# Proof of Structure of Steroid Carboxylic Acids in a California Petroleum by Deuterium Labeling, Synthesis, and Mass Spectrometry[1]

**Wolfgang K. Seifert,* Emilio J. Gallegos, and Richard M. Teeter**

*Contribution from the Chevron Oil Field Research Company and Chevron Research Company, Richmond, California 94802.* *Received August 31, 1971*

**Abstract:** Discovery and proof of structure of four stereoisomeric $C_{22}$–$C_{24}$ steroid acids in virgin petroleum in admixture with thousands of $C_{16}$–$C_{31}$ petroleum carboxylic acids are described. The position of the carboxyl groups was fixed *via* reduction to alcohols and tosylation and reduction with lithium aluminum deuteride. The identity of synthetic deuterium-labeled and unlabeled 20-methyl-5$\alpha$-pregnanes and 5$\alpha$- and 5$\beta$-cholanes with the derivatives of the natural products is demonstrated by gas chromatography (gc) combined with mass spectrometry (ms). Final proof for the presence of two stereoisomers of 5$\alpha$-pregnane-20$\xi$-carboxylic acid and of 5$\alpha$- and 5$\beta$-cholanic acid was obtained by gcms of natural product and synthetic perfluoroalcohol esters. Both animal sources (bile acids) and plant sources (*i.e.*, unsaturated sterols) are proposed to explain the presence of the steroid acids in their observed ratios.

The structure of carboxylic acids in petroleum has been the subject of investigations throughout the century.[2] One incentive for scientists in this area of research lies in the proposal[3] that acids may be the precursors of petroleum hydrocarbons. The theory that petroleum is of biological origin had its beginning with the discovery of porphyrins[4] supplemented by recent work of Baker, *et al.*,[5] and others. It has now become a generally accepted concept supported by the presence[6] of a variety of compounds in sediments that appear to be related to similar ones occurring in the living organism (*e.g.*, amino acids, fatty acids, carbohydrates, sterols[7]). Similarly, biological marker hydrocarbons such as isoprenoids,[8] pentacyclic triterpanes,[9] tricyclic diterpanes,[10d] and steranes[10] in sediments[10a–e] and petroleum[10f,g] have become identifiable by modern means of instrumentation. The progress in petroleum carboxylic acid research has been retarded because of the enormous complexity of the mixtures and the difficulty of isolation and separation. Thus, with the exception of fatty acids, the $C_{10}$ limit had not been exceeded until 1964 when Cason[11] and coworkers first discovered and proved the structure of four $C_{14}$–$C_{20}$ isoprenoid[11a] and several[11b,c] cyclic and acyclic $C_{10}$–$C_{11}$ carboxylic acids in petroleum. This finding was followed by a series of papers on isoprenoid acids in sediments, the most recent of which[12a–c] summarizes the literature.

We now wish to report the first discovery of *steroid* carboxylic acids in any petroleum source with structure proof for two $C_{22}$ and two $C_{24}$ stereoisomeric steroid acids. Our interest in the structure of individual carboxylic acids grew out of our work on classes[13a] of carboxylic acids in a virgin California petroleum of Pliocene age Field (Midway–Sunset, $10^7$ years); it resulted in the addition of some 40 new compound classes of carboxylic acids[13b,c] from $C_{16}$ to $C_{31}$ to the existing knowledge of a few classes known to date and summarized in one of our papers.[13d]

Despite the extensive separation[13e] depicted by Scheme I, a rather "pure" thin-layer chromatographic subfraction of D-4, which is free of phenols,[13f] was estimated by mass spectrometry[13g] to contain at least 1500 compounds. Reduction of this fraction yielded hydrocarbons which could be more readily and extensively separated. The location of the carboxyl group was fixed by parallel reduction to deuteriomethyl and methyl (Scheme I). This approach resulted in working on a milligram scale at the hydrocarbon level; therefore, the only instrumental method of analysis applicable to the nanogram quantities of individual steroid acid derivatives present in the mix-

(1) Partial preliminary communication: W. K. Seifert, E. J. Gallegos, and R. M. Teeter, *Angew. Chem., Int. Ed. Engl.*, 10, 747 (1971); *Angew. Chem.*, **83**, 805 (1971).

(2) H. L. Lochte and E. R. Littman, "Petroleum Acids and Bases," Chemical Publishing Co., New York, N. Y., 1955.

(3) J. E. Cooper and E. E. Bray, *Geochim. Cosmochim. Acta*, **27**, 1113 (1963).

(4) A. Treibs, *Angew. Chem.*, **49**, 682 (1936), and references therein.

(5) (a) E. W. Baker, *J. Amer. Chem. Soc.*, **88**, 2311 (1966); (b) E. W. Baker, T. F. Yen, J. P. Dickie, R. E. Rhodes, and L. F. Clark, *ibid.*, **89**, 3631 (1967).

(6) (a) G. Eglinton and M. T. J. Murphy, Ed., "Organic Geochemistry," Springer Verlag, New York, N. Y., 1969; (b) P. Albrecht and G. Ourisson, *Angew. Chem., Int. Ed. Engl.*, 10, 209 (1971).

(7) (a) G. Mattern, P. Albrecht, and G. Ourisson, *Chem. Commun.*, 1570 (1970); (b) W. Henderson, W. E. Reed, G. Steel, and M. Calvin, *Nature (London)*, **231**, 308 (1971).

