Tetrahedron Vol. 48, No. 19, pp. 3915-3924, 1992
Printed in Great Britain

0040-4020/92 $5.00+.00

Novel Triterpene-derived Hydrocarbons of Arborane/Fernane Series in Sediments. Part I.

Verena HAUKE, Roland GRAFF, Patrick WEHRUNG, Jean M. TRENDEL and Pierre ALBRECHT*

Faculté de Chimie, Université Louis Pasteur, 1 rue Blaise Pascal, 67000 Strasbourg, France

Lorenz SCHWARK

Lehrstuhl für Geologie, Geochemie und Lagerstätten des Erdöls und der Kohle, RWTH Aachen, Lochnerstr. 4-20, 5100 Aachen, Federal Republic of Germany

Brendan J. KEELY and Torren M. PEAKMAN

School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, Great Britain

(*Received in Belgium* 11 *March* 1992)

Key Words: des-*A*-arbora-5,7,9-triene; 24,25-dinorarbora-1,3,5(10),9(11)-tetraene; 24,25-dinorarbora-1,3,5,7,9-pentaene; pentacyclic triterpene-derived aromatic hydrocarbons; isolation from sediments and structural identification by NMR spectroscopy.

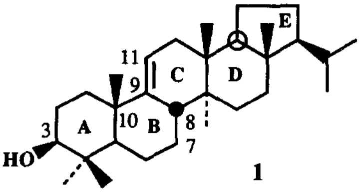
***Abstract:** Three aromatic polycyclic hydrocarbons were isolated from Eocene shales and the Permian Kupferschiefer, and their molecular structures established by 1D and 2D NMR spectroscopy. These compounds (des-A-arbora-5,7,9-triene; 24,25-dinorarbora-1,3,5(10),9(11)-tetraene; 24,25-dinorarbora-1,3,5,7,9-pentaene) derive from microbial transformations of pentacyclic triterpenes oxygenated at C-3 belonging to the arborane or fernane families.*

INTRODUCTION

Aromatisation reactions count among the most important geochemical transformations undergone by the polycyclic skeletons of biological molecules (steroids, terpenoids) in the sub-surface. They lead to the formation of natural polycyclic aromatic hydrocarbons by processes which could occur as well at advanced stages of maturation as during initial steps of diagenesis. In the latter case, where mediation of micro-organisms appears essential,[1] two major pathways of early aromatisation of terrestrial plant pentacyclic triterpenes oxygenated at C-3 (*e.g.* **1** or **2**) have been recognised : a) starting from ring A, triggered by the elimination of the oxygenated function at C-3, and proceeding sequentially towards ring D(E),[2] and b) from ring B to ring D(E), after loss of ring A.[3]

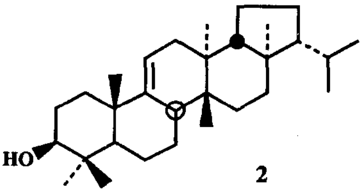
We wish to report here the precise structural elucidation by high-field NMR spectroscopy of three new aromatic polycyclic hydrocarbons, isolated from the Eocene Messel (Germany) and Ménat (France) shales[4] and the Permian Kupferschiefer (Germany),[5] which are related to C-3 oxygenated triterpenes, most likely to the uncommon arborane series : des-*A*-arbora-5,7,9-triene **3**, 24,25-dinorarbora-1,3,5(10),9(11)-tetraene **4** and

24,25-dinorarbora-1,3,5,7,9-pentaene **5**. These compounds constitute important links of the triterpene aromatisation sequence, which further strenghten the aromatisation pathways described previously. Moreover, their predominance, as well as that of their functionalised precursors in sediments, in particular of lacustrine origin, suggest an alternative source to that solely from terrestrial plants, an hypothesis reinforced by carbon isotopic composition measurements.



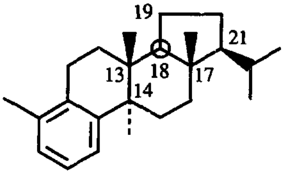
1

arbor-9(11)-en-3β-ol
(= isoarborinol)

