**Tetrahedron Letters No. 14, pp 1349 - 1352, 1974. Pergamon Press. Printed in Great Britain.**

# Degraded and extended hopane derivatives ($C_{27}$ to $C_{35}$) as ubiquitous geochemical markers

A. Van Dorsselaer, A. Ensminger, C. Spyckerelle, M. Dastillung,
O. Sieskind, P. Arpino, P. Albrecht and G. Ourisson

Laboratoire de Chimie Organique des Substances Naturelles,
Associé au CNRS, Institut de Chimie, Université Louis Pasteur,
1 rue Blaise Pascal, 67008 Strasbourg, France.

and

P.W. Brooks, S.J.Gaskell, B.J. Kimble, R.P. Philp, J.R. Maxwell
and G. Eglinton

Organic Geochemistry Unit, School of Chemistry, University of
Bristol, England.

(**Received in UK 1 February 1974; accepted for publication 20 February 1974**)

Of the wide variety of triterpene families which occur in the contemporary biosphere[1], only a few, usually as the parent alkanes or dehydrogenated derivatives, have been recognized among the organic compounds isolated from geological sources. Thus oleanane and rearranged oleananes occur in crude oils[2], 3-desoxyallobetulin derivatives in a crude oil[3] and in peats and brown coals[4], (17αH) hopane[5] and gammacerane[6] in Green River oil shale, isoarborinol and the corresponding ketone in Messel oil shale[7], (17αH) hopane and a rearranged trisnorhopane in crude oils[8], and hopane **1** and homohopane **4** in Messel shale[9]. These complex molecules can be useful in tracing the chemical reactions which occur in the geological environment during the decay of biological organic materials ; ultimately, after millions of years, these reactions lead to the fossil energy sources such as coal, petroleum or natural gas.

*We note that pentacyclic triterpenes of the hopane family ($C_{27}$ - $C_{35}$) are, in fact, important biological markers which are found in almost every sample of sedimentary organic matter we have investigated to date. Their ubiquitous occurrence, taken with the fact that they are often abundant components of the organic extracts, implies previously unsuspected similarities in their origin and in the processes leading to their formation.*

The hydrocarbon, ketone, alcohol and acid hopane derivatives so far identified are shown in Table 1, together with some other hopane derivatives whose structures are inferred from their glc-ms mass spectra and glc retention times or in one case, glc-ms mass spectrum and uv spectrum. They include members of the true hopane series (17βH, 21βH), the moretane series (17βH, 21αH) and the (17αH, 21βH)-hopane series. The latter series has not yet been found in living organisms.

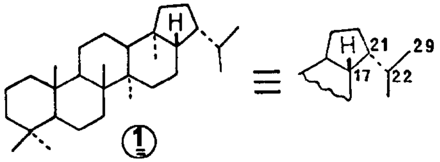
Occurrences of these compounds are listed in Table 2. In most cases the hopane derivatives are important, or even major, components in the corresponding chromatographic fractions from the various samples ; for example, acid 7 is the most abundant individual organic compound in the Messel shale at a concentration of 100-150 ppm in the rock. In most cases, no derivative of any other triterpane series is present in significant concentrations in the same fractions. It is significant that the hopane derivatives are also important constituents of very young sediments ; for example acid 7 is the major free acid in Rostherne Mere and Grasmere bottom sediments. It is reasonable to assume that these compounds are syngenetic (i.e. same age) with the sediments. However, any sediment laid down may contain a proportion of much older organic matter from erosion of pre-existing sedimentary formations.

The significance of the wide occurrence of derivatives of *one* particular triterpene skeletal type is noteworthy, especially one not generally considered as being widespread in living organisms. We attach great importance to the fact that the series extends up to the $C_{35}$ compound, 29-pentylhopane 24 because a closely related $C_{35}$ skeleton has recently been proposed (but not unambiguously proved) for bacteriohopane, a tetrahydroxy derivative of which has been isolated from *Acetobacter xylinum*[10]. 3-Desoxy $C_{29}$ and $C_{30}$ hopane derivatives are found in several contemporary organisms, eg: ferns[11], mosses[12], lichens[13], several cyanophyceae[14] and chemotrophic bacteria[15]. A $C_{31}$ pentacyclic triterpane of unknown structure has also been isolated in trace amounts from a bacterium and $^{14}C$ labelling studies with the same organism have confirmed its ability to synthesise hopane derivatives[16].

