C(14a)-Homo-26-nor-17α-hopanes, a Novel and Unexpected Series of Molecular Fossils in Biodegraded Petroleum

Jean-Michel Trendel,[a] Roland Graff,[a] Patrick Wehrung,[a] Pierre Albrecht,*[a] Daniel Dessort[b] and Jacques Connan[b]

[a] Institut de Chimie, Université Louis Pasteur, URA 31 du CNRS, 1 rue Blaise Pascal, 67000 Strasbourg, France
[b] Elf-Aquitaine, CST Jean Feger, 64018 Pau Cedex, France

A series (C_{27}–C_{35}) of pentacyclic triterpene hydrocarbons, the C(14a)-homo-26-nor-17α-hopanes, has been characterized in a biodegraded petroleum by gas chromatography–mass spectrometry and NMR structural determination of the C_{29} and C_{30} members; these biological markers reveal an as yet unknown transposition of the hopane skeleton.

Hopane molecular fossils originating from polar constituents of procaryotic organisms[1] are widespread in sediments and petroleum,[2] and thereby count among the most useful biological markers in applied geochemistry for maturity assessment of organic matter and correlation studies.[3] In the alkane fractions, they usually appear in series extending up to C_{35}, under various structural forms which result from the transformations undergone during their sedimentary history: regular hopanes with stereochemical variations at C-17, C-21 and C-22,[4a] rearranged hopanes,[4b] hexahydrobenzo-

hopanes,[4c] 17,21-secohopanes,[4d] 8,14-secohopanes,[4e,f] and also demethylated hopanes, frequently at C-10.[4f,5] The latter, which often occur in geological sources in small quantities and are the most easily detected in biodegraded petroleum, have been definitely identified recently from a heavy oil-impregnated sandstone (Loufika outcrop, Congo).[5]

In the complex alkane mixture of the above sample, gas chromatography–mass spectrometry (GC–MS) studies allowed detection of unknown compounds, the distribution of which could best be studied by using the characteristic m/z 259
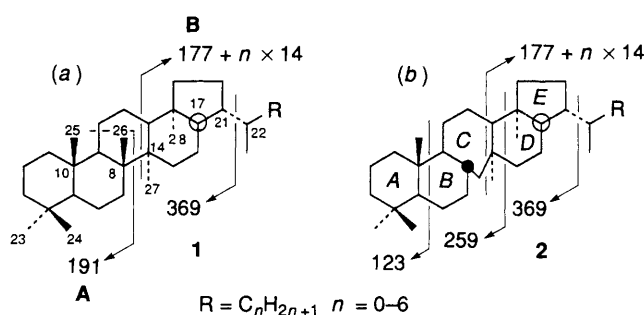
Table 1 ^{1}H and ^{13}C chemical shifts for C(14a)-homo-26-nor-17α-hopane **2** (R = Me) (400 MHz, C_6D_6, 300 K)

C	δ^{13}C	δ^{1}H		C	δ^{13}C	δ^{1}H	
1	39.67	1.63 (β)	0.83 (α)	15	34.74	1.59 (β)	1.04 (α)
2	19.37	~1.50 (β)	~1.42 (α)	16	22.21	1.81 (α)	1.49 (β)
3	42.35	1.39 (β)	1.16 (α)	17	49.39	1.42	
4	33.45			18	46.18		
5	55.48	0.81		19	41.25	1.84	1.16
6	22.71	1.54 (α)	1.25 (β)	20	25.20	1.85	1.44
7	38.75	1.65 (β)	1.07 (α)	21	47.77	1.89	
8	31.89	1.63		22	31.65	1.60	
9	58.90	0.77		23	33.77	0.888 s	
10	38.28			24	22.13	0.844 s	
11	30.98	1.84 (α)	1.05 (β)	25	14.41	0.821 s	
12	23.92	1.43 (β)	~1.26 (α)	27	23.59	0.944 s	
13	45.98	1.24		28	23.63	1.026 s	
14	36.07			29	22.24	0.959 d J 6.8 Hz	
14a	53.95	1.37 (α)	1.03 (β)	30	18.49	0.872 d J 6.8 Hz	

and 369 mass fragmentograms. The latter highlight a series of pentacyclic triterpene structures which range from C_{27} to C_{35} ($M^+ = 370$; $398 + n \times 14$, $n = 0$–6) with two isomers for higher pseudohomologues ($>C_{30}$), thus paralleling that of the regular hopanes. Although their mass spectra† are appreciably different from those of the regular hopanes, their mass fragmentation pattern still indicates a close relationship with the hopane skeleton (Fig. 1). The most striking differences come from the considerably reduced intensities of ions **A** and **B** [Fig. 1(a)] and consequently from the relative enhancement of usually less important ions (M^+, 369, 231, 123) together with the emergence of an ion of m/z 259. As in the case of regular hopanes the major ions **A** and **B** arise from primary cleavage of the particularly weakened C8–C14 bond, these variations indicate important alterations at these positions. An additional clue came from the great similarity which could be observed between the mass spectrum of the unknown C_{30} component and those of the seven-membered ring C 14ζ-serratanes[6] (prominent ions at m/z 259, 231, 191, 123). Hence, we envisaged from these preliminary data that we were dealing with a series of enlarged ring C hopanes.