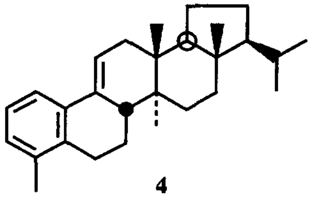


2

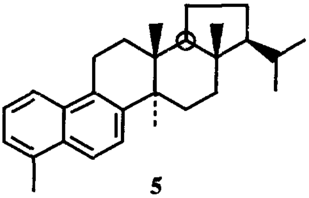
fern-9(11)-en-3β-ol



3



4

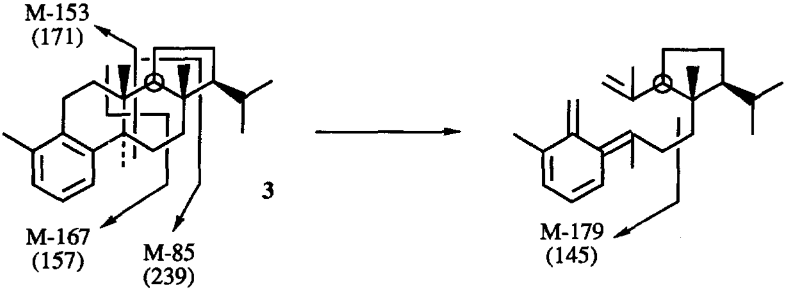


5

RESULTS

The investigated compounds were isolated in mg quantities by successive chromatographic steps (silica-gel CC and TLC, silver nitrate-impregnated silica-gel CC, reverse phase HPLC) from the sediment organic extracts as indicated in the experimental part.

Electron impact mass spectra of the three hydrocarbons (see experimental part, **3** : M^+=324, $C_{23}H_{34}$; **4** : M^+=376, $C_{28}H_{40}$; **5** : M^+=374, $C_{28}H_{38}$) show common and caracteristic patterns of fragmentation, as illustrated below for **3**.



In each case, a structurally significant ion at M-85 is observed, explained by decomposition of ring E (cleavage of the 18,19 and 17,21 bonds), whereas the fragment M-43 (loss of the isopropyl group) is generally very weak, two features found in the arborane/fernane series.[6] Also characteristic of pentacyclic triterpenoid structures with methyl substituents at C-13 and C-14 and with an unsaturation at 7, 8 or 9(11) position, is the fragment triplet at M-153, M-167 and M-179 which could be interpreted according to Budzikiewicz *et al.*[7]

A rigorous identification of the three isolated products from the Eocene shales was achieved by high-field ^1H and ^{13}C NMR studies which led to complete assignment of proton and carbon chemical shifts (tables 1, 2 and 3, see experimental part) using, for each compound, the same set of experiments. Compound **3**, isolated from the Permian Kupferschiefer, was similarly identified on the basis of ^1H NMR experiments. The ^1H NMR spectra (figures 1a, 2a, 3a) show, in each case, the presence of four methyl singlets (or slightly split, J<1 Hz) - the absorption at 2.1-2.7 ppm being associated with a benzylic methyl group -, two methyls as doublets (J 6.5 Hz) and deshielded signals ranging from 6 to 8 ppm which correspond to aromatic protons or conjugated vinylic protons. ^{13}C NMR ^1H broad band decoupled and DEPT spectra allowed carbon multiplicities to be determined. Analyses of phase-sensitive ^1H-^1H COSY and inverse one-bond ^1H-^{13}C correlation experiments permitted establishment of the proton connectivity patterns (figures 1b, 2b, 3b) - and subsequently the corresponding sequences of carbons - into sectors which are not through-coupled (except with long-range effects) and which are delimitated by quaternary carbons. Linkage of the methyl groups located on non aromatic quaternary carbons to a sector was possible through long-range ^4J (W-type) couplings involving the axial proton of the methylene/methine groups vicinal to the quaternary carbon. Connectivities between isolated ^1H-coupled sectors were established *via* phase-sensitive NOESY experiments which also furnished essential stereochemical information (figures 1c, 2c, 3c). In particular, it was possible to assign to the right part of the molecules (ring C, D and E) an all-*trans* skeleton with a five membered E-ring bearing one isopropyl chain. More precisely, from n.O.e's observed, 26-CH_3, H-18 and H-21 on the one hand, 27-CH_3, 28-CH_3, and 22-CH on the other hand, are on opposite sides of the molecule mean plane. Concerning compound **4** which possesses a 9(11) double bond, H-8 has been clearly shown to be on the same side as 27-CH_3 from NOESY experiment. Finally, inverse long-range ^1H-^{13}C ($^{2,3}J_{CH}$) correlation experiments allowed unequivocal and full establishment of the carbon sequences of components **3** and **4** (except C-19/C-20, C-21/C-20 and C-15/C-16 connections for **4**), starting from the methyl groups and some individual protons, thus confirming the previous proton and carbon assignments. It should be noted, in addition, that carbon signals for the right part of the molecules (ring E with its substituents and part of ring C) are in excellent agreement with those recorded for fern-9(11)-ene,[8] especially for the most closely related structure **4**.

DISCUSSION

In respect with their structure, the three identified hydrocarbons obviously result from the transformation of pentacyclic triterpenes possessing an oxygenated function at C-3 which seems necessary for inducing ring-A degradation as well as initiating ring-A and further aromatisation processes (compound **4** is the direct precursor of **5**).