(8) (a) J. G. Bendoraitis, B. L. Brown, and L. S. Hepner, *Anal. Chem.*, **34**, 49 (1962); (b) R. A. Dean and E. V. Whitehead, *Tetrahedron Lett.*, **21**, 768 (1961).

(9) I. R. Hills, E. V. Whitehead, D. E. Anders, J. J. Cummins, and W. E. Robinson, *Chem. Commun.*, 20, 752 (1966).

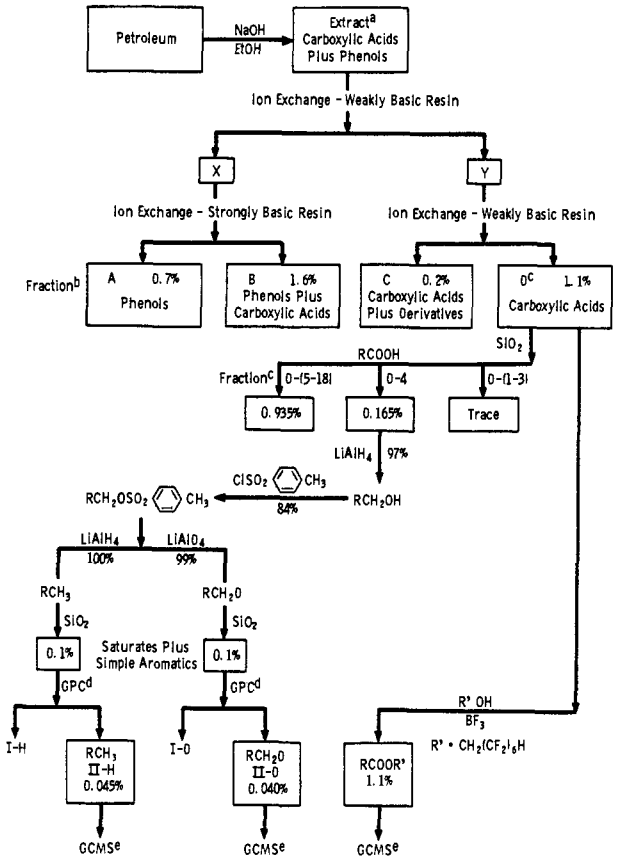
(10) (a) A. L. Burlingame, P. Haug, T. Belsky, and M. Calvin, *Proc. Nat. Acad. Sci. U. S.*, **54**, 1406 (1965); (b) W. Henderson, V. Wollrab, and G. Eglinton, *Chem. Commun.*, 710 (1968); P. A. Schenck and I. Havenaar, *Proc. Int. Meet. 4th, Org. Geochem. 1968*, 181 (1969); (c) P. C. Anderson, P. M. Gardner, E. V. Whitehead, D. E. Anders, and W. E. Robinson, *Geochim. Cosmochim. Acta*, **33**, 1304 (1969); (d) E. J. Gallegos, *Anal. Chem.*, **43**, 1151 (1971); (e) D. E. Anders and W. E. Robinson, *Geochim. Cosmochim. Acta*, **35**, 661 (1971); (f) I. R. Hills and E. V. Whitehead, *Nature (London)*, **209**, 977 (1966); (g) T. C. Hoering, personal communication.

(11) (a) J. Cason and D. W. Graham, *Tetrahedron*, **21**, 471 (1965); (b) J. Cason and K. L. Liauw, *J. Org. Chem.*, 30, 1763 (1965); (c) J. Cason and A. I. A. Khodair, *ibid.*, 31, 3618 (1966); **32**, 3430 (1967).

(12) (a) A. E. Douglas, M. Blumer, G. Eglinton, and K. Douraghi-Zadeh, *Tetrahedron*, **27**, 1071 (1971); (b) R. E. Cox, J. R. Maxwell, G. Eglinton, and C. T. Pillinger, *Chem. Commun.*, 1639 (1970); (c) R. C. Murphy, M. V. Djuricic, S. P. Markey, and K. Bieman, *Science*, **165**, 695 (1969).

(13) (a) W. K. Seifert and R. M. Teeter, *Chem. Ind. (London)*, 1464 (1969); (b) W. K. Seifert and R. M. Teeter, *Anal. Chem.*, **42**, 750 (1970); (c) W. K. Seifert and R. M. Teeter, *ibid.*, **42**, 180 (1970); (d) W. K. Seifert, R. M. Teeter, W. G. Howells, and M. J. R. Cantow, *ibid.*, **41**, 1638 (1969); (e) W. K. Seifert and W. G. Howells, *ibid.*, **41**, 554 (1969); (f) W. K. Seifert, *ibid.*, **41**, 562 (1969); (g) W. K. Seifert and R. M. Teeter, *ibid.*, **41**, 786 (1969).

**Scheme I**




[a] This extract represents 3.54% of the total crude oil; 2.5% based on crude oil is RCOOH, see Tables I–III of reference 13e. [b] Percentages are based on total crude oil. [c] Represents 40% of all RCOOH present in crude oil. [d] Gel permeation chromatography, see Figure 1. [e] Gas chromatography plus mass spectrometry.

tures was capillary gas chromatography (gc) combined with mass spectrometry (ms).

