The Structures of Tricyclic Terpenoid Carboxylic Acids and their Parent Alkanes in the Alberta Oil Sands

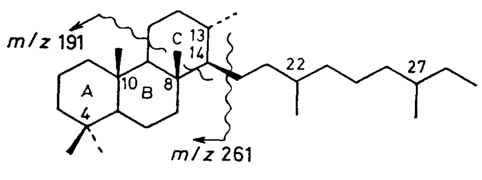
Terry D. Cyr and Otto P. Strausz*

Department of Chemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2G2

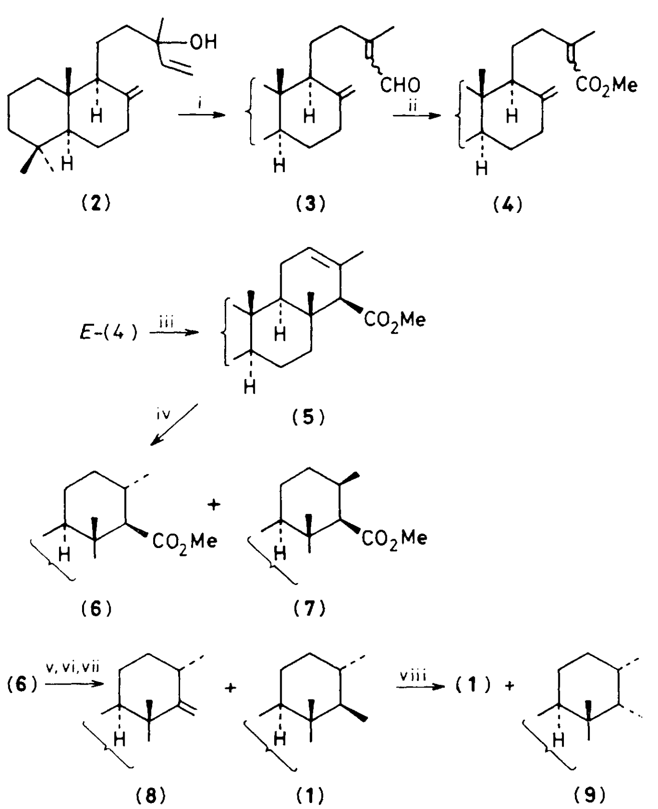
A series of novel tricyclic terpenoid carboxylic acids were isolated from several Alberta oil sand deposits; their structures and those of a related series of tricyclic terpenoid alkanes were elucidated *via* an unambiguous synthesis from manool.

Recently, there has been a surge of interest in the structures,[1,2] distribution,[3,4] and source[5] material of tricyclic terpene hydrocarbons, which are present in many petroleums and sediments. These compounds, which range from C_{19} to C_{30},[3] are usually identified by their characteristic m/z 191 fragment peak[6] using capillary g.c.–mass spectrometry (g.c.–m.s.).[7] The most abundant member (C_{23}) has been assigned the structure 18,19-dinor-13βH,14αH-cheilanthane.[3] Their biological origin is not well understood, but probably involves the cyclization[8,9] of a regular hexaprenoid,[3] its 2,3-oxide, or its 1-ol.[2]

These terpene hydrocarbons have already been used in oil–source rock and oil–oil correlations,[10] although their detailed structures were not known. In order to be useful as thermal maturity indicators one must determine the origin of these compounds.[11] One likely geochemical precursor for the tricyclic alkanes could be the corresponding tricyclic carboxylic acids,[12] some of which had been found prior to this study but not fully characterized.[13]

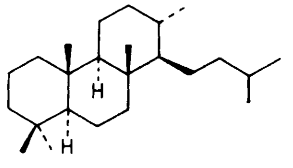


The carboxylic acids were obtained as their methyl esters by the basic hydrolysis, extraction, and then diazomethane methylation of a sample of bitumen from the Athabasca oil sand deposit[14] of Alberta. Capillary g.c.–m.s. analysis of the crude ester mixture indicated that at least four tricyclic methyl esters were present: C_{20}, C_{21}, and a pair of C_{24} diastereoisomers. A concentrate of the esters was obtained by column chromatography on silica gel, then alumina, to obtain a mixture of saturated monocarboxylic methyl esters, followed by Sephadex LH-20 to remove the material outside the molecular weight range of interest, and finally reversed-phase C_{18} bonded silica gel to remove the unresolved complex mixture. Capillary g.c.–m.s. of the tricyclic ester concentrate revealed that the major series consisted of C_{20}, C_{21}, C_{22}, C_{23}, diastereoisomeric C_{24}, C_{25}, and C_{26} esters. The C_{21} and the pair of C_{24} diastereoisomeric esters comprised about 75% of the series.