It was not possible, however, solely on the basis of our studies, to determine unambiguously to which triterpene family - arborane or fernane - the isolated compounds are related. Indeed, loss of chirality at position 10 consecutive to ring A and/or ring B aromatisation, leads to enantiomeric structures in both series. Nevertheless, in our case, the three hydrocarbons from the Eocene samples are most likely related to the arborane series for the following reasons. First, in both shales, isoarborinol **1** and the corresponding ketone arborinone occur as major components in their respective fraction,[4,9,10] suggesting that our aromatic hydrocarbons could derive from such potential precursors. Secondly, the carbon isotopic composition of compound **3** (Messel shale) measured by gas chromatography - isotopic ratio mass spectrometry (GC-IRMS) on the mono- and di-aromatic fraction (see experimental part) is nearly identical with that of isoarborinol previously isolated from the same shale, suggesting closely related compounds. Indeed, the $\delta^{13}C$-value[11] recorded for **3** is -28.7‰, whereas that found for **1** is -28.0‰.[12] At the same time, all $\delta^{13}C$-values found for other triterpene-derived hydrocarbons (lupane, ursane series) occurring in the same shale are significantly different and range from -23 to -25‰.[13,14] In the case of the Permian Kupferschiefer no alcohols or ketones have been observed due to its greater level of maturity and so an origin from an arborane precursor is less certain.

Hydrocarbons **3** and **5** have been detected previously in Messel and Ménat shales,[15] and more recently the