More specifically, acid fraction D-4, representing 6% of all acids and equal to 0.17% of the petroleum, was investigated (Scheme I). Fractions of predominantly saturated hydrocarbons isolated from both the labeled and unlabeled portions in identical fashion by silica gel chromatography were further separated by gel permeation chromatography (Scheme I). The deviation of the refractive index from that of the solvent is given in Figure 1. It illustrates the duplication between labeled and unlabeled fractions for the three steps of reduction, silica gel chromatography, and gel permeation chromatography. The investigation was focused on the pair II-H and II-D (Figure 1) because they showed the highest concentration of steroid hydrocarbons. The gas chromatogram of the unlabeled hydrocarbon is given in Figure 2A and is identical with that of the labeled sample. The two largest peaks were recognized as terpanes from their mass spectra.[14]

(14) A complication which interfered with the interpretation of the gcms data was the continual emergence, from the chromatograph, of a mixture of unresolved components which resulted in a significant background to each steroid mass spectrum. Correction for these background contributions was made by averaging backgrounds taken before and after the steroid peak and deducting the average from the observed mass spectrum of the steroid. The error inherent in this method is due to the variation of the background during the emergence of the individual steroid compounds. Therefore, the significance of the mass spectra depends upon the general shape of the pattern rather than the exact intensities of the fragment peaks.

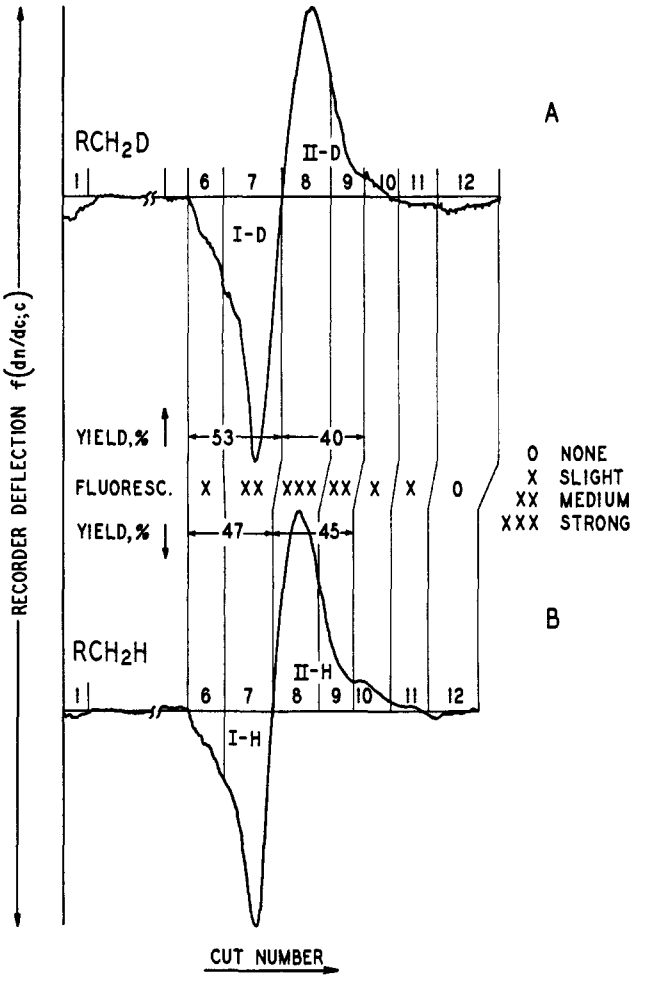



Figure 1. Gel permeation chromatograms of deuterium-labeled hydrocarbons (silica gel fractions **I** of hydrocarbons derived from carboxylic acid fraction D-4) and hydrogen analogs.

**$C_{22}$ Acids.** The mass spectrum of the first significant steroid hydrocarbon emerging after the $C_{24}$-terpane is shown in Figure 3A and that of its deuterated counterpart in Figure 3C. The dominant fragments in steroids at *m/e* 218, 217, and 232 have been shown to arise from cleavage of ring D[15] (see formula I for numbering). Because the deuterated steroid (Figure 3C) shows no increase in mass at these positions, the carboxylic acid group cannot be attached to rings A, B, or C, but must be at ring D. Moreover, the presence of the fragment at *m/e* 232 shows $C_{15}$ to be unsubstituted because this particular fragment arises from cleavage of the 13–17 and 15–16 bonds.[15] Furthermore, the appearance of the spectrum (Figure 3A) as a whole resembles closely that of a naturally occurring steroid with substitution on $C_{17}$, and on that basis substitution at $C_{16}$ can be excluded.[16] The parent minus methyl fragments in the labeled and unlabeled compounds at *m/e* 288 and 287, respectively, must represent loss of the 18- and 19-methyl groups because no loss of deuteriomethyl (mass 16) is observed.

Reported spectra of 5α- and 5β-pregnane[15] show that the stereochemistry of rings A and B can be recognized as trans by the near absence of a fragment at *m/e* 151

(15) L. Toekes, G. Jones, and C. Djerassi, *J. Amer. Chem. Soc.*, 90, 5465 (1968).
(16) C. Djerassi, personal communication, 1971.

