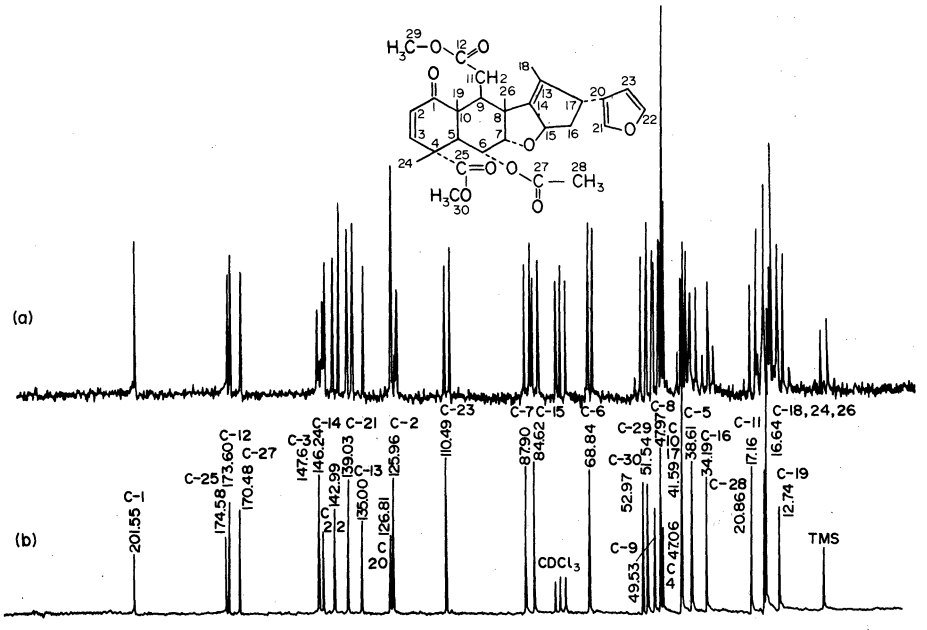


Fig. 16. 22.63 MHz PFT ^{13}C {^1H} NMR spectra of nimbin, 350 mg/1.5 ml $CDCl_3$, temperature: 30°C, pulse width: 10 μ sec, pulse interval: 0.4 sec/4k interferogram; phase corrected. (a) proton off-resonance decoupled; 9550 accumulated interferograms; (b) proton broad band decoupled; 1600 accumulated interferograms.

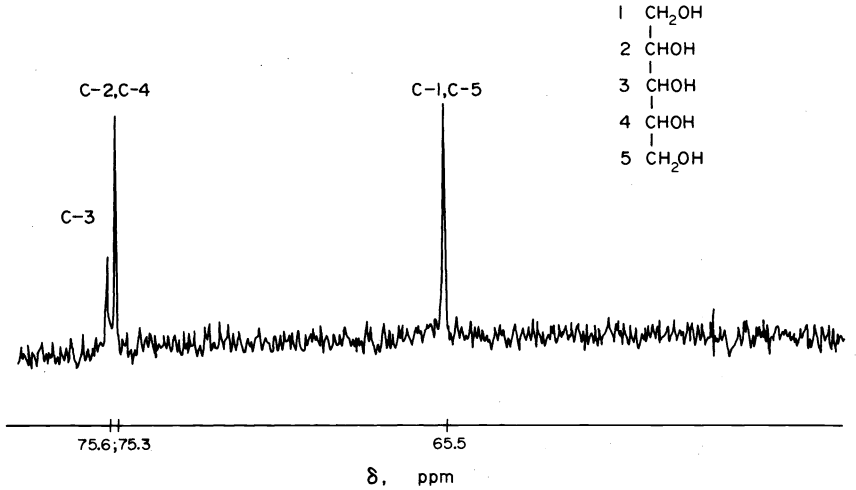


Fig. 17. 22.63 MHz PFT $^{13}C\{^1H\}$ NMR spectrum of ribitol, 20% in D_2O, temperature: 30°C, pulse width: 10 μsec, pulse interval: 0.4 sec/4 K interferogram, ppm values are given relative to TMS = 0.

TMS=0) rules can be applied for the signal assignments of carbohydrates:

(a) The anomeric carbon atom is the only one which is attached to two electron withdrawing oxygen atoms and therefore resonates at lowest field (90–98 ppm).

(b) The signals of CH_2-groups (CH_2–OH residues of hexoses or ring CH_2-groups of pentoses) generally occur at highest field in the ^{13}C NMR spectrum of a free monosaccharide.

(c) The ppm range of ring carbons bearing hydroxyl groups is in the range of 65–76 ppm.

(d) Generally in pyranose molecules equatorial hydroxyl groups cause, compared to axial ones, an electron withdrawing effect on the attached carbon atom. The ^{13}C shifts of the anomeric carbon atoms of the pair α- and β-D-mannose are exceptions for this rule and the exceptional behavior can be explained on the basis of the Reeves effect.

(e) 1.3-Diaxial interaction causes the resonance of a carbon atom in γ position of an axial substituent to resonate at higher field.

(f) The signals of carbon atoms of anomers can easily be assigned comparing the ^{13}C spectra of a carbohydrate solution before and after mutarotational equilibrium is achieved.

The above made statements (a–f) are demonstrated by Fig. 18 showing the ^{13}C NMR spectra of D-glucose before and after mutarotational equilibrium has adjusted.

As an example for an enzyme cofactor and a nucleoside the ^{13}C NMR spectrum of riboflavin is discussed (Fig. 19).

Application of general chemical shift rules, spectral comparison with similar compounds and proton off-

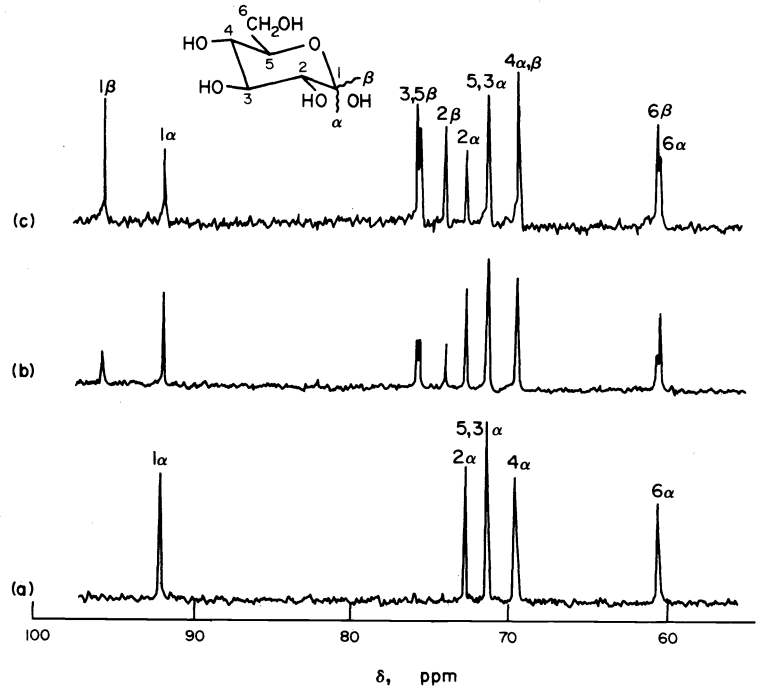


Fig. 18. PFT ^{13}C {^1H} NMR spectrum of D-glucose, 22.63 MHz, 1 M in D_2O, temperature: 30°C, pulse width: 5 μsec, pulse interval: 0.8 sec, 2500 Hz, the numbers indicate the numbering of the C-atoms (δ-values relative to TMS = 0); (a) recorded immediately after dissolving the sugar (accumulation of 512 pulse interferograms); (b) the same solution measured after 2 hr (accumulation of 1024 pulse interferograms); (c) the same solution measured after 8 hr (accumulation of 1024 pulse interferograms).[54]

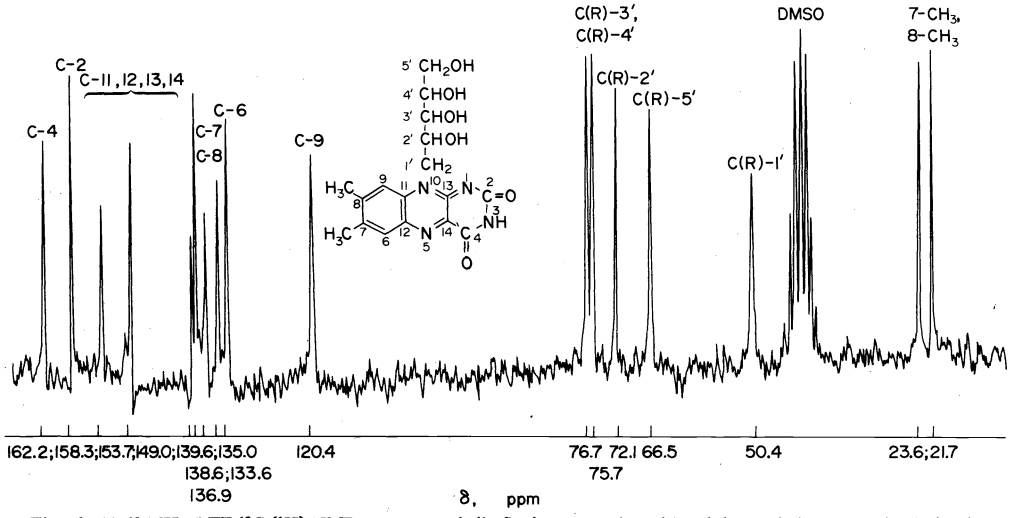


Fig. 19. 22.63 MHz PFT ^{13}C {^1H} NMR spectrum of riboflavin, proton broad band decoupled, saturated solution in DMSO-D_6, temperature: 30°C, pulse width: 10 μsec, pulse interval: 0.4 sec/4 K interferogram; ppm values relative to TMS = 0.

resonance spectroscopy leads to the signal assignment given in Fig. 19.

