Tetrahedron Letters No. 43, pp 4163 - 4166.
© Pergamon Press Ltd. 1978. Printed in Great Britain.

0040-4039/78/1022-4163$02.00/0

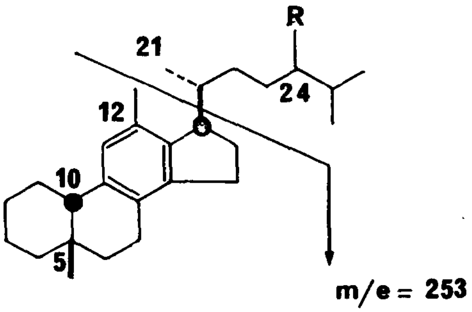
AROMATIC HYDROCARBONS FROM GEOLOGICAL SOURCES. VI
NEW AROMATIC STEROID DERIVATIVES IN SEDIMENTS AND CRUDE OILS

J. SCHAEFLE, B. LUDWIG, P. ALBRECHT and G. OURISSON

Institut de Chimie, Laboratoire de Chimie Organique des Substances Naturelles,
Associé au CNRS, 1, rue Blaise Pascal, 67000 Strasbourg, France.

Precise structural elucidation of molecular markers of the geological environment can give useful clues as to the origin and alteration processes of the organic matter in sediments and crude oils[1,2]. Especially the steroid skeleton, which has been extensively studied over the past years, undergoes various degradation pathways during sedimentation. Some of them already take place in the recent sediment stage under microbial influence[3]. Further changes occur during maturation due to the catalytic effect of clay minerals[4]. Several authors have noticed the presence of partially aromatized hydrocarbons related to steroids in geological samples where they often appear as major constituents of the aromatic fraction[5-8]. Until now, none of these compounds had been identified, mainly due to the complexity of the mixtures involved.

We would now like to report the first identification of one member, **1**, of a series of monoaromatic steroid derivatives which occurs in several sediments and crude oils of various origins and ages. The samples we have studied (marine oil shales of the lower Toarcian formation of the Paris basin, 180 x 10^6 yrs ; Green River oils (3 samples) ; crude oils from Canada, Iran, Venezuela, Lybia) are only a limited selection ; from the results described by Seifert and Moldowan[8], it appears probable that the range of distribution of the derivatives now identified is quite large.