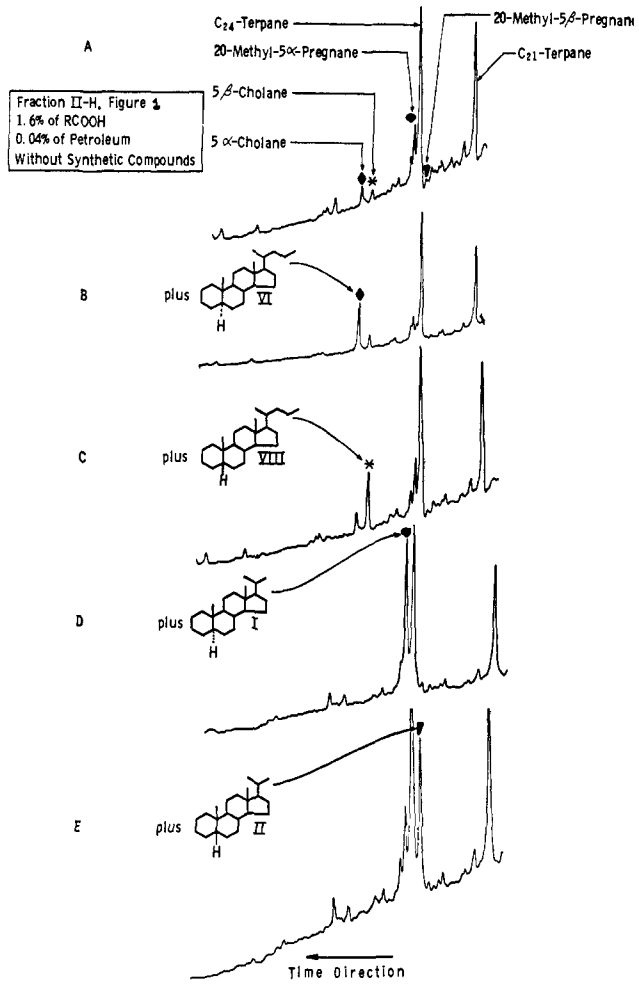



Figure 2. Gas chromatogram of carboxylic acid derived hydrocarbons coinjected with synthetic steroids. Time is measured on a variable scale.

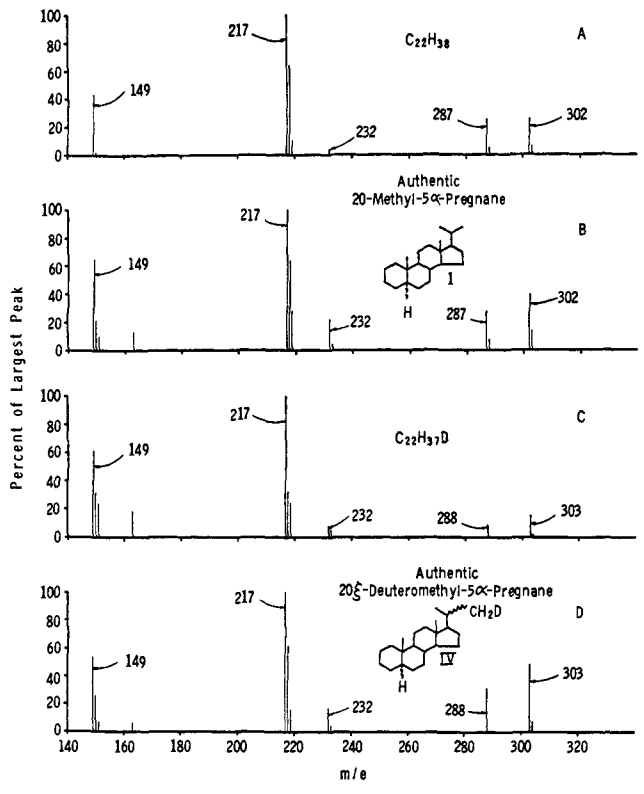
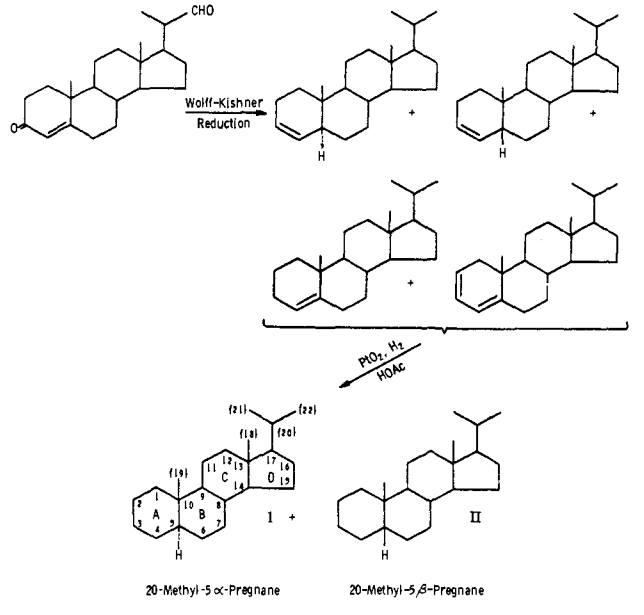



Figure 3. Mass spectra of 5α-pregnane-20ξ-carboxylic acid derived petroleum hydrocarbons (gcms; after background deduction in A and C) *vs.* synthetic compounds.

**Scheme II**




and as cis by a fragment at *m/e* 151 of intensity about equal to or greater than that at *m/e* 149; this observation was confirmed in this laboratory by comparison of the spectra of 5α-cholestane and 5β-cholestane. Figures 3A and 3C indicate the trans A/B configuration for this naturally occurring steroid carboxylic acid.