(c) ^{13}C *NMR spectra of amino acids and peptides*[51,52,54]

The different carbon atoms of amino acids resonate in the following ppm ranges (see opposite).

Deprotonation of NH_3^+, SH or COOH groups usually shifts the resonance of neighbouring carbon atoms downfield.

^{13}C NMR spectroscopy is also used more and more for structural proofs of amino acid and peptide derivatives which are intermediates in peptide synthesis.

Carbon atom	ppm range (relative to TMS = 0)
Carboxyls	168–183
α-carbons	40–65
β-carbons	17–70
γ and δ carbons	17–50
Aromatic and heteroaromatic Carbons	110–140

Figure 20 shows the ^{13}C NMR spectrum of

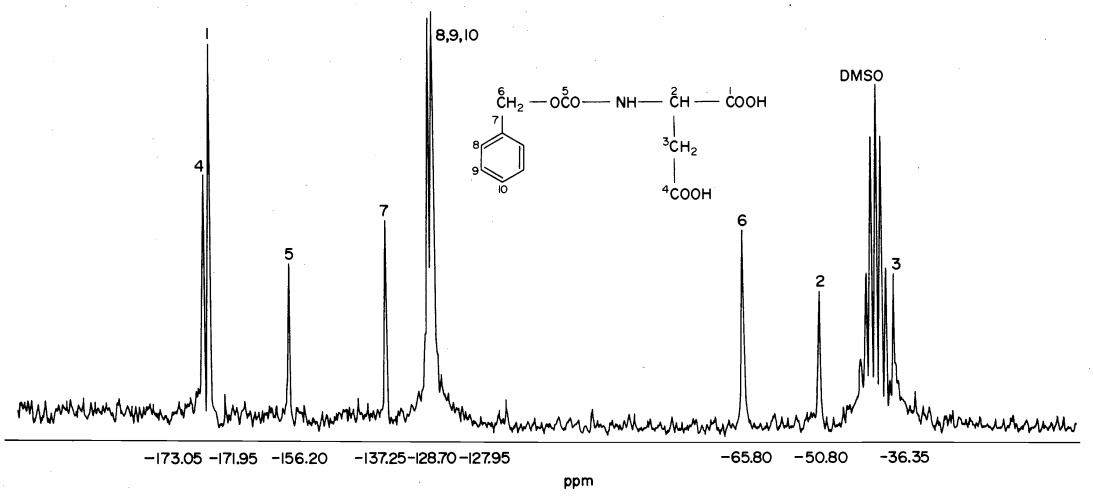


Fig. 20. 22.63 MHz PFT ^{13}C {^1H} NMR spectrum of benzyloxycarbonyl-L-aspartic acid, 250 mg/ml DMSO-D_6, temperature: 30°C, pulse width: 5 μsec, accumulation of 4096 pulse interferograms, ppm values relative to TMS = 0.

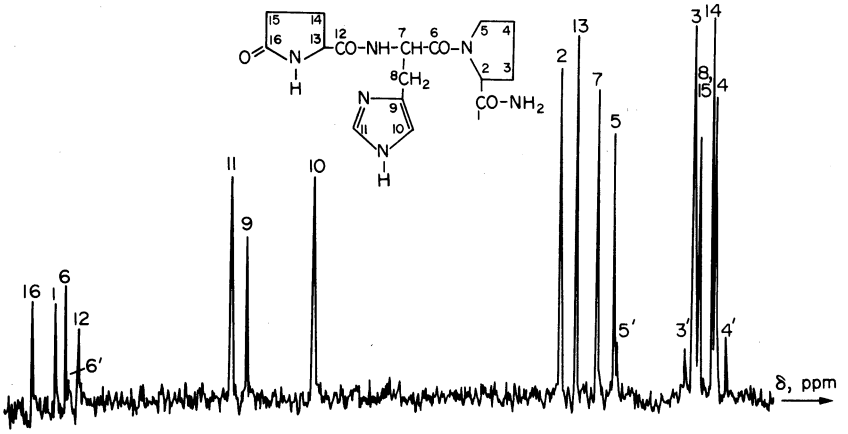


Fig. 21. 22.63 MHz PFT ^{13}C {^1H} NMR spectrum of TRH, 200 mg/1.5 ml D_2O, temperature: 30°C, accumulation of 16384 pulse interferograms, pulse width: 5 μsec, ppm values relative to TMS = 0.[40] C-1: 176.75; C-2: 60.50; C-3: 29.45; C-3′: 31.40; C-4: 24.50; C-4′: 21.80; C-5: 47.90; C-5′: 47.05; C-6: 174.50; C-6′: −173.80; C-7: 51.80; C-8: 29.15; C-8′: 28.15; C-9: 132.80; C-10: 117.40; C-11: 136.50; C-12: 171.55; C-13: 56.65; C-14: 25.15; C-15: 29.15; C-15′: 28.15; C-16: 182.10.

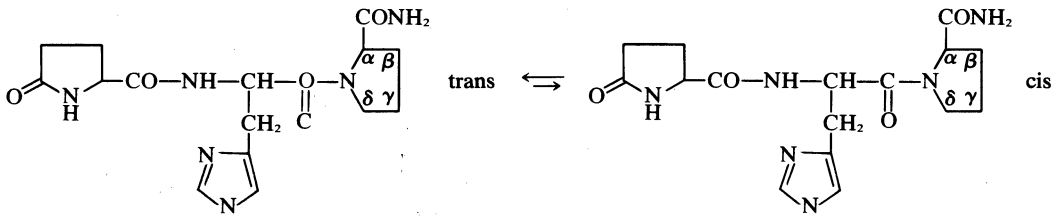
benzyloxycarbonyl-L-aspartic acid which is often used as a derivative for the synthesis of aspartic acid containing peptides.

For demonstration of a ^{13}C NMR spectrum of a peptide hormone that of thyrotropin-releasing hormone (TRH) is given in Fig. 21.

The ^{13}C NMR spectrum of TRH shows four different groups of resonances: at lowest field those of C=O groups, in the ppm range from 115 to 140 the signals of histidine, the group of the α-carbons around 50 ppm and at highest field the resonances of CH_2 carbons. Close to the signals of the prolinamide residue smaller resonances are observed and this fact can be explained only by assuming that the TRH exists in a cis and a trans form in solution.

5. COMPUTER-AIDED INTERPRETATION OF ^{13}C NMR SPECTRA

CMR spectroscopy gives information about the carbon-skeleton of an unknown compound. Compared to almost all other spectroscopic methods it is possible to characterize all the information of a CMR spectrum by a small set of digital values, and the number of signals of a CMR spectrum is less or equal to the number of carbon atoms of the molecule. Up to 1972 there was no systematic computer handling of CMR data in spite of the suitability



for data processing of the problem. Therefore, at that time, we started work on the computer-aided interpretation of CMR spectra.[59-61]

Already the first tests have shown good results in computer-aided structural analysis of unknown compounds.[59] We have build up a data bank containing the most important information from 3000 spectra. The following parameters are stored on magnetic tapes and discs: Names of the compounds, structural formulae, molecular weights, CMR chemical shifts (relative to TMS = 0) multiplicities of the resonances in the proton off-resonance decoupled spectra, lists of the assignments of the signals to carbon atoms, solvents used for the measurements, and the references under which these data were published.[62]

For the input and output of chemical structure a system which is based on the conventional representation of structural formulae is used. This method requires rules for the drawing of structural formulae on schemes for punching. A special print chain must be used for the output of structural formulae. This print chain contains special characters such as double bonds and triple bonds. Figures 22a, b and c show examples of computer prints with the stored information.

The (stored) conventional representation of structural formulae is very well fitted for the dialogue machine user. But this representation is not computer-fitted. The complexity of the conventional structural formula representation and the great matrix required for storage makes data processing nearly impossible. Therefore we used the conventional representation of structural formulae only for input and output. For all other internal purposes of data processing the structural formulae must be coded.