Conclusive information was obtained from NMR experiments performed on the C_{29} and C_{30} members of the series, after their isolation from the asphalt using repeated reversed-phase HPLC fractionation (for sandstone extraction and extract gross fractionation procedures, see ref. 5).

In the case of the C_{30} product (see Table 1 for ^{1}H and ^{13}C assignments and Fig. 2 for carbon numbering), ^{1}H NMR reveals the presence of two methyl doublets and only five methyl singlets, *i.e.*, the lack of a methyl on a quaternary carbon in comparison with a regular C_{30} hopane. This was further confirmed by ^{13}C ^{1}H broad-band decoupled and DEPT spectra, which disclose the replacement of one methyl group and one quaternary carbon by methylene and methine groups, affording evidence of a ring enlargement involving a methyl group. The basic skeleton was deduced from ^{1}H–^{13}C long-range ($^{2,3}J$) correlations, essentially starting from the methyl groups, as follows. The two methyl doublets (29-Me and 30-Me) are correlated with the same two methines (C-22 and C-21), thus confirming the isopropyl moiety. The geminal methyl groups (23-Me and 24-Me) display remote connections to each other and to the same quaternary (C-4), methylene (C-3) and methine (C-5) carbons. The latter has a connection with a methyl group (25-Me) which also 'sees' quaternary (C-10), methylene (C-1) and methine (C-9) carbons. None of

† *MS data* (Finnigan MAT TSQ70) EI (70 eV), m/z (rel. int.) for **2**; R = H: 398 (M^+, 42%), 383(21), 369(22), 313(3), 299(3), 285(8), 274(7), 259(35), 245(11), 231(21), 219(10), 205(12), 204(21), 193(22), 191(31), 177(95), 163(32), 149(35), 137(48), 123(74), 109(79), 95(100); **2**; R = Me: 412(M^+, 29%), 397(19), 369(39), 327(2), 299(8), 285(5), 274(4), 259(30), 245(11), 231(20), 219(15), 218(15), 205(12), 193(22), 191(84), 177(35), 163(31), 149(44), 137(53), 123(84), 109(59), 95(100).
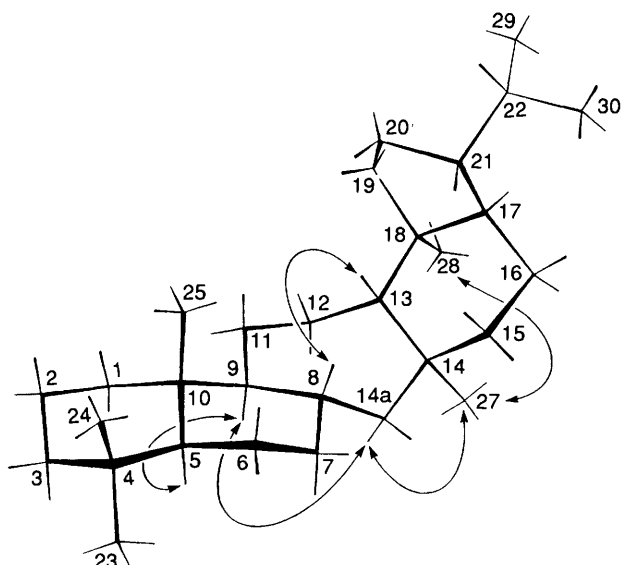

Fig. 1 Mass spectral fragmentations observed for (*a*) regular 17α-hopanes **1** and (*b*) C(14a)-homo-26-nor-17α-hopanes **2**

these carbons, and in particular not the methine (C-9), shows remote connection with any of the two remaining methyl groups, thus demonstrating the absence of the 26-Me at C-8. The two remaining methyl singlets, which have in common only one connection with a methine, must be located at their original position (C-14 and C-18) on the hopane skeleton: indeed, their location at C-13 and C-17 is precluded as 28-Me should have a connection with the same methine (C-21), which correlated with the methyl doublets, and this is not the case. Furthermore, one of the methyl groups (27-Me) shows connections with quaternary (C-14) and methine (C-13) carbons and with two methylene carbons (C-15 and C-14a), one of the latter (C-14a) resulting from the ring enlargement implicating the methyl group previously vicinal to 27-Me. In addition, the proton which later on was assigned as being the 14aα one, has clear connectivities with C-27, C-15, C-14, C-8 and particularly with C-9. All these features prove that ring enlargement involves the C(8)–C(14) bond, to the detriment of the 26-Me.