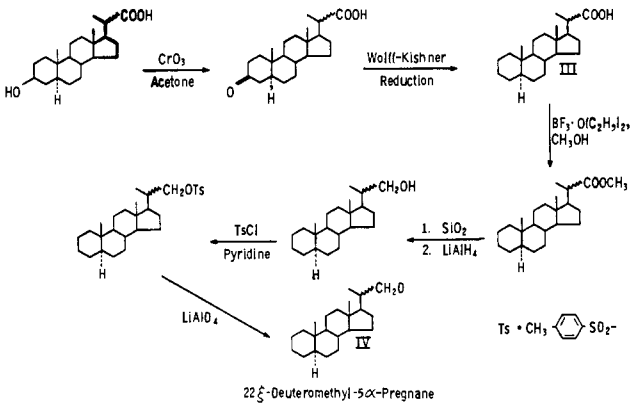
The unlabeled trans (I) and cis (II) hydrocarbons were synthesized as a mixture as shown in Scheme II, and separated by preparative gas chromatography.

The mass spectral fragmentation pattern of I was obtained from gcms data collected under the same conditions as was that of the petroleum-derived hydrocarbon, and the two are shown in Figures 3A and 3B. Differences in relative intensities are due to uncertainties in the subtraction of background. Gas chromatographic coinjection of the synthetic 5α-trans compound I with the petroleum carboxylic acid derived hydrocarbons illustrates the enhancement of the natural product peak of interest (Figure 2D). The synthetic 5β-cis compound shows a fragment at *m/e* 151 of approximately the same intensity as its *m/e* 149 peak. Coinjection with the petroleum-derived hydrocarbon (Figure 2E) showed it to have a retention time identical with that of a very small peak emerging just prior to the $C_{24}$ terpane. A weak steroid mass spectrum was detected at this position in the deuterium-labeled mixture with *m/e* 151 equal to *m/e* 149. Our conclusion is that cis isomeric pregnanecarboxylic acid is present in this petroleum in much smaller (trace) quantity than the trans isomer.

In order to confirm the proposed position of the carboxyl group at $C_{20}$ in this $C_{22}$-steroid acid, the required deuterium-labeled hydrocarbon IV was synthesized *via* the authentic carboxylic acid (III) as shown in Scheme III.

The general mass spectral pattern of this 20ξ-deuteriomethyl-5α-pregnane (Figure 3D) obtained under gcms conditions is very similar to that derived from the nat-

**Scheme III**




ural product (Figure 3C); gas chromatographic co-injection of IV with the deuterium-labeled hydrocarbon mixture results in peak enhancement the same as that shown in Figure 2D for the unlabeled hydrocarbon I.

A comparison of the intensities of the fragments at *m/e* 44 of the labeled and unlabeled pairs additionally supports the proposed structure. This peak, in the unlabeled species, is due to $C_2{}^{13}CH_7$. In the labeled species, *m/e* 44 is enhanced by the presence of $C_3H_6D$ from the side chain at $C_{17}$ of the steroid nucleus. Again, background corrections reduce the accuracy of the figures, but the height increase (relative to the height of the *m/e* 217 peak) is the same (12%) for the synthetic pair (4–16%) and the petroleum-based pair (1–13%). This indicates that the side-chain structures are probably the same. Proof would require a demonstration that the *n*-propyl group yields a different increase.

To further prove the presence of the isopropyl group and to determine whether the carboxyl group at $C_{20}$ is stereochemically $\alpha$ or $\beta$ or both, an ester of crude carboxylic acid fraction D (Scheme I, representing 40% of all carboxylic acids in this petroleum) with 1,1,7-trihydroperfluoroheptanol[13g] was investigated by gcms. Figure 4 shows the mass chromatogram (see Experimental Section) using the intensities of the parent peak of the ester at *m/e* 646 to monitor the occurrence of the various isomers of the $C_{22}$ acids. Three isomers, A, B, and C, of very similar mass spectral fragmentation pattern were observed (Table I). The ratio of the retention times of B/A was 1.023 and of C/A, 1.042. To clarify this situation, the fluoroalcohol ester of synthetic acid III was prepared; it showed the presence of two isomers (gcms) whose mass spectral fragmentation patterns are very similar to one another (Table I). Their retention time ratio was 1.023, indicating that A and B of the petroleum sample are the same as A and B of the authentic sample and that isomer C occurs in the petroleum only. When the mixture of these two synthetic perfluoroalcohol ester 5$\alpha$ isomers was coinjected with the natural product ester mixture, compounds A and B (Figure 4) were enhanced, not compound C. The agreement of the mass spectral fragmentation patterns of compounds A[17] and B with their synthetic counterparts of equal retention time is seen from Table I. The height ratio of the peaks at *m/e* 149 and 151 indicates both A and B to be 5$\alpha$ (Table I). They are believed to be the isomers with the carboxyl group at $C_{20}$, either $\alpha$ or $\beta$. The latter distinction is no longer detectable after reduction to deuterium-labeled hydrocarbon IV which gives the appearance of one single compound by gcms. Our conclusion is that both 5$\alpha$-pregnane-20$\alpha$-carboxylic acid and its 20$\beta$ isomer are present in this petroleum.

The mass spectrum of C (Table I) indicates an A/B trans 5$\alpha$ isomer. It may differ from A and B by having either a branched chain at $C_{17}$ in the $\alpha$ configuration or a straight chain in either $\alpha$ or $\beta$ configuration at $C_{17}$. From Figure 4 the amount of the unidentified $C_{22}$-steroid carboxylic acid (isomer C) is estimated to be about one-tenth that of the sum of the identified acids (isomers A and B).