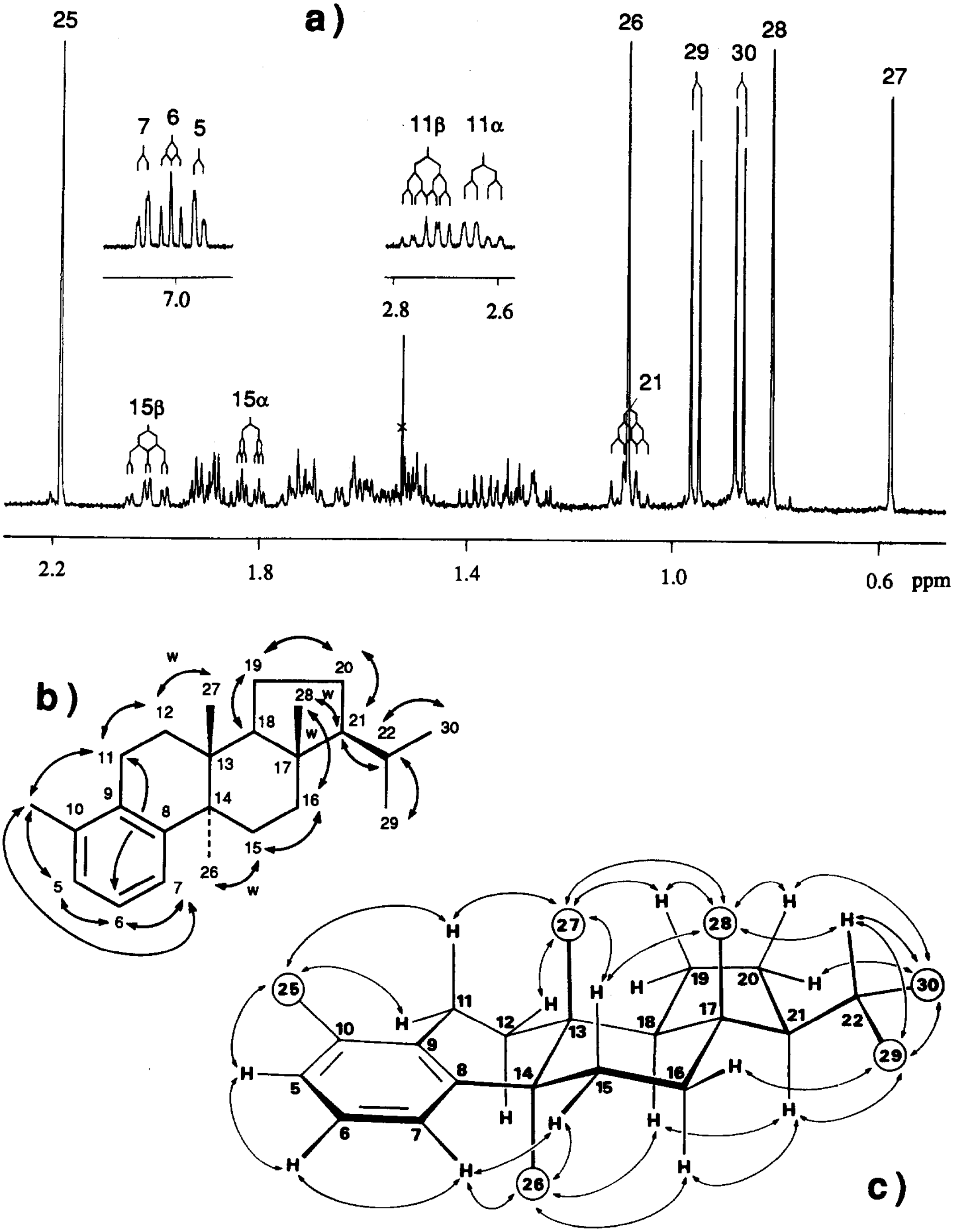


Figure 1 : ^{1}H-NMR spectrum and connectivity patterns observed for des-*A*-arbora-5,7,9-triene 3
a) ^{1}H-NMR spectrum (400 MHz, CD_2Cl_2, 300 K)
b) Connectivities from COSY experiment (W = long-range W-type coupling)
c) Connectivities from NOESY experiment

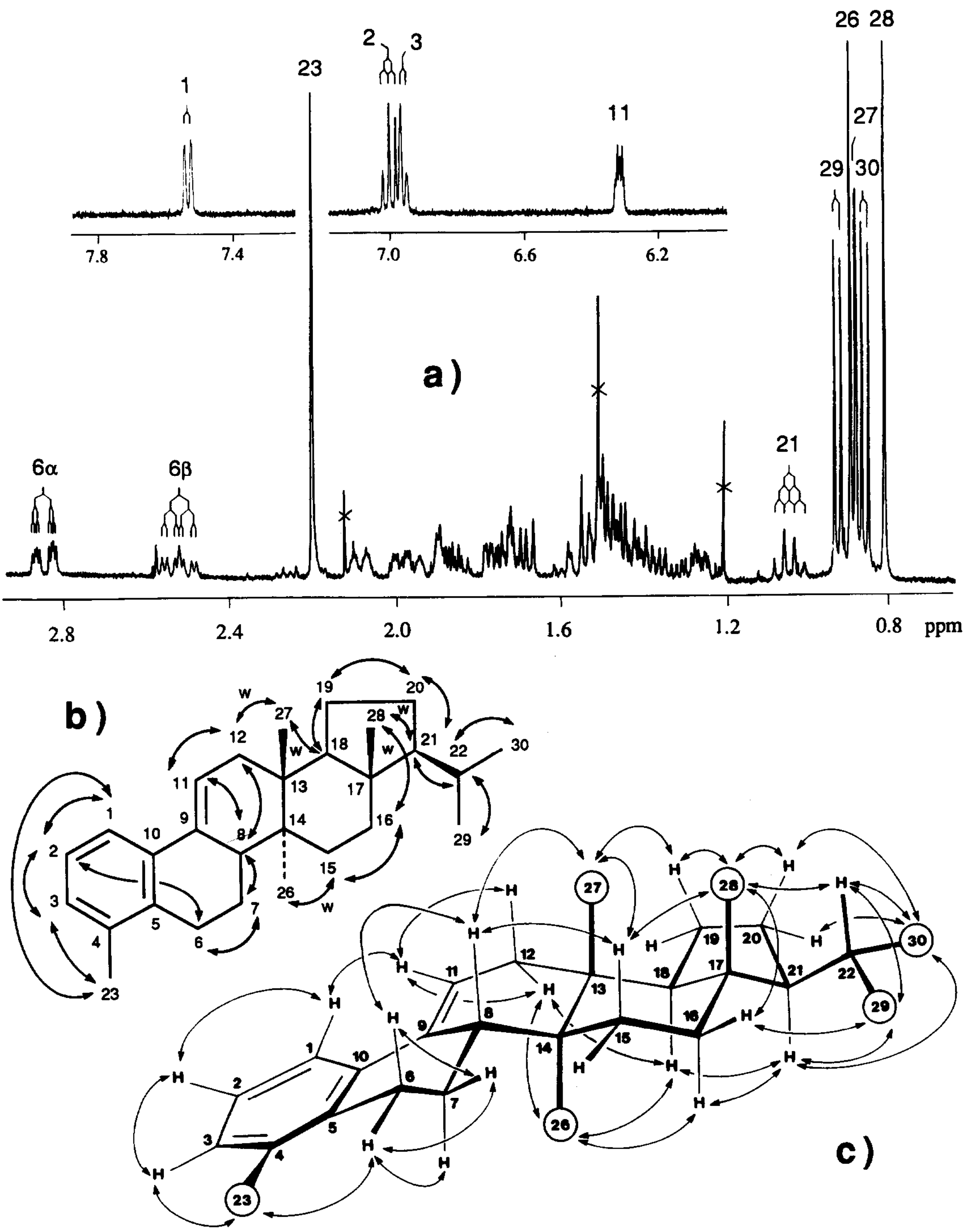


Figure 2 : ^{1}H-NMR spectrum and connectivity patterns observed for 24,25-dinorarbora-1,3,5,9(11)-tetraene 4
a) ^{1}H-NMR spectrum (400 MHz, CD_2Cl_2, 300 K)
b) Connectivities from COSY experiment (W = long-range W-type coupling)
c) Connectivities from NOESY experiment