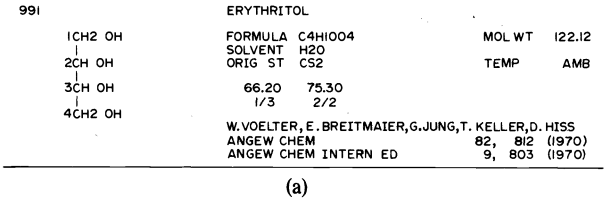
There are some computer-fitted topological representations of chemical structure such as connection tables and linear notations.[63] The principles of the linear notations are outlined using the Wiswesser line-notation (WLN).[63] The forty symbols of the WLN include among others the upper-case alphabetic characters and the ten numerals. All are included in the character sets of standard card-punching and computer line-printing equipment. The symbols serve various functions. First they may represent particular atoms or functional groups. Thus the symbol 1 represents a methyl group and the carbonyl group is respectively V. Second, they may have a syntactic function.

These functions can be distinguished by examination of the context. A simple example of a Wiswesser code is given for a linear molecule: Acetone e.g. is encoded 1 V 1.

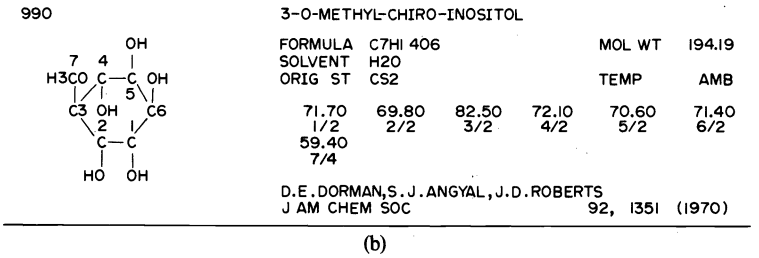
In symmetrical linear structures, the same notation is obviously obtained by starting at either the end of the chain or if unsymmetrical chains are encoded ordering rules are necessary.

The individual carbon atoms of rings are seldom explicitly cited in the WLN, instead a description of the graph is given followed by an indication of the position of the heteroatoms and the degree of saturation. The benzene ring, because of its frequency, receives special treatment—it is encoded as R. Nevertheless a complex series of rules is applied to ensure an unequivocal description of ring systems and of substituents attached to them.

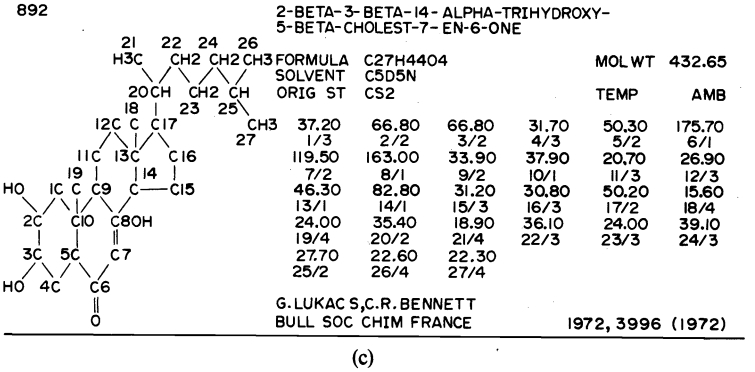
The specific properties of ^{13}C NMR spectra demand an



(a)



(b)



(c)

Fig. 22(a, b, c). Computer prints obtained from the ^{13}C NMR data bank.

adequate documentation of structural formulae. In a system of linear or formal representation and storage of structural formulae the assignment to some specific carbon atom ought to be possible. The usefulness of verbal notations or topological representations for electronic data processing of large collections of formulae must be considered. Bearing this in mind, after investigating the possibilities for adapting linear codings (e.g. WLN)[63] or fragmentation codes (e.g. Gremas)[63] we have developed a modified linear notation.[59] The steps of coding are defined by simple rules. With this system input and output of structural formulae is largely analogous to the standard representation in chemistry: the assignment of ^{13}C NMR signals to carbon atoms in the structural formulae makes use of the conventional numbering of the atoms. Moreover for data processing problems such as structural information retrieval the modified linear notation itself shows several advantages:

1. Only standard character sets are used.
2. Storage problems are reduced.
3. A screen hierarchy gives excellent results for speed, retrieval and redundance.
4. Searching of structures is no problem.

Though manual coding of structural formulae is possible programs for automatic conversion are developed.[59] Other programs e.g. the program SEARCH[59]—which uses the ^{13}C NMR data collection of 3000 compounds are already accomplished and give good results. The program SEARCH is used for the identification of unknown compounds. With the aid of this program, the ^{13}C NMR data collection and the ^{13}C NMR spectrum of an unknown compound the latter can be unambiguously identified, if its data are stored on the file. For compounds whose ^{13}C NMR data are not stored in the data collection, a list of similar compounds and partial structures is found. This list is ordered by decreasing fitting factors, a factor describing the similarity of compounds.[61] Sometimes it is possible to build up the whole unknown molecule with the help of the listed substructures. If e.g. nicotin is assumed to be the unknown compound which is not yet stored in the data collection the program gives a list which shows the substructures pyridine, 3-picoline, *N*-methylpyrrolidine.

The quality of results is scarcely influenced by machines, solvents etc. Our tests have shown essential advantages of computer-aided structural analysis of unknown compounds with the help of ^{13}C NMR data.

Therefore ^{13}C NMR spectroscopy is better qualified for data processing than all other important spectroscopic methods for the following reasons:

1. All information in a ^{13}C NMR spectrum is digitalized.
2. A small number of digital values represents all the information of a ^{13}C NMR spectrum.
3. A small number of digital values gives an extremely specific information of a compound.
4. The comparison of two spectra is very fast.
5. There are no problems of data reduction and information selection. There are also no subjective rules for selection and no subjective weight factors for selected spectral information.

The combination of ^{13}C NMR software with data processing of other spectroscopic or analytic methods is being evaluated. Especially for mass spectroscopy data collections and software are available.[64]

6. STRUCTURE ELUCIDATION OF SOME NATURAL PRODUCTS BY SYNOPTICAL EVALUATION OF DIFFERENT SPECTROSCOPIC PARAMETERS

To speed up the structure elucidation of unknown natural compounds it is necessary to collect as many physicochemical properties as possible (melting point, optical rotation, solubility, absorption, optical rotatory dispersion, circular dichroism, infrared, mass and nuclear magnetic resonance spectra). The final aim is to compare these parameters with computer file-stored values of known compounds by means of a computer and finally to receive suggestions for the structure by electronic data processing.

(a) *Isolation and structure identification of substance* G

Several authors have observed a substance in children's urine which shows similar reactions[65–69] and R_f values in the Bush B 3[70] system to 11-β-hydroxy-androsterone.

Nineteen litres of pre-adolescent children's urine is adjusted to pH 2, saturated with ammonium sulfate and then extracted three times with a mixture of ether/propanol-2. The extract is filtered and evaporated under reduced pressure. The remaining residue is treated for 48 hr at 37°C with acetate buffer of pH 4.3 and then extracted with a 1:1 mixture of ether/ethylacetate. This extract is separated on Whatman No. 3 MM paper in the solvent system benzene: light petroleum: methanol: water (2.5:2.5:3.0:2.0). The spot with an R_f value of 0.39 forms a blue colour with Zimmermann reagent (*m*-dinitrobenzene/ethanol/KOH) and is extracted with absolute ethanol and further purified on Whatman paper No. 2 using light petroleum: methanol: water (5:4:1) for the separation. Final separation was achieved by Sephadex LH-20 column chromatography and 0.5 mg of pure substance are received from the starting material.[71,72]

As this natural product showed many steroid reactions a steroid structure was suggested first.

The compound showed two absorption maxima at $\lambda = 204$ and 249 nm (Fig. 23).[60]

Steroids with absorption properties similar to substance *G* show Cotton effects in the range of the absorption maxima. Circular dichroism measurements proved, however, that substance *G* is an optically inactive compound. Thus with the first two spectroscopic tools applied the

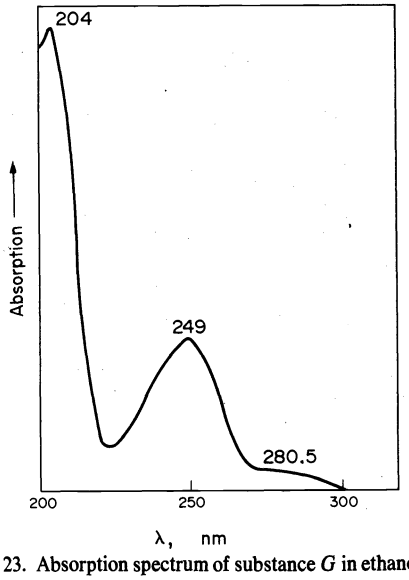


Fig. 23. Absorption spectrum of substance *G* in ethanol.[60]

suggested steroid structure for substance *G* was found to be incorrect.

The i.r. spectrum in CCl_4 shows a concentration dependent N–H stretching vibration at 3450 cm^{-1}. The bands at 3200 and 3100 cm^{-1} are also concentration dependent and allow the suggestion of hydrogen bonding. A carbonyl band at 1710 cm^{-1} has neighbouring shoulders at 1730 and 1745 cm^{-1}. Absorptions at 1945, 1885, 1620 and 712 cm^{-1} could arise from an aromatic part of the molecule (Fig. 24).