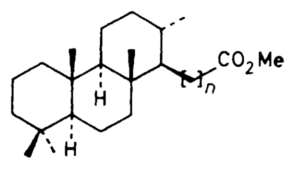


Scheme 1. (i) $H_2Cr_2O_7$, H_2O, H_2SO_4, toluene, AcOH, (ref. 16) 80% (**3**), E:Z = 2.5:1; (ii) MnO_2, NaCN, MeOH, (ref. 17) 95% (**4**); (iii) HCO_2H, 100 °C, (ref. 18), 1 h, 69% (**5**); (iv) 5% Pd–C, pentane, H_2 (1 atm), (ref. 19) 64% (**6**), 36% (**7**); (v) $LiAlH_4$, tetrahydrofuran, 66 °C, 5 h, 75% conversion; (vi) $MeSO_2Cl$, pyridine, 0 to 22 °C, 3 h, 83% conversion; (vii) NaI, Zn, hexamethylphosphoric triamide (ref. 20), 100 °C, 1 h, 88% conversion, (**8**):(**1**) = 95:5; (viii) 5% Pd–C, pentane, H_2 (1 atm), 92% (**1**), 8% (**9**).

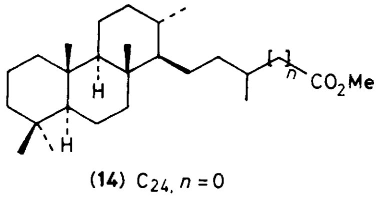
A portion of the crude ester mixture was reduced to an alcohol mixture ($LiAlH_4$, tetrahydrofuran, heat, 5 h, N_2). The methanesulphonates of these alcohols were reduced to the alkanes or monodeuterioalkanes with $LiAlH_4$ or $LiAlD_4$, respectively. The terpenoid alkanes were then concentrated from the unresolved complex mixture in one step.[15] The C_{21} to C_{26} tricyclic alkanes thus obtained and the corresponding native tricyclic alkanes, isolated from the saturate fraction of the bitumen, were identical, as demonstrated by their co-elution on capillary g.c. (SE-30 and Superox-4) and by their very similar mass spectral fragmentation patterns. The mass spectra of the deuteriated tricyclic alkanes indicated that the carboxy group was associated with the C ring since the intensities of the m/z 191 fragment peaks from rings A plus B were not affected by deuteriation. Furthermore, the carboxy group was bonded to an alkyl side chain rather than directly to the C ring since the m/z 261 fragment peaks, from the loss of the C-14 side chain, remained unchanged. The mass spectra (70 eV) of the C_{24} tricyclic methyl esters, alkane, and monodeuterioalkane were as follows: *ester:* 376(M^+, 10%), 361(17), 345(2), 192(19), 191(100), 177(8), 163(10), 149(6), 137(16), 123(23), 109(24), 95(36), 81(27), 69(29), and 55(23); *alkane:* hydride reduction: 332(M^+, 14%), 317(13), 261(5), 247(2), 219(3), 205(2), 192(19), 191(100), 163(3), 137(22), 123(33), 109(31), 95(42), 81(35), 69(54), and 55(41); *monodeuterioalkane:* deuteride reduction: 333(M^+, 20%), 318(18), 261(6), 248(3), 220(7), 205(3), 192(18), 191(100), 163(6), 137(24), 123(27), 109(31), 95(50), 81(3), 69(48), and 55(32).