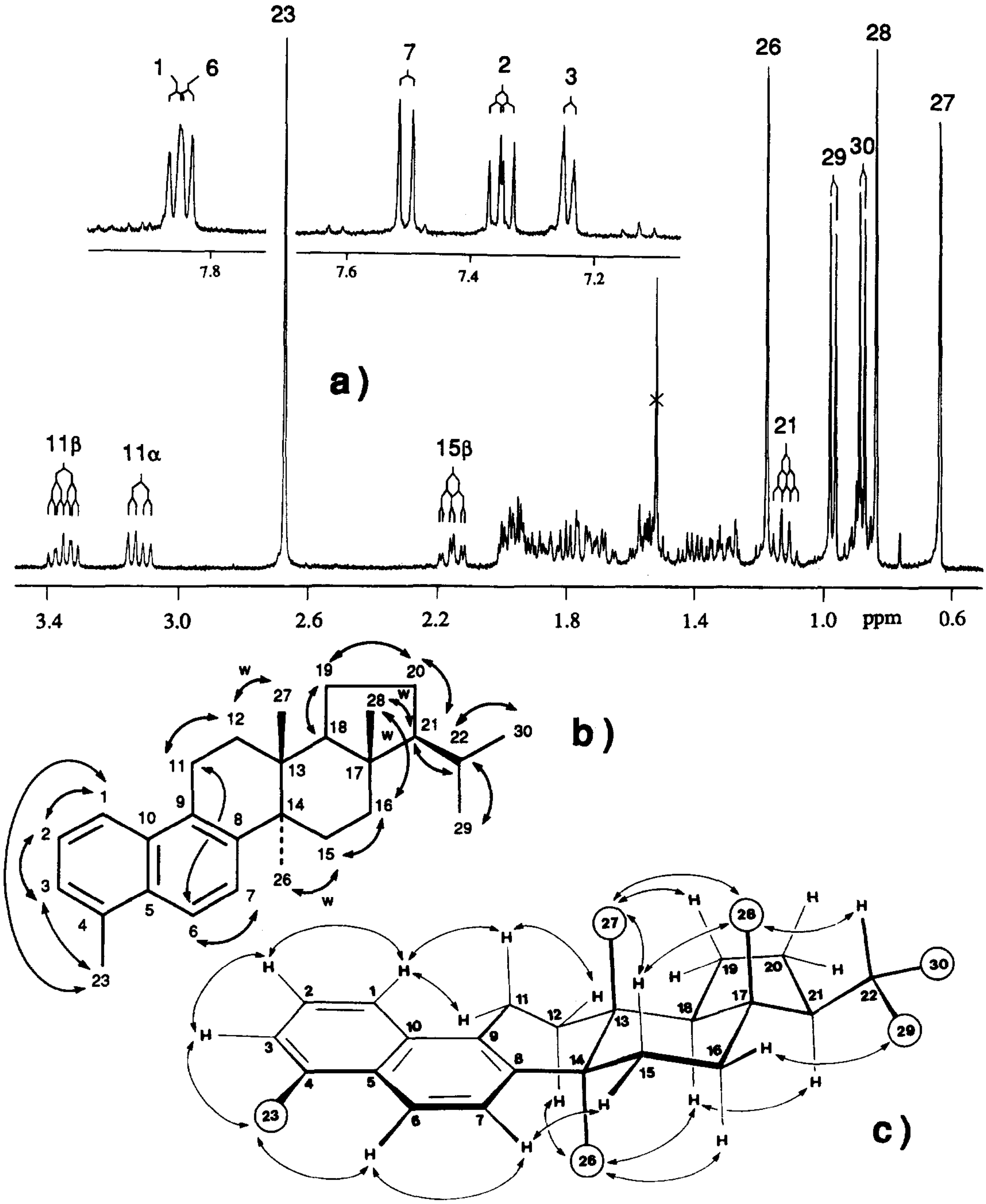


Figure 3 : ^{1}H-NMR spectrum and connectivity patterns observed for 24,25-dinorarbora-1,3,5,7,9-pentaene 5

a) ^{1}H-NMR spectrum (400 MHz, CD_2Cl_2, 300 K)

b) Connectivities from COSY experiment (W = long-range W-type coupling)

c) Connectivities from NOESY experiment

occurrence of 3 - in high concentration - has been reported in a lacustrine shale from southeast Brazil,[16] and in the Permian Kupferschiefer,[5] but with diverse structural hypotheses essentially based on mass spectral data. It should still be kept in mind, however, that depending on the original organic input, all identified aromatic structures could also arise from precursors of the fernane family, as it has been proposed.[16]

The arborane skeleton, and particularly isoarborinol, is rather uncommon in nature, although widely distributed in some plant families, especially in *Graminae*,[17] and its occurrence in geological sources has therefore been interpreted as an indication of a terrestrial contribution to the organic matter.