In the mass spectrum of substance *G* the most intense peak is that of the molecular ion[71,72] at *m/e* 133. The low molecular weight again is unequivocal proof that the unknown compound has no steroid structure. The peak at *m/e* 77 gives strong evidence that an aromatic residue is a partial structure of the compound (Fig. 25).

From the mass spectroscopic fragmentation pattern, given in Table 2, it can be concluded that $\diagdown C{=}O$ and $\diagdown N{-}\overset{|}{C}{-}$ should be structural elements of substance *G*.

Figure 26 shows the ^{13}C NMR spectrum of substance *G*, recorded in $CDCl_3$.

By means of the computer program SEARCH,[59] the computer file-stored data collection[62] and the ppm values of Fig. 26 a computer print was produced (Fig. 27), which identifies the compound as oxindole.

Synthetic oxindole is in all spectroscopic properties identical with substance *G*.

Table 2. Interpretation of main fragments in the mass spectrum of substance *G*

m/e	Rel. int. (%)	Interpretation
133	100	$M^{\oplus}$
132	6	M-1 –H
105	62	M-28 –CO
104	95	M-29 –(H+CO)
78	53	M-55 –(CO+HCN)
77	26	M-56 –(CO+HCN+H)
52	20	M-81 –(CO+HCN+C_2H_2)
51	25	M-82 –(CO+HCN+C_2H_2+H)

Metastable peaks:

m/e	Interpretation		
103	105→104	m'-1	–H
83	133→105	m'-28	–CO
58.5	104→78	m'-26	–CN, C_2H_2
57	104→77	m'-27	–HCN
33.8	77→51	m'-26	–C_2H_2

(b) *Isolation and structure identification of a urinary steroid from a boy with early sexual maturation*[73]

From 7.5 l. of an acidified (pH 2) urine specimen of a boy with signs of early sexual maturation the steroid conjugates were extracted three times with 2 l. ether/isopropanol (3:1) after addition of ammonium sulfate (50 g/100 ml). The extract was filtered and evaporated to dryness. The remaining residue was then extracted four times with 25 ml ethanol and from the combined ethanol solutions the solvent was removed and the remaining residue taken into 100 ml acetate (0.1 M)/KH_2PO_4 (0.05 M) buffer. 4 ml of a β-glucuronidase solution (120,000 Fishman units) were then added and after 24 hr (37°C) a double extraction with 100 ml ether/ethyl acetate was made. The combined extracts were washed with water, evaporated to dryness and the residue was dissolved in 1 ml ethyl acetate/methanol (2:1) and subjected to descending paper chromatography (Whatman No. 3 MM; light petroleum/benzene/methanol/water, 5:5:6:4). From fraction II[74] of the paper chromatograms the substance was extracted and further separated by silica gel thin layer chromatography using benzene/ethylacetate (3:2) as a solvent. The substance was further purified by two preparative silica gel thin layer chromatographic separations using the solvent systems benzene/ethyl acetate (1:1) for the first and chloroform/ethanol/water (74:26:2) for the second run. 0.5 mg of the unknown compound were available for the structure elucidation.[73]

As no more material could be isolated no ^{13}C NMR spectroscopy and computer analysis could be applied for its structure identification.

The absorption spectrum (Fig. 28), measured in ethanol,

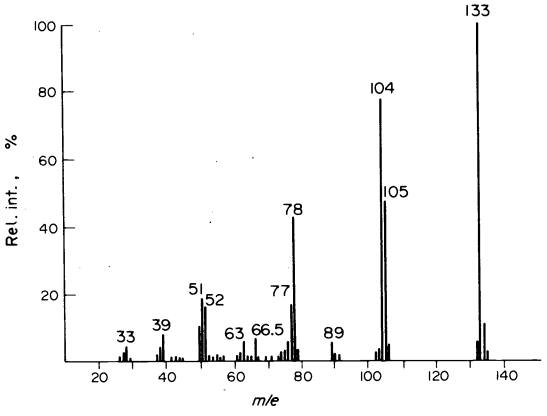


Fig. 25. Mass spectrum of substance *G* (LKB 9000 GC MS instrument, 70 eV ionizing voltage, 3.5 kV accelerating voltage).[60]

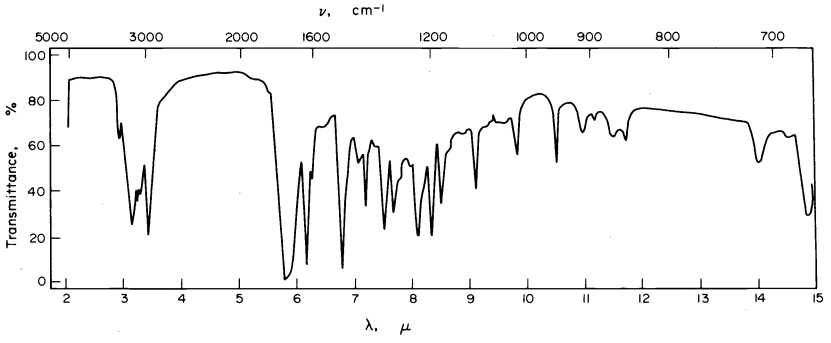


Fig. 24. I.R. Spectrum of substance *G* in $C Cl_4$.[60]

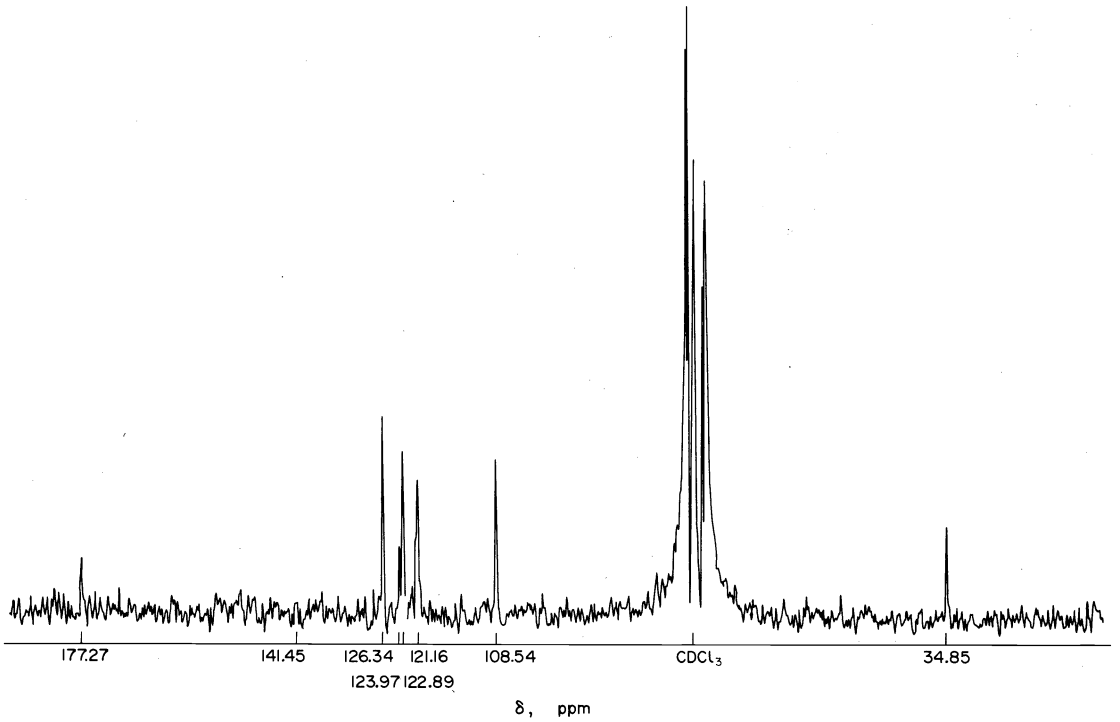


Fig. 26. 22.63 MHz PFT $^{13}C\{^1H\}$ NMR spectrum of substance *G*; 2.8 mg/ml $CDCl_3$, temperature: 30°C, pulse width: 12 μsec, pulse interval: 0.4 sec/4 K interferogram, ppm values relative to TMS = 0, accumulation of 98.032 pulse interferograms.[60]

OXINDOLE

−178.60 − 36.30 − 125.30 − 124.40 − 122.10 − 127.80 − 109.80 − 142.70
MINIMUM CORRESPONDING LINES = 8
FITTING FACTOR = 1.98

ALPHA-ETHYL STYRENE

−148.80 − 109.70 − 140.50 − 126.40 − 126.50 − 125.40 − 126.50 − 126.50
MINIMUM CORRESPONDING LINES = 6
FITTING FACTOR = 1.74

1-METHYL-INDOLE

−129.02 − 101.02 − 120.99 − 121.59 − 119.47 − 109.48 − 129.11 − 137.23 − 31.79
MINIMUM CORRESPONDING LINES = 7
FITTING FACTOR = 1.35

6-METHYL-INDOLE

−124.01 − 102.02 − 120.43 − 121.56 − 131.20 − 111.30 − 126.35 − 137.01 − 21.42
MINIMUM CORRESPONDING LINES = 6
FITTING FACTOR = 1.23