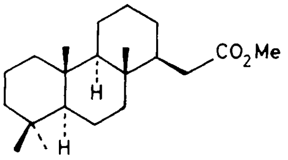
(**10**)



(**11**) C_{21}, n = 1
(**12**) C_{22}, n = 2
(**13**) C_{23}, n = 3



(**14**) C_{24}, n = 0
(**15**) C_{25}, n = 1
(**16**) C_{26}, n = 2



(**17**)

The C_{20} tricyclic alkane (**1**) was synthesized by an unambiguous route from manool (**2**), as shown in Scheme 1. Compound (**1**) was identical to the C_{20} tricyclic alkane present in the saturate fraction of the Athabasca bitumen, as shown by coinjection on capillary g.c. (SE-30 and Superox-4) and by their identical mass spectral fragmentation patterns.[2]

The structures of the tricyclic esters were determined on the basis of their relationship to the corresponding tricyclic alkanes, given that the structure of the C_{20} tricyclane (**1**) was established from the synthesis and the higher members of the series differ only in the length of the isoprenoid side chain at C-14. For example, the reduction of the pair of C_{24} esters (**14**) to the C_{24} alkane (**10**) required that the position of the carboxy group should be as shown. The C_{21}, C_{22}, and C_{23} esters were inferred to have structures (**11**), (**12**), and (**13**), respectively, and C_{25} and C_{26} structures (**15**) and (**16**), respectively.

We synthesized four possible isomers of the C_{20} ester,[19] with a methyl group at C-13 and a methoxycarbonyl group at C-14 (**6**), (**7**), (**6**, α-CO_2Me), and (**7**, α-CO_2Me). None of these compounds matched with the C_{20} tricyclic ester. The plots of the Kovat's index *vs.* carbon number[3,21] were linear for compounds (**11**), (**12**), and (**13**), as well as for compounds (**14**), (**15**), and (**16**); therefore the former and the latter are members of two homologous series. Similarly, the two-stage reduction of the C_{20} ester from the bitumen produced a C_{20} alkane which did not correspond to the C_{20} tricyclic terpene (**1**), from the saturate fraction. Since the C_{20} ester and its corresponding alkane both have m/z 191 as the base peak in their mass spectra, which requires[2] a substituent at C-14, we consider (**17**) as the most likely structure for the C_{20} ester.

Geochemical decarboxylations of the tricyclic carboxylic acids to produce tricyclic alkanes are not sufficient to explain the full range and distribution of the latter. Also, the tricyclic alkanes above C_{24} consisted of pairs of diastereoisomers of approximately equal concentration, as was the case for the C_{24} and C_{25} tricyclic esters; however, for the C_{26} tricyclic ester the ratio of diastereoisomers was 7:1. This low rate of isomerization of a chiral centre more than two carbon atoms removed from the carboxy moiety has been previously reported[22] for acyclic isoprenoid and hopanoid acids. We also observed this phenomenon in the pentacyclic hopanes obtained from the reduction of the corresponding esters.

The tricyclic terpenoid hydrocarbons present in most petroleums are also present in the Alberta oil sands as tetracyclic sulphides and sulphoxides.[23] The presence of their carboxylic acid derivatives as well points to the widespread occurrence and complexity of the, as yet, unknown biological processes responsible for these tricyclic terpenoids.

We thank Drs. D. S. Montgomery, C. M. Ekweozor, E. M. Lown, and Professor H. J. Liu for helpful discussions, Dr. G. A. Poulton for the manool, Dr. J. Payzant for the mass spectra, Dr. A. Dimmler for the capillary columns, and the Natural Sciences and Engineering Research Council of Canada and the Alberta Oil Sands Technology and Research Authority for financial support.