**$C_{24}$ Acids.** In addition to the above-described $C_{22}$-steroid acids, work on the acid-derived, labeled, and unlabeled hydrocarbons revealed two additional stereoisomeric $C_{24}$-steroid acids. Comparison of their mass spectra (Figures 5A and 5C) and the same mass

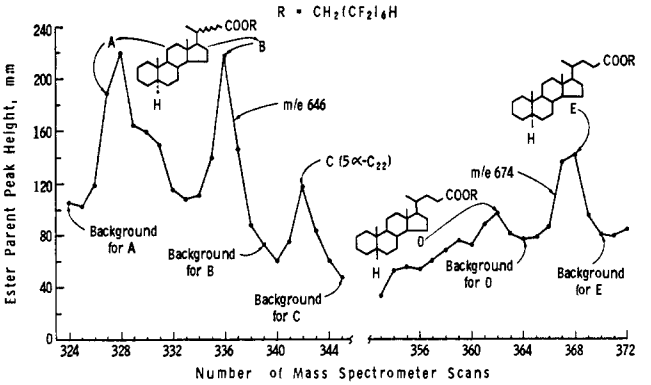
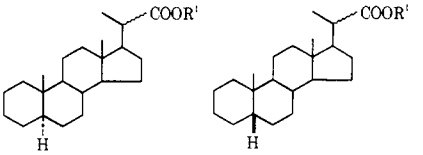



Figure 4. Molecular ion mass chromatogram of esters (with 1,1,7-trihydroperfluoroheptanol) of petroleum steroid carboxylic acids. The number of mass spectrometer scans is from sample injection; The scan cycle is 8.7 sec. The entire spectrum is used to generate Tables I and II.

**Table I.** Mass Spectra[a] of 5$\alpha$-Pregnane-20$\xi$-carboxylic Acid Fluoroalcohol Esters by Gas Chromatography–Mass Spectrometry (R = $CH_2(CF_2)_6H$)




| *m/e* | A, from Synthetic | A, from petroleum[b] | B, from Synthetic | B, from petroleum[b] | C, from petroleum[b] |
|---|---|---|---|---|---|
| 149 | 66 | 98 | 62 | 97 | 74 |
| 151 | 14 | 20 | 12 | 41 | 19 |
| 217 | 100 | 100 | 100 | 100 | 100 |
| 218 | 30 | 33 | 34 | 35 | 43 |
| 232 | 35 | 33 | 42 | 37 | 20 |
| 374 | 0 | 7 | 0 | 0 | 4 |
| 388 | 1 | 2 | 2 | 4 | 1 |
| 631 | 21 | 17 | 23 | 39 | 22 |
| 646 | 32 | 26 | 32 | 41 | 26 |
| Retention time, min | 47.3 | | 48.4 | | 49.3 |

[a] Peak intensities relative to *m/e* 217 [b] After background deduction.

(17) The spectrum of compound A was not taken at the maximum intensity of the *m/e* 646 peak (see Figure 4) because at the time of scan 328, an unidentified component of the mixture was starting to elute and its fragments were interfering. This was revealed when the plots of peaks other than *m/e* 646 failed to change in synchronism with the 646 peak.

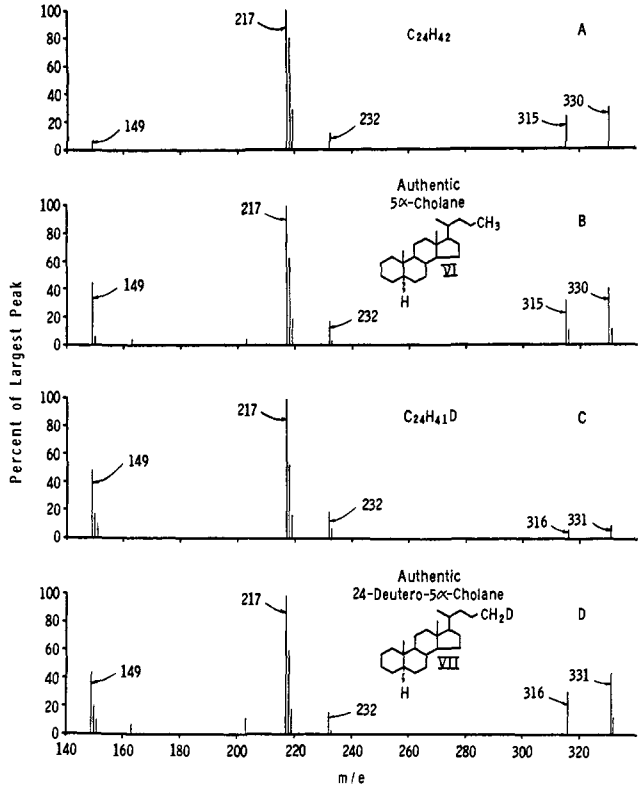



Figure 5. Mass spectra of 5α-cholanic acid derived petroleum hydrocarbons (gcms; after background deduction in A and C) *vs.* synthetic compounds.