4-CL-BENZONITRIL

−114.00 − 107.00 − 130.40 − 126.30 − 140.80 − 126.30 − 130.40
MINIMUM CORRESPONDING LINES = 5
FITTING FACTOR = 1.20

4-ME-BENZONITRIL

−117.00 − 107.30 − 130.80 − 128.50 − 142.80 − 128.50 − 130.80
MINIMUM CORRESPONDING LINES = 5
FITTING FACTOR = 1.20

2-CH_3-THIPHENE

−139.00 − 124.70 − 126.40 − 122.60
MINIMUM CORRESPONDING LINES = 4
FITTING FACTOR = 1.19

PHENANTHRENE

−127.64 − 125.98 − 125.50 − 121.79
MINIMUM CORRESPONDING LINES = 4
FITTING FACTOR = 1.18

Fig. 27. Computer print produced by the aid of the program SEARCH[59] a ^{13}C NMR data collection[62] and the ^{13}C chemical shift values of substance *G*.[60]

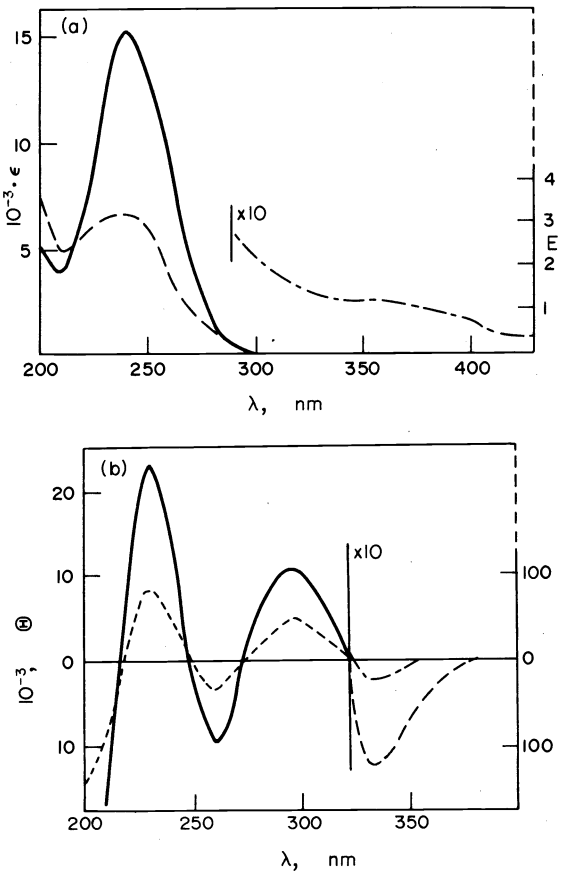


Fig. 28(a). Absorption spectra of the unknown steroid (–·–, ---, 1 mm path length, right scale) and $\Delta^{1,4}$—androstadiene-3.17-dione (left scale) in ethanol.[60] (b). Circular dichroism of the unknown steroid (–·–, ..., 10 mm path length, right scale) and $\Delta^{1,4}$-androstadiene-3.17-dione (left scale) in ethanol.[60]

shows an absorption maximum at $\lambda = 242$ nm. The K-band of α,β-unsaturated ketones ($\pi - \pi^*$ transition) is located at similar wave length.

Four bands are found in the circular dichroism spectra at $\lambda = 332$ (negative), 296 (positive), 260 (negative) and 231 (positive) nm. The negative Cotton effect at 332 nm (R band, $n - \pi^*$ transition) and the circular dichroism extrema in the range of 230–260 nm ($\pi - \pi^*$ transition) are characteristic for α,β-unsaturated keto steroids. The relative intense Cotton effect at 296 nm could arise from an $n - \pi^*$ transition of an isolated keto carbonyl (Fig. 28).

Through the i.r. bands in the finger print region are barely resolved two relatively strong absorptions are observed at 5.98 and 5.73 μm, characteristic for C=O stretching vibrations of α,β-unsaturated cyclohexenones respectively cyclopentanones (Fig. 29).

The molecular weight is obtained from the mass spectrum (*m/e* 284), which also shows characteristic peaks at *m/e* 91, 107, 122 and 159 (Fig. 30).

On the basis of these spectroscopic data an androstane steroid, with at least one α,β-unsaturated keto group and two keto groups in total was suggested.[73]

Spectroscopic comparison of the unknown steroid with $\Delta^{1,4}$-androstadiene—3.17—dione showed that both samples are identical (Figs. 28–30).

(c) *Isolation and structure elucidation of two compounds from the fruit pulp of* Melia azadirachta *Linn*[75,76]

(A) Fruits from *Melia azadirachta* Linn are extracted with ethanol (95%). The extracts are evaporated until green material deposits which is separated from the solution. After evaporation a green residue is received which is dissolved in ether, washed with water and dried over Na_2SO_4. From the solution crystalline needles are received which are chromatographically pure after recrystallization.

The compound has a melting point of 166°C, an $[\alpha]_D^{20}$ value of +11.2° and a molecular weight of 450 (on the basis of its mass spectrum). $C_{28}H_{34}O_5$ is calculated as molecular formula. An absorption at $\lambda = 232.5$ nm ($\epsilon = 25{,}228$) gives evidence for an α,β-unsaturated keto chromophore to be part of the molecule.

Figure 31 shows the i.r. spectrum of the compound registered from a KBr pellet.

From the i.r. spectrum the following sructural information is received:

	band position [cm^{-1}]
Acetyl group	1025
	1250
	1740
α,β-Unsaturated cyclohexenone	1600
	1665
Furan ring	875
	3150

Figure 32 shows the mass spectrum of the natural product.

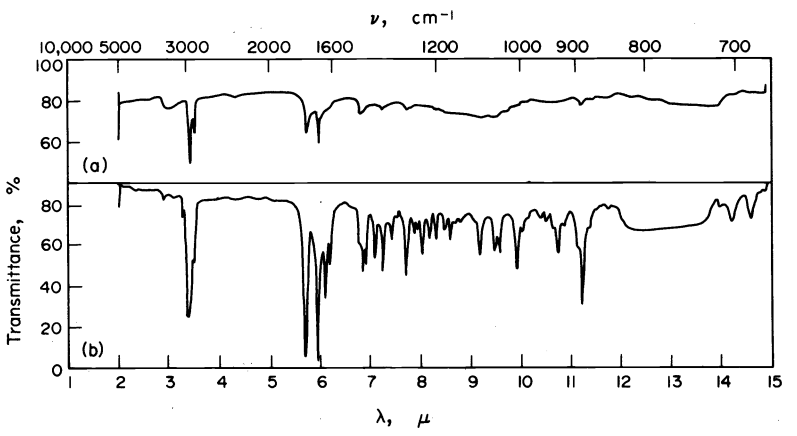


Fig. 29. Infrared spectra of the unknown steroid (a) and $\Delta^{1,4}$-androstadiene-3.17-dione (b) in CCl_4.[60]

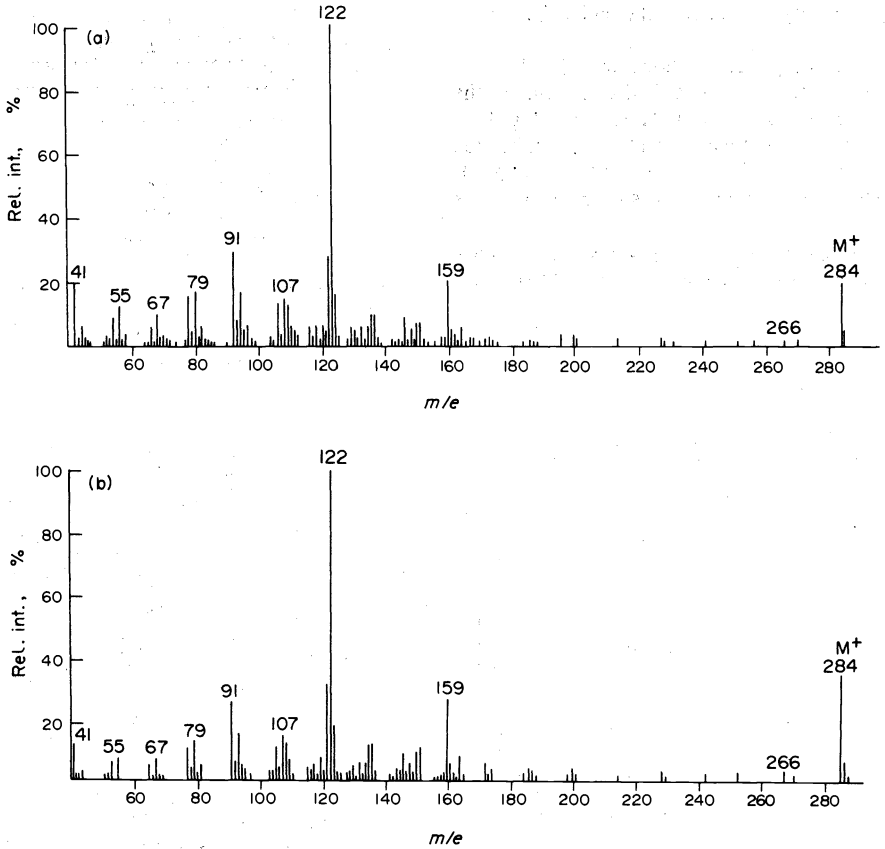


Fig. 30. Mass spectrum of the unknown steroid (a) and $\Delta^{1,4}$-androstadiene-3.17-dione (b) (LKB 9000 instrument, 70 eV).[60]