1 R=H **1'** R=CH_3 **1"** R=C_2H_5

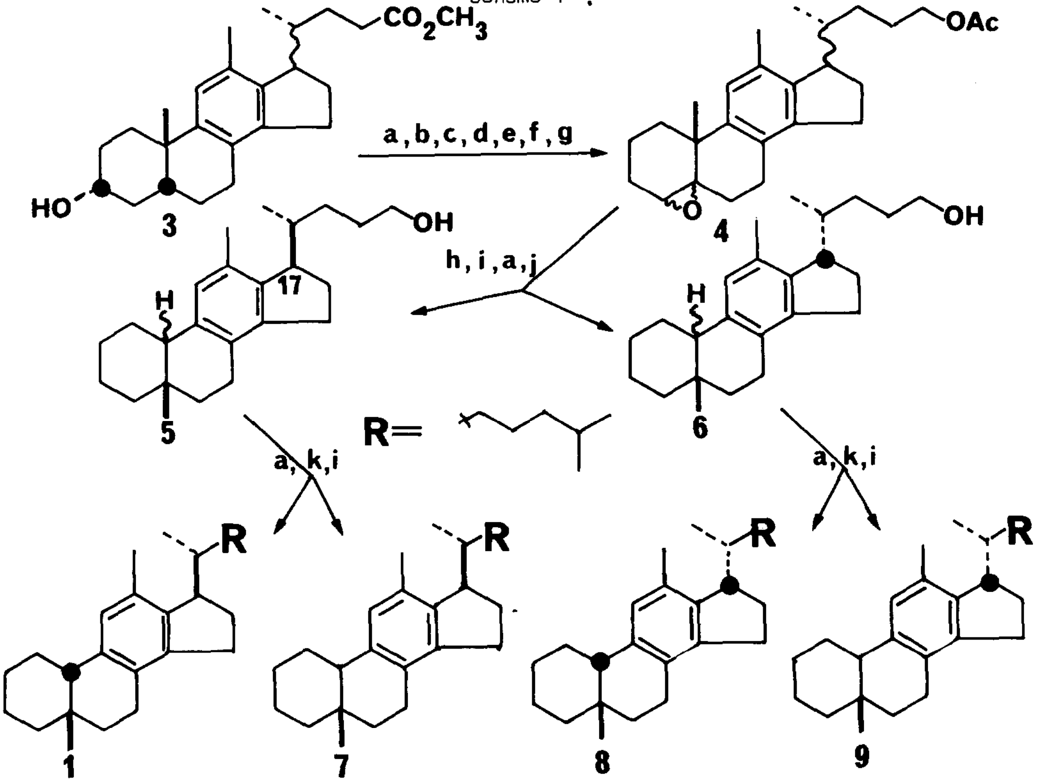
The aromatic hydrocarbons of the geological samples were obtained by extraction and several chromatographic separations as previously described[9,10]. The mono-+ di-aromatic fraction was separated by high performance liquid chromatography ($10\mu SiO_2$-hexane) and usually appears as a very complex mixture requiring the use of computerized gas chromatography- mass spectrometry (GC-MS LKB 9000 S and Finnegan 1500 ; 70 eV) for further analysis. The mass spectra of the major monoaromatic steroids from sediments and crude oils show a very simple fragmentation pattern with the base peak at $m/e = 253$. The distribution of this class of compounds in the mono-+ di-aromatic fraction can be neatly obtained by monitoring this 253 fragment ion, a method which has been sucessfully applied for correlations among crude oils[8]. The parent ions are weak, less than 4 % of the base peak, and often barely discernible from the background. However in favourable cases, good evidence for parent ions corresponding to C_{27}, C_{28} and C_{29} compounds at 366, 380 and 394 amu could be obtained.

These data suggested a steroid skeleton containing an aromatic C ring with the side chain in an easily cleavable benzylic position. Further hints to the structural elucidation of these compounds were obtained from the presence of backbone rearranged sterenes as major components in some of the immature samples which we have analyzed[11], as well as from the results of other authors on the aromatization of ring C in steroids, albeit using stepwise chemical modifications. On this tenuous basis we oriented our synthesis towards a semi-rearranged steroid skeleton with an aromatic C ring bearing a methyl group at position 12. Compound **3** was obtained from cholic acid by the method of MENEY et al.[12] as a mixture of isomers at position 17 (α-chain major) which were separable by SiO_2 chromatography in a later stage as the alcohols **5** and **6** and could be distinguished on the basis of the chemical shift of the 21-methyl group in NMR[12]. The migration of the methyl group from position 10 to 5 was realized by treating the epoxide **4** with $BF_3.Et_2O$ in benzene[13] (scheme 1). The 10α and 10β isomers could be separated on the final mixtures by $SiO_2/AgNO_3$ thin layer chromatography (n-hexane eluent). They were distinguished on the basis of the chemical shift of the aromatic proton appearing further downfield in the trans isomer in NMR[14]. They are oily. Compounds **1**, **7**, **8** and **9** display nearly identical mass spectra (**1** : $M^+ = 366$ (4 %) ; $m/e = 253$ (100 %), 143 (23 %)), but are easily separable by gas chromatography. Only compound **1** (NMR δ $CDCl_3$ 0.69 (3H,s) ; 0.73 (6H, d, J = 6.6 Hz) ; 0.96 (3H, d, H = 6.6 Hz), 2.21 (3H, s), 6.81 (1H, s) coeluted with a major component of the geological samples on the 253 mass fragmentogram obtained in GC-MS (SE 52, 15 m x 0.30 mm, film thickness 0.15μ, open tubular glass column ; Dexsil 1 %, 2.5 x 2 mm i.d. glass column).