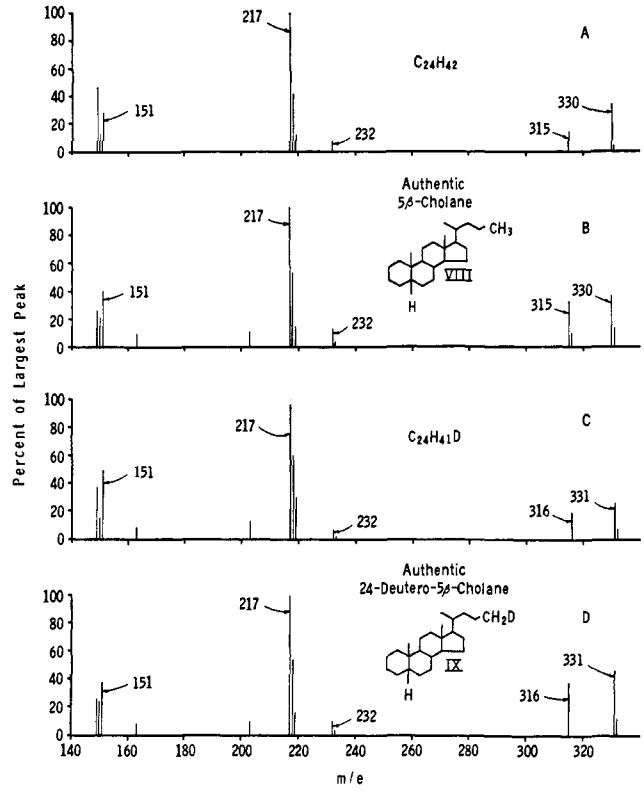
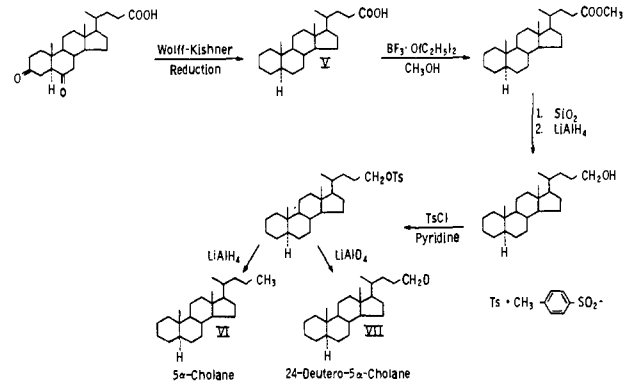



Figure 6. Mass spectra of 5β-cholanic acid derived petroleum hydrocarbons (gcms; after background deduction in A and C) *vs* synthetic compounds.

spectral arguments as described above for the deuterium-labeled and unlabeled $C_{22}$ compounds led us to synthesize 5α-cholane (VI) and 24-deuterio-5α-cholane (VII) *via* 5α-cholanic acid (V) as shown in Scheme IV.

**Scheme IV**

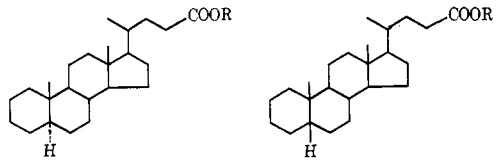



Analogously 5β-cholane (VIII) and 24-deuterio-5β-cholane were synthesized by the above shown route from 5β-cholanic acid (X). It is noteworthy that the experimental conditions for converting the authentic steroid acids to the corresponding hydrocarbons of known stereochemistry were identical with those applied to the natural product acids, thus demonstrating the stability of the hydrocarbon skeletons of the latter during this transformation. The mass spectral similarity of the four synthetic labeled and unlabeled hydrocarbons with their natural product counterparts is seen in Figures 5 and 6. Gc coinjection is illustrated in Figures 2B and 2C showing the shorter retention time of the A/B cis isomer.

Final proof of the presence of 5α- and 5β-cholanic acids in this petroleum was obtained from the gcms data of the fluoroalcohol ester of fraction D (Scheme I). The mass chromatogram (Figure 4) of the petroleum acid fluoroalcohol ester monitored by the intensity of the parent peak at *m/e* 674 for the $C_{24}$ acids demonstrates the presence of two isomers, D and E, whose mass spectral fragmentation patterns are compared with those of the esters from synthetic 5α-cholanic acid (V) and 5β-cholanic acid (X) in Table II. Gc coinjection with the natural product esters enhanced compounds E and D (Figure 4), respectively, the 5β isomer having the slightly shorter retention time. Additional evidence comes from the retention times of synthetic and petroleum-derived 5α-cholanic acid esters when measured relative to the most abundant of the pregnane carboxylic acid esters. Retention time ratios of 1.122 for the ester of acid V/A and 1.123 for E/A were measured.

As in the mass spectra of hydrocarbons, the 5α configuration of the synthetic (V) and naturally occurring 5α-cholanic acid E is illustrated by the predominance of *m/e* 149 over *m/e* 151 in the fluoroalcohol ester spectra (Table II); analogously, the esters of naturally occurring 5β-cholanic acid D and of synthetic acid X show the increase of the intensity at *m/e* 151 required by the 5β configuration. The absolute values of some of the fragment intensities of the ester of the naturally occurring 5β-cholanic acid D are distorted by background which has a particularly large effect on this isomer because its amount is the smallest of all of the steroid acid esters of Figure 4. Fragments due to McLafferty rearrangement at *m/e* 374 in 5α- and 5β-

**Table II.** Mass Spectra[a] of 5α- and 5β-Cholanic Acid Fluoroalcohol Esters by Gas Chromatography–Mass Spectrometry (R = $CH_2(CF_2)_6H$)




| m/e | E, from Synthetic | E, from petroleum[b] | D, from Synthetic | D, from petroleum[b] |
|---|---|---|---|---|
| 149 | 42 | 51 | 26 | 60 |
| 151 | 8 | 11 | 34 | 80 |
| 217 | 100 | 100 | 100 | 100 |
| 218 | 51 | 38 | 55 | 40 |
| 232 | 37 | 36 | 19 | 20 |
| 373 | 3 | 0 | 4 | 8 |
| 374 | 3 | 4 | 4 | 16 |
| 564 | 10 | 11 | 9 | 12 |
| 659 | 34 | 30 | 44 | 76 |
| 674 | 54 | 28 | 85 | 80 |
| Retention time, min | 53.1[c] | | 52.4[c] | |