However, isoarborinol and its derivatives are often found as major components in various sediments,[4,9,10,18] in particular of lacustrine origin - 5 represents about 60% of the Ménat aromatic fraction - whereas their usual companions in *Graminae* (3β-methoxy-derivatives, fernenol) could generally not be detected.[10] This has led to the proposal of a microbiological origin (aerobic bacteria ?) for arborane derivatives in sediments.[19] This point of view, or at least the hypothesis of a particular origin (algal ?) appears reinforced by the carbon isotopic composition of arborane derivatives in the Messel shale as compared with those of all co-occurring terrestrial plant triterpene related hydrocarbons which are markedly different. In this respect it is noteworthy that isoarborinol has also recently been found in a lichen.[20] In the case of the Permian Kupferschiefer, a sample whose age predates the evolution of *Graminae* and other angiosperms, hydrocarbon 3 could either derive from arborane precursors of the same sources as that envisaged for Eocene shales (*i.e.* microbial or algal material, lichens), or from 3β-fernenols which could indicate a contribution from organisms such as ferns.[21]

CONCLUSION

Three aromatic polycyclic hydrocarbons related to 3-oxygenated pentacyclic triterpenes have been conclusively identified by NMR spectroscopy after their isolation from two Eocene shales and one Permian shale. They derive most likely by microbiological transformations from precursors belonging to the relatively rare arborane family. The striking abundance of arborane derivatives in some sediments, often of lacustrine origin, as well as their carbon isotopic composition, further support the hypothesis of an alternative source (microbiological ?) to that generally accepted, *i.e.* terrestrial plants.

EXPERIMENTAL

Physical measurements. Electron Impact Mass Spectra were obtained on a GC-MS Finnigan MAT TSQ 70 equipment. GC-IRMS analyses were performed on a Finnigan MAT DeltaS-GC. NMR experiments for compounds 3, 4 and 5 from the Eocene shales were done on Bruker AM-400 and AMX-500 spectrometers operating at observation frequencies of 400/500 and 100/125 MHz for ^{1}H and ^{13}C nuclei respectively and data were recorded at 300 K. The chemical shifts (δ) are reported in ppm from TMS using the solvent (CD_2Cl_2 δ^1H 5.32 ppm, δ^{13}C 53.84 ppm) as internal reference. J values result from first-order interpretations and are given to the nearest 0.5 Hz. Samples were not degased and concentrations used ranged from 0.01 to 0.06M. NOESY experiments were performed, after T_1 measurements, with both relaxation delay and mixing time set at 1.5s. For inverse long-range ^{1}H-^{13}C correlation experiments, the J-selection delay was tuned for $^nJ_{CH} = 7$ Hz. Similar ^{1}H NMR experiments on compound 3 from the Permian Kupferschiefer were obtained on a Jeol GX400 spectrometer operating at 400 MHz but using $CDCl_3$ instead of CD_2Cl_2.

Isolation procedure of the aromatic hydrocarbons. 3 : a mono- and di-aromatic fraction obtained by liquid chromatography (CC, silica-gel, hexane) of the Messel shale extract (toluene/methanol 3:1 once, and then 2:1 twice, 50°C, 6h) was subjected to preparative HPLC (DuPont Zorbax ODS 250x21.2mm, 8µm 80Å, MeOH 25 ml/min). Further separation of the collected fractions on an analytical HPLC column (Baker 250x4.6mm, Bakerbond WP Octadecyl, 5µm 300Å, MeOH/H_2O 98:2 1ml/min) led to the isolation of 3 with a purity of 92% (GC). 3 : the aromatic fraction obtained by liquid chromatography (CC, silica-gel, CH_2Cl_2 after hexane) of the Kupferschiefer extract (CH_2Cl_2, 24h) was further separated by liquid chromatography (CC, silver nitrate impregnated-silica gel, hexane/ether 99:1) to give **3** in 95% purity (GC). **4** and **5** : the Ménat shale extract ($CHCl_3$/MeOH 3:1 twice, and Toluene/MeOH 3:1 twice, 50°C, 1h) was roughly fractionated by CC (silica-gel, dichloromethane, diethylether, MeOH/$CHCl_3$/H_2O 25:60:4). The less-polar first-eluting fraction (dichloromethane) was again chomatographed on column (silica-gel, gradient of dichloromethane in hexane) in order to obtain the aromatic fraction. Further fractionation of the latter on TLC (silica-gel, hexane) afforded **5** (90% purity by GC) and a fraction containing **4** which was subjected to reverse phase HPLC (Baker 250x4.6mm, Bakerbond WP Octadecyl, 5µm 300Å, MeOH 1ml/min) allowing **4** to be isolated (93% purity by GC).