Further 2D NMR experiments (phase-sensitive ^{1}H–^{1}H COSY and NOESY, one-bond ^{1}H–^{13}C correlations) were carried out, and the complete set of NMR experiments was repeated with 17α-hopane **1** (R = Me) for comparison purposes. These fully confirmed the preceding structural considerations (in particular the ^{1}H–^{1}H COSY connectivity pattern observed: H-14as–H-8–H-9–H-11s) and allowed the stereochemistry of the molecule to be determined (Fig. 2). In brief, the latter must have the thermodynamically most stable 17αH,21βH configuration, since all chemical shifts (^{1}H and ^{13}C) concerning positions of rings C and E (with its isopropyl substituent) that should not be notably affected by ring C enlargement (*i.e.* 16, 17, 20, 21, 22, 29 and 30), are indeed very close to those of 17α-hopane [Δδ^{13}C < 0.3 ppm; Δδ^{1}H < 0.1 ppm except for one H-20 (β ?) for which Δδ^{1}H ≈ 0.12 ppm]. The 8β-H configuration at C-8 follows from the NOEs (nuclear Overhauser effects) observed between H-8–H-13 and H-14aα–H-9, the expected cross-peak between H-8 and 25-Me being almost completely obscured by the intense H-1α–H-1β correlation, and from the ^3J coupling constant

Fig. 2 3D drawing of C(14a)-homo-26-nor-17α-hopane **2** (R = Me) showing some essential NOE connectivities

measurement for H-9, the signal of which appears as a broad doublet of triplet (J ca. 11, 11, 6 Hz, plus a long-range coupling). Molecular-mechanics data (SYBYL software, Tripos Associates Inc., St Louis, Missouri), obtained by modelling the most probable conformers for the two configurations at C-8, indicated an energy minimum for the 8β-H-structure with the conformation shown on Fig. 2, which agrees with NMR data.

The ^{1}H NMR spectrum‡ of the C_{29} member also shows the lack of a methyl group on a quaternary carbon as compared with a regular C_{29} hopane, and methyl singlets with chemical shifts quasi identical to those of the previously identified C_{30} compound, thus confirming the presence of a homologous series; the structures of the C_{27} and higher members ($>C_{30}$) are inferred from their mass spectra and their gas chromatography retention times.

Until now, this novel hopanoid series could be the most easily detected in biodegraded samples (the mass spectrum of an 'unknown compound' that fits very well with that of our C_{30} structure has been previously reported,[7] also from a severely biodegraded bitumen), generally accompanied by the 8,14-secohopanes. Hence, as for the 25-norhopanes and 8,14-secohopanes, the question is open as to whether the C-homo-26-norhopanes are generated during biodegradation in the reservoir, or are concentrated by selective removal of more readily degradable polycyclic alkanes, e.g. the 17α-hopanes. Recently, reliable arguments have been put forward, at least for the 25-norhopanes, in favour of both hypotheses.[8,9]

Concerning the formation of the C-homo-26-norhopanes, a direct production from bacteria involving a pathway similar to that which leads to the serratane backbone appears quite unlikely, although it cannot be completely ruled out (serratene, an unsaturated hydrocarbon, has been found in ferns which are known to contain hopane or rearranged hopane skeletons lacking oxygen at the C-3 position[10]). Otherwise, a possibility of formation could for instance lie in a biological oxidation process of the 26-Me, either in the reservoir or at an early stage of diagenesis, leading to an alcohol function. The latter could be eliminated under acidic conditions, in a more or less concerted way, by migration of the C-8–C-14 bond. However, abiotic geochemical processes operating on alkanes could probably also promote (most likely via free radical intermediates) such a transposition, especially as the over-

stretched C-8–C-14 bond appears a privileged site of alteration, as shown by the cooccurrence in geological samples of the 8,14-secohopanes. It also appears from GC–MS data that at least a second less important series of C-homonorhopanes is present in our sample, which clearly results from a ring enlargement involving the 27-Me instead of the 26-Me (the typical fragments at m/z 259 and 231 in the first series are now replaced by ions 217, 245 + $n \times 14$ and 189, 217 + $n \times 14$, respectively; in comparison with the first series, the fragment at m/z 123 is relatively more important, whereas the ion m/z 369 becomes very weak). In addition, a C_{30} compound which most likely is a C-homonorgammacerane (= serratane skeleton) has also been detected. This apparent lack of specificity with regard to the methyl group implicated in the transposition, as well as to the type of skeleton, argues against an original production of these compounds, which obviously arise from secondary transformation processes operating in the subsurface. From the thermodynamic view point, it should be noted that the C-homonorhopanes appear to be substantially more stable than the corresponding 17α-hopanes (SYBYL software).