We postulate (and will attempt to verify) that the presence of abundant hopane derivatives is a sign of the occurrence of active microbial processes during sedimentation, and that their exact nature in a given sediment reflects, in a decipherable manner, the conditions of sedimentation and geological maturation. At this stage, we cannot conclude whether our results indicate the *general* intervention of *Acetobacter* spp., or indicate the *general* occurrence of hopane derivatives in the prevalent microorganisms of the sedimentary environment. The total global amount of hopane derivatives in sediments can now be conservatively estimated at $10^8 - 10^9$ tonnes.

We thank ELF-ERAP, the Natural Environment Research Council (GR/3/655), the National Aeronautics and Space Administration (subcontract from NGL 05-003-003) and the Petroleum Research Fund (PRF-6350-AC2) for financial support.

## Table 1. Hopane derivatives from geological sources*.

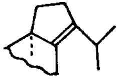


①

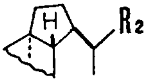
②

③ $R_1 = H$ ④ $R_1 = Et$ 5 $R_1 = n\text{-}C_3H_7$ 6 $R_1 = CH_2CO_2H$

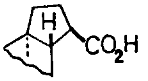
⑦ $R_1 = (CH_2)_2CO_2H$ 8 $R_1 = (CH_2)_3CO_2H$ 9 $R_1 = (CH_2)_3OH$



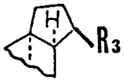
10



⑪ $R_2 = H$ ⑫ $R_2 = Me$ ⑬ $R_2 = Et$ 14 $R_2 = O$

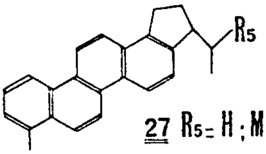


⑮

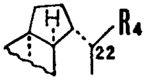


⑯ $R_3 = H$

17 $R_3 = O$



27 $R_5 = H; Me$



⑱ $R_4 = H$ ⑲ $R_4 = Me$ ⑳ $R_4 = Et$ ㉑ $R_4 = n\text{-}C_3H_7$ ** 22 $R_4 = n\text{-}C_4H_9$

23 $R_4 = n\text{-}C_5H_{11}$ ㉔ $R_4 = n\text{-}C_6H_{13}$ 25 $R_4 = CH_2CO_2H$ ** 26 $R_4 = (CH_2)_2CO_2H$

---

* Circled numbers: structures substantiated by glc-ms and glc (coinjection on capillary columns) with authenticated synthesised samples. The other structures are tentative (glc-ms and UV), no direct comparison with authentic samples.

** Compounds 20-26 are often present as a mixture of 22 S and 22 R epimers.

### References

1) T.G. HALSALL and R.T. APLIN, Fortsch. Chem. Org. Naturst., 22, 153 (1964).

2) I.R. HILLS, G.W. SMITH and E.V. WHITEHEAD, Nature (London), 219, 243 (1968).

3) D.H.R. BARTON, W. CARRUTHERS and K.H. OVERTON, J. Chem. Soc., 788 (1956).

4) M. STREIBL and V. HEROUT, in Organic Geochemistry, G. Eglinton and M.T.J. Murphy, Ed., Springer, Berlin, 402 (1969).

5) B. BALOGH, D.M. WILSON, P.C. CHRISTIANSEN and A.L. BURLINGAME, Nature (London), 242, 603 (1973).

6) I.R. HILLS, E.V. WHITEHEAD, D.E. ANDERS, J.J. CUMMINS and W.E. ROBINSON, Chem. Commun., 752 (1966).

7) P. ALBRECHT and G. OURISSON, Science, 163, 1192 (1969).

8) E.V. WHITEHEAD, Chem. and Ind. (London), 1116 (1971).

9) A. ENSMINGER, P. ALBRECHT, G. OURISSON, B.J. KIMBLE, J.R. MAXWELL and G. EGLINTON, Tetrahedron Letters, 3861 (1972).