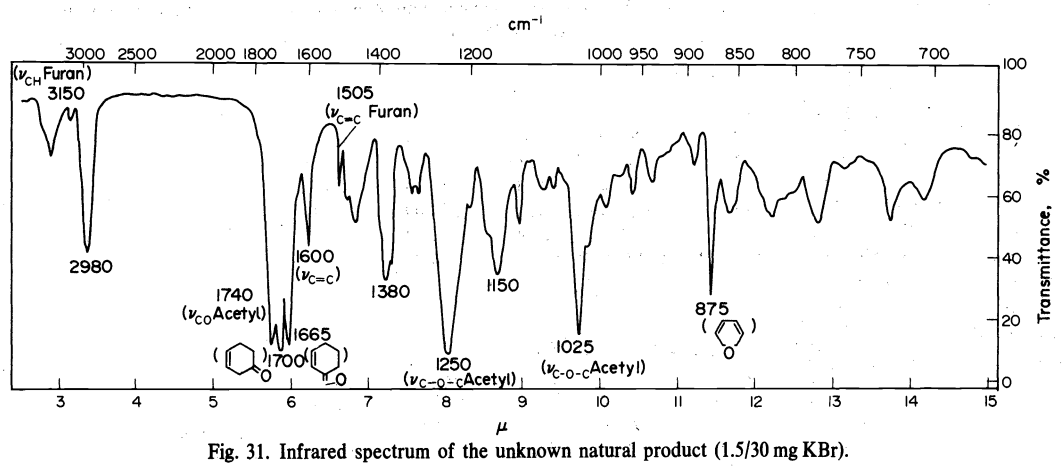


Fig. 31. Infrared spectrum of the unknown natural product (1.5/30 mg KBr).

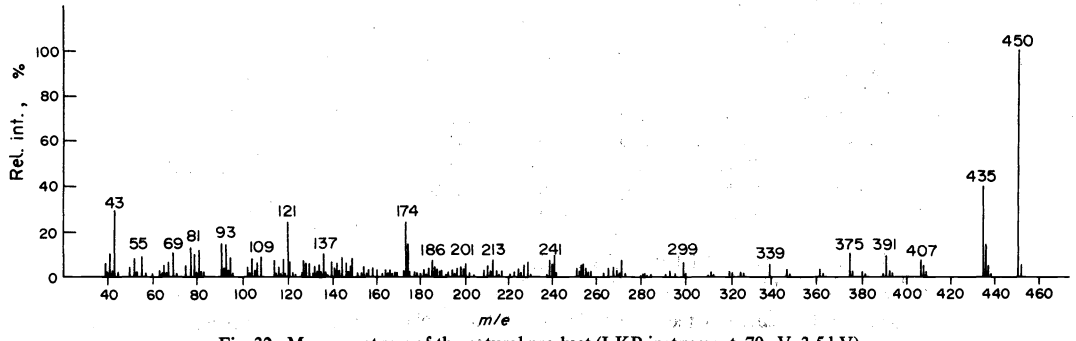


Fig. 32. Mass spectrum of the natural product (LKB instrument, 70 eV, 3.5 kV).

According to the mass spectrum the molecular weight of the product is 450.

Some ions are interpreted as shown in Table 3.

The pulse Fourier transform ^{13}C NMR spectrum of the substance is measured in DMSO and is shown in Fig. 33.

Using the ppm values of Fig. 33 and by the application of the computer program SEARCH[59] and the ^{13}C NMR data collection[62] the computer print shown in Fig. 34 is received.

According to the computer output the most similar file-stored structures to the natural product are cholestanes

Table 3. Interpretation of the most important ions in the mass spectrum of the natural product

m/e	Interpretation
450	M^+
435	$M-CH_3$
407	$M-CH_3CO$
391	$M-CH_3COO$
390	$M-CH_3COOH$
43	CH_3CO

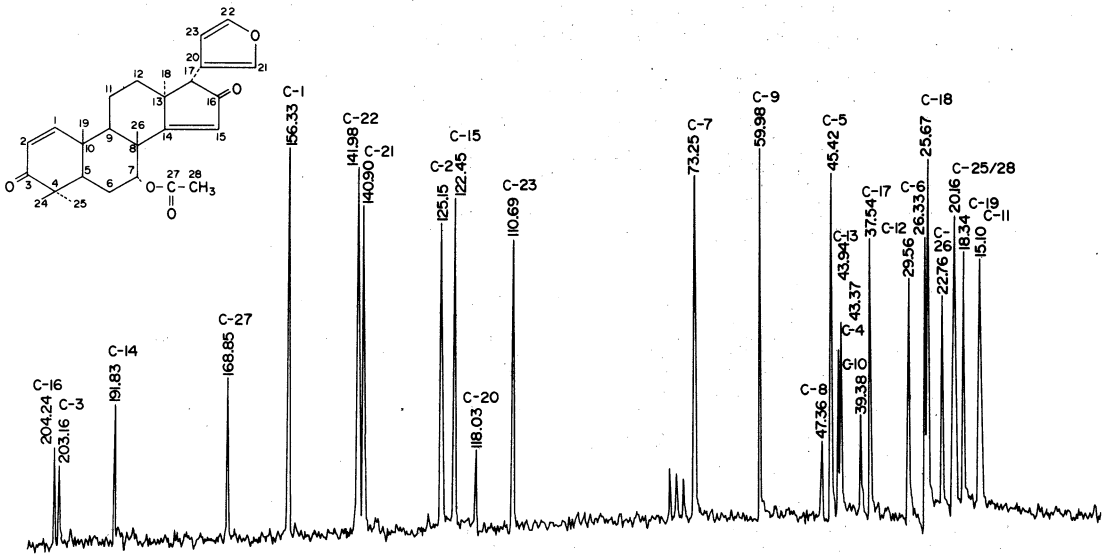
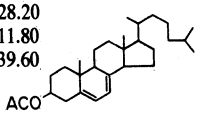


Fig. 33. PFT ^{13}C {^1H} NMR spectrum of the natural product, 22.63 MHz, 500 mg/1.5 ml $CDCl_3$, temperature: 30°C, pulse interval: 0.4 sec/4 K interferogram, phase corrected, accumulation of 2048 pulse interferograms, ppm values relative to TMS = 0.[75]

NR 1(NR* 30)
7-DEHYDROCHOLESTERYL ACETATE

−38.10	−28.20	−72.60	−36.80	141.00	−120.40
−116.70	−138.50	−46.20	−37.20	−21.10	−28.20
−43.00	−54.50	−23.00	−39.40	−56.20	−11.80
−16.00	−36.30	−18.90	−36.30	−24.00	−39.60
−28.00	−22.50	−22.70	−20.90	−169.60	

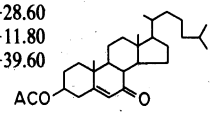
NUMBER OF CORRESPONDING LINES = 22
FITTING FACTOR = 0.748



NR 2(NR* 28)
CHOLEST-5-EN-7-ON-3-BETA-YL-ACETATE

−36.10	−27.50	−72.30	−37.70	−163.20	−126.70
−200.10	−45.30	−49.90	−38.30	−21.30	−28.60
−43.20	−50.20	−26.40	−39.00	−55.50	−11.80
−16.90	−35.90	−18.90	−36.40	−24.10	−39.60
−28.00	−22.40	−22.60	−20.50	−169.30	

NUMBER OF CORRESPONDING LINES = 21
FITTING FACTOR = 0.715



NR 3(NR* 31)
ERGOSTEROL

−38.60	−32.20	−69.70	−41.10	−140.70
−119.40	−116.70	−140.60	−46.50	−37.20
−21.20	−28.30	−43.00	−54.60	−23.10
−39.40	−56.00	−11.80	−16.00	−40.50
−19.40	−132.20	−136.00	−43.00	
−33.20	−19.70	−21.00	−17.40	

NUMBER OF CORRESPONDING LINES = 22
FITTING FACTOR = 0.712

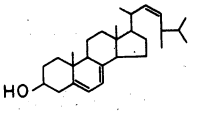


Fig. 34. Computer print produced by the program SEARCH,[59] a ^{13}C NMR data collection[62] and the ^{13}C chemical shifts of Fig. 33.

with one or two carbon–carbon double bonds one acetyl and one keto group.

The signal assignments in Fig. 33 are made on the basis of spectral comparison with similar structures, proton off-resonance spectroscopy and general chemical shift rules.

Lavie *et al.*[77] have isolated from the seed oil of *Melia azadirachta* Linn a compound which could not be crystallised. According to all these investigations we suggest our crystalline product to be azadiradione.