Be that as it may, these C-ring enlarged hopanes may, like the 25-norhopanes, turn out to be useful biomarkers in petroleum geochemistry, particularly as maturity[11] and palaeobiodegradation[9] parameters.

References

1 G. Ourisson and M. Rohmer, *Acc. Chem. Res.*, 1992, **25**, 403.
2 G. Ourisson and P. Albrecht, *Acc. Chem. Res.*, 1992, **25**, 398.
3 W. K. Seifert and J. M. Moldowan, *Geochim. Cosmochim. Acta*, 1978, **42**, 77; A. S. MacKenzie, in *Advances in Petroleum Geochemistry*, ed. J. Brooks and D. Welte, Academic Press, London, 1984, vol. 1, p. 115; K. E. Peters and J. M. Moldowan, *The Biomarker Guide*, Prentice Hall, Englewood Cliffs, 1993.
4 *E.g.* (*a*) A. Van Dorsselaer, A. Ensminger, C. Spyckerelle, M. Dastillung, O. Sieskind, P. Arpino, P. Albrecht, G. Ourisson, P. W. Brooks, S. J. Gaskell, B. J. Kimble, R. P. Philp, J. R. Maxwell and G. Eglinton, *Tetrahedron Lett*, 1974, 1349; A. Ensminger, A. Van Dorsselaer, C. Spyckerelle, P. Albrecht and G. Ourisson, in *Advances in Organic Geochemistry 1973*, ed. B. Tissot and F. Bienner, Technip, Paris, 1974, p. 245; (*b*) J. M. Moldowan, F. J. Fago, R. M. K. Carlson, D. C. Young, G. Van Duyne, J. Clardy, M. Schoell, C. T. Pillinger and D. S. Watt, *Geochim. Cosmochim. Acta*, 1991, **55**, 3333; (*c*) J. Connan and D. Dessort, *Org. Geochem.*, 1987, **11**, 103; (*d*) J. M. Trendel, A. Restlé, J. Connan and P. Albrecht, *J. Chem. Soc., Chem. Commun.*, 1982, 304; (*e*) J. M. Schmitter, W. Sucrow and P. J. Arpino, *Geochim. Cosmochim. Acta*, 1982, **46**, 2345; (*f*) J. Rullkötter and D. Wendisch, *Geochim. Cosmochim. Acta*, 1982, **46**, 1545.
5 J. M. Trendel, J. Guilhem, P. Crisp, D. Repeta, J. Connan and P. Albrecht, *J. Chem. Soc., Chem. Commun.*, 1990, 424.
6 B. J. Kimble, J. R. Maxwell, R. P. Philp and G. Eglinton, *Chem. Geol.*, 1974, **14**, 173.
7 M. G. Fowler, L. R. Snowdon, P. W. Brooks and T. S. Hamilton, *Org. Geochem.*, 1988, **13**, 715.
8 K. E. Peters and J. M. Moldowan, *Org. Geochem.*, 1991, **17**, 47.
9 P. Chosson, J. Connan, D. Dessort and C. Lanau, in *Biological Markers in Sediments and Petroleum*, ed. J. M. Moldowan, P. Albrecht and R. P. Philp, Prentice Hall, Englewood Cliffs, 1992, p. 320; P. Blanc and J. Connan, *Org. Geochem.*, 1992, **18**, 813.
10 G. Berti, F. Bottari, A. Marsili, I. Morelli and A. Mandelbaum, *J. Chem. Soc., Chem. Commun.*, 1967, 50; H. Ageta, K. Shiojima and K. Masuda, *Chem. Pharm. Bull.*, 1982, **30**, 2272; Y. Arai, M. Yamaide, S. Yamazaki and H. Ageta, *Phytochemistry*, 1991, **30**, 3369.
11 J. K. Volkman, R. Alexander, R. I. Kagi and G. W. Woodhouse, *Geochim. Cosmochim. Acta*, 1983, **47**, 785.

‡ *^{1}H NMR data* for **2**: R = H (Bruker AM-400) 400 MHz, C_6D_6, 300 K; δ (SiMe₄): 0.819 (3H, s, H-25), 0.845 (3H, s, H-24), 0.889 (3H, s, H-23), 0.936 (3H, s, H-27), 0.952 (3H, t, J 7.4 Hz, H-29), 1.026 (3H, s, H-28).