NMR data.

Table 1. ^1H and ^{13}C Chemical Shift Data of des-*A*-arbora-5,7,9-triene **3** (CD_2Cl_2, 300K)

C	$\delta^{13}C$ ppm	δ^1H ppm	*J* Hz	δ^1H ppm	*J* Hz
5	127.18	6.96	7.5 (br.)		
6	125.25	7.01	7.5, 7.5		
7	121.20	7.06	7.5 (br.)		
8	147.85				
9	134.91				
10	136.39				
11	22.73	2.74 (β)	18.0, 10.0, 7.5	2.63 (α)	18.0, 10.0 (br.)
12	31.18	1.72 (α)		1.59 (β)	13.0, 10.0, 1.5
13	36.97				
14	41.33				
15	29.65	2.02 (β)	13.5, 13.5, 3.5	1.82 (α)	13.5, 3.5, 3.5
16	36.62	1.90 (β)	13.5, 3.5, 3.5	1.61 (α)	13.5, 13.5, 3.5
17	43.48				
18	53.38	1.72 (α)	13.0, 7.0		
19	20.96	1.51 (α)		1.36 (β)	13.0, 13.0, 12.0, 5.5
20	28.73	1.90 (α)		1.28 (β)	
21	60.22	1.08 (α)	9.5, 9.5, 9.5		
22	31.24	1.51			
25	19.94	2.184			
26	25.57	1.087	0.7		
27	16.95	0.577	1.0		
28	14.22	0.806			
29	22.39	0.956	6.5		
30	23.19	0.869	6.5		

Table 2. ^1H and ^{13}C Chemical Shift Data of 24,25-dinorarbora-1,3,5(10),9(11)-tetraene **4** (CD_2Cl_2, 300K)

C	$\delta^{13}C$ ppm	δ^1H ppm		C	$\delta^{13}C$ ppm	δ^1H ppm	
1	121.59	7.53		**15**	29.76	~1.57 (β)	~1.50 (α)
2	125.65	6.98		**16**	36.44	~1.75 (β)	~1.52 (α)
3	128.09	6.96		**17**	43.44		
4	136.82			**18**	52.35	1.70 (α)	
5	135.65			**19**	20.54	1.33-1.40	1.33-1.40
6	28.14	2.88 (α)	2.55 (β)	**20**	28.63	1.87 (α)	1.26 (β)
7	22.96	2.00 (β)	1.47 (α)	**21**	60.10	1.04 (α)	
8	43.27	2.11 (β)		**22**	31.24	1.48	
9	135.09			**23**	19.93	2.221	
10	135.27			**26**	15.88	0.891	
11	118.83	6.32		**27**	16.31	0.877	
12	37.70	1.93 (α)	1.76 (β)	**28**	14.22	0.805	
13	37.11[a]			**29**	22.34	0.923	
14	38.56[a]			**30**	23.17	0.854	

a : signals may be interchanged

Table 3. ^{1}H and^{13}C Chemical Shift Data of 24,25-dinorarbora-1,3,5,7,9-pentaene **5** (CD_2Cl_2, 300K)

C	δ^{13}C ppm	δ^1H ppm		C	δ^{13}C ppm	δ^1H ppm	
1	121.62	7.86		15	29.61	2.15 (β)	1.96 (α)
2	125.55^a	7.35		16	36.59	1.96 (β)	1.68 (α)
3	125.59^a	7.24		17	43.45^c		
4	131.18^b			18	53.40	1.79 (α)	
5	131.25^b			19	20.93	1.56 (α)	1.40 (β)
6	122.07	7.84		20	28.72	1.92	1.30
7	122.47	7.51		21	60.17	1.12 (α)	
8	132.76^b			22	31.22	1.52	
9	135.04^b			23	19.82	2.669	
10	144.61^b			26	25.90	1.174	
11	22.24	3.35 (β)	3.12 (α)	27	16.86	0.638	
12	31.08	1.86 (α)	1.73 (β)	28	14.32	0.838	
13	37.13^c			29	22.34	0.972	
14	41.57^c			30	23.17	0.880	

a-c : signals may be interchanged

Mass spectral data.

3 : EIMS (70 eV), *m/z* (rel. int.) : 324(M^+,20%), 309(100), 281(<1), 239(13), 171(61), 157(75), 145(72), 143(20), 137(28), 131(25), 105(15), 95(32).

4 : EIMS (70 eV), *m/z* (rel. int.) : 376(M^+,100%), 361(38), 333(5), 291(5), 237(4), 223(53), 219(68), 218(50), 209(67), 197(48), 195(29), 184(27), 170(44), 155(48), 137(24), 131(33), 121(26), 109(28), 95(38).