(B) Immediately after harvesting fruits from *Melia azadirachta* Linn are cut into pieces, stored in methanol and after several weeks the extracts are evaporated. During evaporation green material deposits which is removed by filtration. After evaporation to dryness the residue is dissolved in benzene and purified by neutral

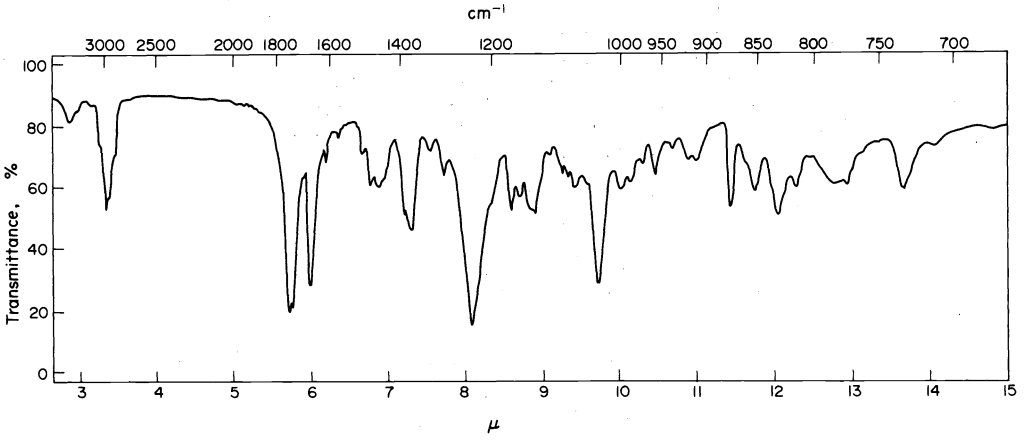


Fig. 35. Infrared spectrum of the compound registered from a KBr pellet (1.5/300 mg KBr).[76]

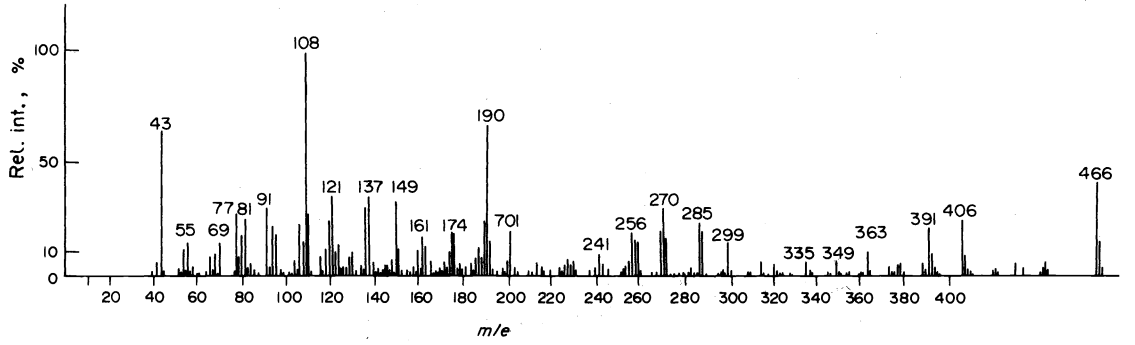


Fig. 36. Mass spectrum of the natural product (LKB instrument, 70 eV, 3.5 KV).[76]

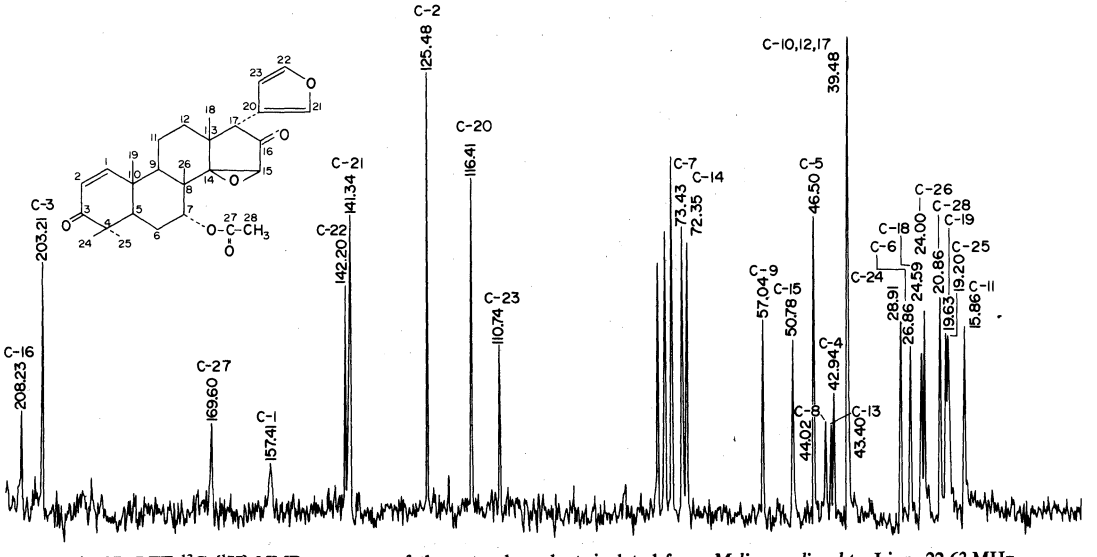


Fig. 37. PFT ^{13}C {^{1}H} NMR spectrum of the natural product, isolated from *Melia azadirachta* Linn, 22.63 MHz, 120 mg/1.5 ml $CDCl_3$, temperature: 30°C, pulse interval: 0.4 sec/4 K interferogram, phase corrected, accumulation of 48,000 pulse interferograms, ppm values relative to TMS = 0.[76]

Al_2O_3. A crystalline product is received from a benzene solution which can be separated into two components by thin layer chromatography. From a benzene methanol (1:1) mixture a crystalline product (melting point 205°C) deposits which is removed by filtration.

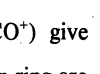
The compound has a melting point of 205°C, and, according to its mass spectrum, a molecular weight of 466. From elemental analysis the molecular formula is calculated to be $C_{28}H_{34}O_6$.

The absorption spectrum shows an intense band at $\lambda = 219$ nm ($\epsilon = 17{,}450$) and low absorption in the range from 310–350 nm; both bonds could arise from $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ transitions of an α,β-unsaturated ketone chromophore respectively.

The i.r. spectrum gives evidence for the following structural elements of the compound:

band position [cm^{-1}]	molecular group
870	furan ring
1030	C–O–C stretching
1240	of acetoxyl group
1665	α,β-unsaturated keto group
1735	five-membered ring keto group
1750	ester carbonyl group

The i.r. spectrum of the natural product is shown in Fig. 35.

The molecular weight of the compound is obtained from its mass spectrum (Fig. 36). The peaks at *m/e* 406 ($M-CH_3COOH$) and *m/e* 108 () give strong evidence that an acetyl group and a furan ring are constituents of the natural product.[76]

The signal assignments in Fig. 37 are based on general chemical shift rules,[51,52,54] proton off-resonance spectroscopy, spectral comparison with similar structures and the application of the computer program SEARCH.[59] The physical and spectroscopic properties are identical with epoxyazadiradione which was isolated recently from the seed oil of *Melia azadirachta* Linn.[77]

Acknowledgements—Most of the work reported in this article has been achieved with the aid of the following scientists and co-workers: H. Bauer, E. Breitmaier, G. Breitmaier, J. Brun, S. Fuchs, D. Gupta, G. Haas, H. Horn, G. Jung, W. A. König, J. Lücke, H. Ringsdorf, H. Schmid, S. Siddiqui, T. N. Waheed and K. Zech. It is a pleasure to express sincere gratitude to all these scientists for their invaluable assistance.

The author wishes to thank the following editors for the permission to reproduce figures: Dozent Dr. H. F. Ebel, *Verlag Chemie*, Weinheim, Figs. 12, 13, 14, 18,[54] 21;[40] Priv. Doz. Dr. K. Maas, editor of *Chemie Experiment und Didaktik*, Georg Thieme Verlag, Stuttgart, Figs. 1–3;[25] Dr. H. Möllinger, editor of *Chemiker Zeitung*, Dr. Alfred Hüthig Verlag GmbH., Heidelberg, Figs. 23–30;[60] Dr. H. Voelter, editor of *Zeitschrift f.Naturforschung*, Verlag der Zeitschrift f.Naturforschung, Tübingen, Fig. 7.[41] Deutsche Fondung gemeinschaft is thanked gratefully for supporting part of this work.