10) H.J. FÖRSTER, K. BIEMANN, W.G. HAIGH, N.H. TATTRIE and J.A. COLVIN, Biochem. J., 135, 133 (1973).

Table 2. Occurrence of hopane derivatives in sediments and crude oils.

| *Sample* | *Age* | *Origin* | *Compounds identified* |
|---|---|---|---|
| Esthwaite Water, Rostherne Mere and Grasmere sediments (England) | Recent (0-50 yrs) | Lacustrine | 2,4,7*,16,18-20 |
| Atlantic Ocean sediment (Mauritania) | Recent ($< 10^6$ yrs) | Marine | 6,7,8 |
| Baltic sea sediments | Recent ($< 10^5$ yrs) | Marine | 1-4,8,9,11,16,18-20 |
| Gulf of Mexico sediment | Pleistocene ($10^6$ yrs) | Marine | 7*[18] |
| Yallourn lignite (Australia) | Miocene ($25 \times 10^6$ yrs) | Terrestrial | 20,21 |
| Messel oil shale[9,17] (Germany) | Eocene ($50 \times 10^6$ yrs) | Lacustrine | 1-8,10,11,14,16-20,<br>26,27 |
| Bouxwiller shale (France) | Eocene ($50 \times 10^6$ yrs) | Lacustrine | 1-7,14,17 |
| Ménat shale (France) | Eocene ($50 \times 10^6$ yrs) | Lacustrine | 1-4 |
| Stockstadt shale (Germany) | Eocene ($50 \times 10^6$ yrs) | Lacustrine | 16,18-20 |
| Green River oil shale (U.S.A.) | Eocene ($50 \times 10^6$ yrs) | Lacustrine | 18,19[5],20 |
| Gaboon shale | Cretaceous($120 \times 10^6$) | Marine | 2,4,7,11-13,15,16,18-21 |
| Atlantic Ocean sediment (Bermuda) | Cretaceous($140 \times 10^6$) | Marine | 7*[18] |
| Aquitaine basin sediments (France) | Cretaceous and Jurassic($150-200 \times 10^6$ yrs) | Marine | 11-13,16,18-24 |
| Mediano oil shale (Spain) | Jurassic($180 \times 10^6$ yrs) | Marine | 16,18-24 |
| Lorraine coal (France) | Carboniferous ($275 \times 10^6$ yrs) | Terrestrial | 11-13,15,16,18-26 |
| Crude oils from : | | | |
| Aquitaine basin (France) | Cretaceous and Jurassic($150-200 \times 10^6$ yrs) | Marine | 11-13,16,18-24 |
| Nigeria, Lybia, Iran Kuwait, Texas | $>25 \times 10^6$ yrs | Marine | 19[8] |

* Comparison with mass spectrum of authentic 7 alone.

11) G. BERTI and F. BOTTARI, in Progress in Phytochemistry, L. Reinhold and V. Lipschitz, Ed., Interscience, London, vol. 1, 589 (1968).

12) A. MARSILI, I. MORELLI, C. BERNARDINI and M. PACCHIANI, Phytochemistry, 11, 2003 (1972).

13) I. YOSHIOKA, M. YAMAKI and I. KITASAWA, Chem. Pharm. Bull. (Tokyo), 14, 804 (1966).

14) E. GELPI, H. SCHNEIDER, J. MANN and J. ORO, Phytochemistry, 9, 603 (1970).

15) C.W. BIRD, J.M. LYNCH, S.J. PIRT and W.W. REID, Tetrahedron Letters, 3189 (1971).

16) M. DE ROSA, A. GAMBACORTA, L. MINALE and J.D. BU'LOCK, Phytochemistry, 12, 117 (1973).

17) B.J. KIMBLE, J.R. MAXWELL, R.P. PHILP, G. EGLINTON, P. ALBRECHT, A. ENSMINGER, P. ARPINO and G. OURISSON, Geochim. Cosmochim. Acta, in press.

18) B.R. SIMONEIT, in Initial Reports of DSDP, O. Weser, D. Ross et al., Eds., vol. XXIII, in press.