It can be assumed at this stage that component **1** is representative of a new class of geochemical substances which arise through dismutation reactions during the backbone rearrangement of sterenes, a process occurring in sediments at an early stage of maturation under the catalytic influence of clay minerals[4,11].

Further work aiming at the identification of other members of the series, as well as the more aromatized steroid derivatives is,in progress. It is already clear that the homologous C_{28} and C_{29} hydrocarbons must be **1'** and **1"**.

Scheme 1[15].



a. CrO$_3$, Py ; b. PTAB ; c. Li$_2$CO$_3$, LiBr·d.LiAH$_4$; e. Ac$_2$O, Py ; f. Li, EtNH$_2$[16] ; g.m-Cl(C$_6$H$_4$CO$_3$H)
h. BF$_3$.Et$_2$O ; i. H$_2$/Pd , j. N$_2$H$_4$, KOH,k. (CH$_3$)$_2$C = P(C$_6$H$_5$)$_3$, Et$_2$O

Acknowledgements

We thank ELF-Aquitaine and the Institut Français du Pétrole for financial support and supply of samples, Rhône Poulenc for a doctoral fellowship (J.S.), Dr. W. GIGER*, EAWAG, Dübendorf, for use of his mass spectral facilities and Dr. K.P. CHENG for helpful discussions.

REFERENCES

1) P. ALBRECHT and G. OURISSON, Angew. Chem. Int. Edn, 10, 209 (1971).
2) B.M. DIDYK, B.R.T. SIMONEIT, S.C. BRASSELL and G. EGLINTON, Nature, 272, 216 (1978).
3) M. DASTILLUNG and P. ALBRECHT, Nature, 269, 678 (1977) and references therein.
4) A. ENSMINGER, G. JOLY and P. ALBRECHT, Tetrahedron Letters, 1575 (1978).
5) B. TISSOT, Y. CALIFET-DEBYSER, G. DEROO and J.L. OUDIN, Bull. Amer. Assoc. Petrol. Geol., 55, 2177 (1971).
6) B. TISSOT, J. ESPITALIE, G. DEROO, C. TEMPERE and D. JONATHAN, in "Advances in Organic Geochemistry 1973", ed. B. TISSOT and F. BIENNER, Technip, Paris, 315 (1974).

7) C. SPYCKERELLE, Thèse de Doctorat d'Etat, Université Louis Pasteur, Strasbourg (1975).

8) W.K. SEIFERT and J.M. MOLDOWAN, Geochim. Cosmochim. Acta, 42, 77 (1978).

9) A.Ch. GREINER, C. SPYCKERELLE and P. ALBRECHT, Tetrahedron, 32, 257 (1976).

10) C. SPYCKERELLE, A. Ch. GREINER, P. ALBRECHT and G. OURISSON, J. Chem. Res. (S) 332, (M) 3801 (1977).

11) I. RUBINSTEIN, O. SIESKIND and P. ALBRECHT, JCS Perkin I, 1833 (1975).

12) J. MENEY, YOUNG-HO KIM, R. STEVENSON and T.N. MARGULIS, Tetrahedron, 29, 21 (1973).

13) W.G. DAUBEN, G. AHLGREN, T.J. LEITEREG, W.C. SCHWARZEL, M. YOSHIOKO, J. Amer. Chem. Soc., 94, 8593 (1972).

14) W.NAGATA, T. TERASAWA and K. TORI, J. Amer. Chem. Soc., 86, 3746 (1964).

15) All compounds were characterized by spectroscopic methods (IR, NMR, MS), elemental analysis or high resolution MS ; their purity was checked by GC. The yields of the reactions ranged between 50 and 95 %.

16) A.S. HALLWORTH, H.P. HENBEST and T.I. WRIGLEY, J. Chem. Soc. 1969 (1957).

(Received in UK 15 August 1978; accepted for publication 4 September 1978)