5 : EIMS (70 eV), *m/z* (rel. int.) : 374(M^+,56%), 359(34), 331(<1), 289(3), 235(6), 221(63), 207(100), 195(44), 192(16), 181(16), 179(13), 165(17), 151(6), 137(32), 109(10), 95(22).

Acknowledgments : we thank M. Rautenschlein (Finnigan MAT, Bremen, Germany) for the GC-IRMS analyses, M. Vandenbroucke (Institut Français du Pétrole) for a sample of the Ménat shale and the Centre National de la Recherche Scientifique for financial support. L. Schwark is grateful to the Deutsche Forschungsgemeinschaft and B.J. Keely to the Natural Environment Research Council for financial support.

REFERENCES AND NOTES

1. Lohmann, F.; Trendel, J.M.; Hetru, C.; Albrecht, P. *J. Labelled Compd. Radiopharm.* **1990**, *28*, 377-386.
2. a) Jarolim, V.; Hejno, K.; Hemmert, F.; Sorm, F. *Collect. Czech. Chem. Commun.* **1965**, *30*, 873-879. b) Streibl, M.; Herout, V. In *Organic Geochemistry - Methods and Results*: Eglington, G.; Murphy, M.T.J. Eds.; Springer: Berlin, 1969; pp. 401-424. c) Wolff, G.A.; Trendel, J.M.; Albrecht P. *Tetrahedron* **1989**, *45*, 6721-6728.
3. a) Spyckerelle, C.; Greiner, A.Ch.; Albrecht, P.; Ourisson, G. *J. Chem. Res.* **1977**, (*S*) 330-331, (*M*) 3746-3777. b) Spyckerelle, C.; Greiner, A.Ch.; Albrecht, P.; Ourisson, G. *J. Chem. Res.* **1977**, (*S*) 332-333, (*M*) 3801-3828. c) Corbet, B.; Albrecht, P.; Ourisson, G. *J. Am. Chem. Soc.* **1980**, *102*, 1171-1173. d) Trendel, J.M.; Lohmann, F.; Kintzinger, J.P.; Albrecht, P.; Chiaroni, A.; Riche, C.; Cesario, M.; Guilhem, J.; Pascard, C. *Tetrahedron* **1989**, *45*, 4457-4470.
4. Dastillung, M.; Albrecht, P.; Ourisson, G. *J. Chem. Res.* **1980**, (*S*) 166-167, (*M*) 2325-2352.
5. Schwark, L.; Püttmann, W. *Org. Geochem.* **1990**, *16*, 749-761.

6. Kimble, B.J.; Maxwell, J.R.; Philp, R.P.; Eglinton, G. *Chem. Geol.* **1974**, *14*, 173-198.
7. Budzikiewicz, H.; Wilson, J.M.; Djerassi, C. *J. Am. Chem. Soc.* **1963**, *85*, 3688-3699.
8. Wilkins, A.L.; Bird, P.W.; Jager, P.M. *Magn. Reson. Chem.* **1987**, *25*, 503-507.
9. Albrecht, P.; Ourisson, G. *Science* **1969**, *163*, 1192-1193.
10. Dastillung, M.; Albrecht, P.; Ourisson, G. *J. Chem. Res.* **1980**, (*S*) 168-169, (*M*) 2353-2374.
11. $\delta^{13}C=10^3[(R_x-R_s/R_s)]$, where $R={}^{13}C/{}^{12}C$, x : sample, s : PDB standard and R_s=0.0112372.
12. Hayes, J.M.; Takigiku, R.; Ocampo, R.; Callot, H.J.; Albrecht, P. *Nature* **1987**, *329*, 48-51.
13. Freemann, K.H.; Hayes, J.M.; Trendel, J.M.; Albrecht, P. *Nature* **1990**, *343*, 254-256.
14. J.M. Trendel, unpublished results.
15. Spyckerelle, C. *Constituants aromatiques de sédiments*, Université Louis Pasteur de Strasbourg, France 1975.
16. Loureiro, M.R.B.; Cardoso, J.N. *Org. Geochem.* **1990**, *15*, 351-359.
17. Ohmoto, T.; Ikuse, M. *Phytochemistry* **1970**, *9*, 2137-2148.
18. Arpino, P. *Les lipides de sédiments lacustres éocènes*, Université Louis Pasteur de Strasbourg, France 1973.
19. Ourisson, G.; Albrecht, P.; Rohmer, M. *Trends Biochem. Sci.* **1982**, *7*, 236-239.
20. Gonzalez, A.G.; Barrera, J.B.; Rodrigez P.; Elsa, M. *Z. Naturforsch., C: Biosci.* **1991**, *46*, 12-18.
21. Kamaya, R.; Tanaka, Y.; Hiyama, R.; Ageta, H. *Chem. Pharm. Bull.* **1990**, *38*, 2130-2132.