REFERENCES

[1]E. Lederer and M. Lederer, *Chromatography. A Review of Principles and Applications.* 2nd Edn, Elsevier, Amsterdam (1959).
[2]I. M. Hais and K. Macek, *Handbuch der Papierchromatographie.* 1st Edn, Vol. 2, VEB Gustav Fischer, Jena (1960).
[3]F. Cramer, *Papierchromatographie.* 5th Edn, Verlag Chemie, Weinheim/Bergstrasze (1962).
[4]Y. Hashimoto, *Thin Layer Chromatography.* Hirokawa, Tokyo (1962).
[5]K. Randerath, *Dünnschicht-Chromatographie.* 2nd Edn, Verlag Chemie, Weinheim/Bergstraße (1966).
[6]E. Stahl, *Dünnschicht-Chromatographie.* 2nd Edn, Springer, Berlin (1967).
[7]J. C. Giddings, *Dynamics of Chromatography—1. Principles and Theory.* E. Arnold, London (1965).
[8]E. Lederer and M. Lederer, *Chromatography. A Review of Principles and Applications.* 2nd Edn, Elsevier, Amsterdam (1957).
[9]J. J. Kirkland, *Modern Practice of Liquid Chromatography.* Wiley Interscience, New York (1971).
[10]E. Bayer, *Gas-Chromatographie.* 2nd Edn, Springer, Berlin (1962).
[11]R. Kaiser, Chromatographie in der Gasphase, Vol. I–IV, Hochschultaschenbücher Bibliographisches Institut, Mannheim (1960–68).
[12]H. M. McNair and E. I. Bonelli, *Basic Gas Chromatography.* Consolidated Printers Berkeley, California (1969).
[13]N. Hadden, F. Baumann, F. MacDonald, M. Munk, R. Stevenson, D. Gere, F. Zamaroni and R. Majors, *Basic Liquid Chromatography.* Varian Aerograph, Walnut Creek, California (1971).
[14]L. R. Snyder and J. J. Kirkland, *Introduction to Modern Liquid Chromatography*, John Wiley, New York (1974).
[15]F. Hyashi, *J. Biochem. (Japan)* **16**, 1 (1932).
[16]W. H. McNeely, W. W. Binkley and M. L. Wolfrom, *J. Am. chem. Soc.* **67**, 527 (1945).
[17]R. L. Whistler and D. F. Durso, *J. Am. chem. Soc.* **72**, 677 (1950).
[18]H. Rückert and O. Samuelson, *Sätryck ur Svensky Kemisk Tidskrift* **66**, 337 (1954).
[19]J. K. N. Jones, R. A. Wall and A. O. Pittet, *Can. J. Chem.* **38**, 2285 (1960).
[20]J. K. N. Jones and R. A. Wall, *Can. J. Chem.* **38**, 2290 (1960).
[21]J. X. Khym and L. P. Zill, *J. Am. chem. Soc.* **73**, 2399 (1951).
[22]R. B. Kesler, *Anal. Chem.* **39**, 1416 (1967).
[23]D. D. Chilcote, C. D. Scott and W. W. Pitt, *J. Chrom.* **75**, 175 (1973).
[24]W. Voelter and H. Bauer, *G-I-T Fachz. Lab.* **17**, 846 (1973).
[25]W. Voelter and H. Bauer, *Chemie Experiment und Didaktik* **1**, 203 (1975).
[26]J. Böeseken, *Adv. Carbohyd. Chem.* **4**, 189 (1949).
[27]W. Voelter, *Chemiker-Ztg.* **98**, 554 (1974).
[28]R. B. Merrifield, *Recent Prog. Hormone Res.* **23**, 451 (1967).
[29]D. H. Spackman, W. H. Stein and S. Moore, *Anal. Chem.* **30**, 1190 (1958).
[30]K. Zech and W. Voelter, *Chromatographia* **8**, 350 (1975).
[31]W. Voelter and K. Zech, *J. Chromatography* **112**, 643 (1975).
[32]M. Weigele, S. L. DeBernardo, J. P. Tengi and W. Leimgruber, *J. Am. chem. Soc.* **94**, 5927 (1972).
[33]S. Udenfried, S. Stein, P. Böhlen and W. Dairmann, in: *Chemistry and Biology of Peptides*, S.655 (editor J. Meinhofer), Ann Arbor Science Publishers, Michigan (1972).
[34]S. Udenfried, S. Stein, P. Böhlen, W. Dairmann, W. Leimgruber and M. Weigele, *Science* **178**, 871 (1972).
[35]K. Biemann, *Mass Spectrometry, Organic Chemical Applications.* McGraw-Hill, New York (1962).
[36]F. W. McLafferty, *Interpretation of Mass Spectra.* W. A. Benjamin, New York (1966).
[37]G. Spiteller, *Massenspektrometrische Strukturanalyse organischer Verbindungen.* Verlag Chemie, Weinheim/Bergstraße (1966).
[38]H. Budzikiewicz, C. Djerassi and D. H. Williams, *Mass Spectrometry of Organic Compounds.* Holden-Day, San Francisco (1967).
[39]H. Kienitz, *Massenspektrometrie.* Verlag Chemie, Weinheim/Bergstraße (1968).
[40]D. Gupta and W. Voelter, *Hypothalamic Hormones.* Verlag Chemie, Weinheim (1975).
[41]W. A. König, S. Fuchs, K. Zech and W. Voelter, *Z. Naturforsch.* **28b**, 820 (1973).

[42]K. Zech, H. Horn and W. Voelter, *Chemiker-Ztg.* **98**, 209 (1974).
[43]F. Korthe, *Methodicum Chimicum.* Vol. 1, Georg Thieme, Stuttgart (1973).
[44]H. Ringsdorf *et al.* in preparation.
[45]J. A. Pople, W. G. Schneider and H. J. Bernstein, *High-Resolution Nuclear Magnetic Resonance.* McGraw Hill, New York (1959).
[46]H. Suhr, *Anwendung der kernmagnetischen Resonanz in der organischen Chemie.* Springer, Berlin (1965).
[47]J. W. Emsley, J. Feeney and L. H. Sutcliffe, *High Resolution Nuclear Magnetic Resonance Spectroscopy.* Pergamon Press, Oxford (1966).
[48]E. F. Mooney, *Annual Review of NMR Spectroscopy.* Academic Press, London (1968, 1969).
[49]A. F. Casy, *PMR Spectroscopy in Medicinal and Biological Chemistry.* Academic Press, London (1971).
[50]The Sadtler Standard Spectra, Sadtler Research Laboratories, Philadelphia.
[51]G. C. Levy and G. L. Nelson, *Carbon-13 Nuclear Magnetic Resonance for Organic Chemists.* Wiley-Interscience, New York (1972).
[52]J. B. Stothers, *Carbon-13 NMR Spectroscopy.* Academic Press, New York (1972).
[53]T. C. Farrar and E. D. Becker, *Pulse and Fourier Transform NMR.* Academic Press, New York (1971).
[54]E. Breitmaier and W. Voelter, ^{13}C *NMR Spectroscopy.* Verlag Chemie, Weinheim (1974).
[55]J. J. Lowe and R. E. Norberg, *Phys. Rev.* **107**, 46 (1957).
[56]W. Bremser, H. D. W. Hill and R. Freeman, *Meßtechnik* **78**, 14 (1970).
[57]R. A. Hoffmann and S. Forsén, *Prog. NMR Spectroscopy* **1**, 82 (1966).
[58]K. F. Kuhlmann and D. M. Grant, *J. Am. chem. Soc.* **90**, 7355 (1968).
[59]G. Haas, Thesis of the University of Tübingen, FRG, in press.
[60]W. Voelter, E. Breitmaier, G. Breitmaier, D. Gupta, G. Haas and W. A. König, *Chemiker-Ztg.* **97**, 239 (1973).
[61]W. Voelter, G. Haas and E. Breitmaier, *Chemiker-Ztg.* **97**, 507 (1973).
[62]E. Breitmaier, G. Haas and W. Voelter, *Atlas of Carbon-13 NMR Spectra* Heyden & Son, London (1975).
[63]D. C. Veal, *Fortschritte der chemischen Forschung* **39**, 65 (1973); E. G. Smith, *The Wiswesser Line-Formula Chemical Notation.* McGraw-Hill, New York (1968).
[64]United Kingdom Atomic Energy Authority, Aldermaston, Berkshire.
[65]D. Gupta and J. M. Tanner, *Nature, Land.* **203**, 187 (1964).
[66]D. Gupta, Dissertation Faculty of Science, University of London (1965).
[67]D. Gupta and J. Goodwin, *Steroids* **8**, 195 (1966).
[68]J. M. Tanner and D. Gupta, *J. Endocrinol.* **41**, 139 (1968).
[69]D. M. Cathro, K. Birchall, F. L. Mitchell and C. C. Forsyth, *J. Endocrinol.* **27**, 53 (1963).
[70]I. E. Bush, *Biochem. J.* **50**, 370 (1952).
[71]W. Voelter, G. Jung, E. Breitmaier, W. König, D. Gupta and G. Breitmaier, *Z. Naturforsch.* **26b**, 1380 (1971).
[72]D. Gupta, G. Breitmaier, W. Voelter, G. Jung, E. Breitmaier and W. König, *Nature New Biol. (London)*, **235**, 55 (1972).
[73]G. Jung, W. A. König, W. Voelter, G. Breitmaier and D. Gupta, *Hoppe-Seyler's Z. Physiol. Chem.* **353**, 1005 (1972).
[74]D. Gupta and W. Marshall, *Acta Endocrin.* **68**, 141 (1971).
[75]S. Siddiqui, T. N. Waheed, J. Lücke and W. Voelter, *Z. Naturforsch.* **306**, 961 (1975).
[76]S. Siddiqui, T. N. Waheed, J. Lücke and W. Voelter, *Chemiker-Ztg.* **99**, 504 (1975).
[77]D. Lavie and M. K. Jain, *Chem. Commun.* 278 (1967).