Received, 5th May 1983; Com. 571

References

1 C. M. Ekweozor and O. P. Strausz, *Tetrahedron Lett.*, 1982, 2711.
2 F. R. Aquino Neto, A. Restle, J. Connan, P. Albrecht, and G. Ourisson, *Tetrahedron Lett.*, 1982, 2027.
3 C. M. Ekweozor and O. P. Strausz, Abstracts of Papers, 10th International Meeting on Organic Geochemistry, Bergen, Norway, September 1981; *Adv. Org. Geochem.*, 1983, in the press.
4 B. R. T. Simoneit and P. F. Lonsdale, *Nature*, 1982, **295**, 198.
5 J. Stuart Richardson and D. E. Miiller, *Anal. Chem.*, 1982, **54**, 765.
6 W. E. Reed, *Geochim. Cosmochim. Acta*, 1977, **41**, 237.
7 W. K. Seifert and J. M. Moldowan, *Geochim. Cosmochim. Acta*, 1978, **42**, 77.
8 E. E. van Tamelen and R. G. Nadeau, *J. Am. Chem. Soc.*, 1967, **89**, 176; E. E. van Tamelen and S. A. Marson, *ibid.*, 1975, **97**, 5614.
9 J. Hellou, R. J. Anderson, S. Rafii, E. Arnold, and J. Clardy, *Tetrahedron Lett.*, 1981, 4173.
10 W. K. Seifert, J. M. Moldowan, and R. W. Jones, 10th World Petroleum Congress, Bucharest, 1979, Special Paper SP8, Heyden, London, 1980, p. 425.
11 A. S. Mackenzie, N. A. Lamb, and J. R. Maxwell, *Nature*, 1982, **295**, 223.
12 B. P. Tissot and D. H. Welte, 'Petroleum Formation and Occurrence,' Springer-Verlag, Berlin, 1978, ch. 3.
13 W. K. Seifert, *Fortschr. Chem. Org. Naturst.*, 1975, **32**, 1.
14 The oil sand sample was obtained from the Syncrude Mine (Co-ordinates 50845E, 46200N) and the carboxylic acid concentration was 1.9% of the bitumen.
15 A. Dimmler and O. P. Strausz, *J. Chromatogr.*, submitted for publication.
16 R. C. Cambie, K. N. Joblin, and A. E. Preston, *Aust. J. Chem.*, 1971, **24**, 2365.
17 S. F. Hall and A. C. Oehlschlager, *Tetrahedron*, 1972, **28**, 3155.
18 S. Bory, D. D. K. Manh, M. Fetizon, M. Kone, and N. T. Anh, *Bull. Soc. Chim. Fr.*, 1975, 2347. In our hands, it was necessary to heat the reaction mixture to 100 °C for 1 h. Using the conditions described by Bory *et al.* we isolated only starting material plus two products due to rearrangement of the exocyclic double bond into the B ring. See also G. Cimino, D. DeRosa, S. De Stefano, and L. Minale, *Tetrahedron*, 1974, **30**, 645.
19 D. K. Manh Duc, M. Fetizon, and M. Kone, *Bull. Soc. Chim. Fr.*, 1975, 2351. The following yields of (**6**) and (**7**) respectively, were obtained in our laboratory, using a high ratio of catalyst to substrate, for the hydrogenation (H_2, 1 atm) of compound (**5**): PtO_2, AcOH (8%, 92%); PtO_2, MeOH (39,61); Pd–C, MeOH (58,41); Pd–C, pentane (64,36). Compound (**5**, α-CO_2Me), obtained from *Z*-(**4**) in 87% yield using the conditions described in Scheme 1 (iii) was also hydrogenated and gave the following yields of (**6**, α-CO_2Me) and (**7**, α-CO_2Me), respectively: Pd–C, MeOH or pentane (5%, 95%); PtO_2, MeOH, 0.02 M $HClO_4$ (8,90) and 5% Rh–C, pentane (33,67). The isomers separated by preparative t.l.c. (1—3% Et_2O–pentane; 10% $AgNO_3$–silica gel; 3—4 successive elutions) were obtained in 96—99% purity by capillary g.c.
20 Y. Fujimoto and T. Tatsuno, *Tetrahedron Lett.*, 1976, 3325.
21 A. G. Douglas, M. Blumer, G. Eglinton, and K. Douraghi-Zadeh, *Tetrahedron*, 1971, **27**, 1071.
22 A. S. MacKenzie, R. L. Patience, D. A. Yon, and J. R. Maxwell, *Geochim. Cosmochim. Acta*, 1982, **46**, 783, and references therein.
23 J. D. Payzant, D. S. Montgomery, and O. P. Strausz, *Tetrahedron Lett.*, 1983, **24**, 651.