[a] Peak intensities relative to *m/e* 217. [b] After background deduction. [c] Of the petroleum-derived esters.

cholanic acids are small in synthetic and natural product esters and even smaller at *m/e* 388 in pregnane carboxylic acid ester. The reason for the latter is probably the rigidity of ring D of the steroid system[16] coupled with interference by the $C_{19}$ methyl with the planarity of the six-membered ring intermediate.[18]

**Amounts of Acids and Their Significance.** An estimate of the relative amounts of the individual steroid acids in this petroleum can be based on the intensities (relative to *m/e* 217) of the parent peaks of the four synthetic acid esters given in Tables I and II combined with the relative heights of parent peaks of the derivatives A–E of the naturally occurring acids in the mass chromatogram (Figure 4). The sum of the two 5α-pregnane-20ξ-carboxylic acids, A and B, is about seven times the quantity of 5α-cholanic acid E; the latter occurs in four–five times greater amount than 5β-cholanic acid D. From the hydrocarbon work described above, 5β-pregnane-20ξ-carboxylic acid appears to be present in an amount smaller than 5β-cholanic acid.

The sum of the two predominant 5α-pregnane-20ξ-carboxylic acids A and B, both of which become 20-methyl-5α-pregnane after reduction, is estimated from the occurrence of the latter in reduced fraction D-4 (Scheme I and Figure 2A) and in reduced fraction D (Scheme I) to be about 6 ppm of the petroleum. Therefore, from the ratios described above, 5α-and 5β-cholanic acids are estimated at about 1 and 0.2 ppm of the petroleum, respectively. Because only 40% of all acids in this petroleum were investigated and since the complexity of the mixture prevented in-hand isolation, these analytical values allow an estimate of the order of magnitude of all steroid acids identified as about 0.1–10 ppm of the petroleum. It is therefore obvious that proof of structure, including stereochemical details of such polar compounds in the enormously complex mixtures in which they naturally occur, requires the elaborate separation sequences combined with labeling techniques (which exclude contamination) and modern instrumental analysis plus organic synthesis described here. The fortuitous factors in this work were good sensitivity of steroid mass spectra and the ease of interpretation of the labeled compounds due to the extensive mass spectral work done on steroid systems by others.[15]

Because of the presence of functional groups and the recognition of stereochemical features, the detailed information which can be derived biogeochemically from steroid acids in petroleum may exceed that of any biological marker hydrocarbon and that of any carboxylic acid except those isoprenoid acids recently shown[12b] to be linked to the phytyl side chain of chlorophyll. While these latter acids supplement the abundant evidence for the predominant plant genesis of petroleum, the identification of the steroid acids reported here justifies postulation of some animal contribution. Because the observed abundance ratio of 5α-$C_{22}$ to 5β-$C_{22}$ (30:≪1) is different from the corresponding ratio of 5α-$C_{24}$ to 5β-$C_{24}$ (4:1), thermodynamic equilibrium between 5α and 5β in both pairs is most unlikely. In addition, the greater amount of 5β-$C_{24}$ over 5β-$C_{22}$, which is inverse to the ratio in the 5α series (5α-$C_{22}$'s > 5α-$C_{24}$), warrants some speculation.

Most $C_{24}$ bile acids presently known[19] possess A/B cis configuration (*i.e.*, 5β), and no abundant 5β-$C_{24}$ steroids of plant origin are known. Therefore, the most likely sources for the 5β-$C_{24}$ acid are 5β-$C_{24}$ bile acids which have various numbers of hydroxyl groups (*i.e.*, cholic acid). Upon dehydration and hydrogenation, transformations generally accepted by petroleum chemists, all these 5β bile acids would yield the observed 5β-cholanic acid. The 5α-$C_{24}$ acid, being more abundant than the 5β-$C_{24}$ isomer, could be derived by isomerization of the latter, because the trans-5α isomer is the thermodynamically more stable one,[20] or in part from the less abundant 5α-$C_{24}$ bile acids. Allowing β oxidation of 5α-$C_{24}$ to 5α-$C_{22}$, the latter can also be explained from bile acids.

However, plants also have to be considered as possible sources, *i.e.*, the 5α-$C_{22}$ steroid acid isomers are derivable by oxidative $C_{22}$–$C_{23}$ double bond cleavage of sterols which occur in large variety in the plant kingdom (*i.e.*, dihydroergosterol, algae) or, less likely, oxidation of steroidal sapogenins. Thus, the larger quantity and number of 5α-$C_{22}$ steroid acid isomers may reflect a greater variety of sources. Analogously, 5α-$C_{24}$ can be derived by oxidative $C_{24}$–$C_{25}$ double bond cleavage of zymosterol.

In summary, animal sources alone can explain all steroid acid isomers found. Plant sources, although they alone cannot explain all the isomers found in their observed ratios, are very likely contributors. Consequently, both, which existed at Pliocene age, are proposed to explain the results, unless recent bacterial activity has occurred voiding the above conclusions.

## Experimental Section

**Instrumentation.** For gcms analysis of hydrocarbons, an F&M Model 810 gas chromatograph without a splitter was used in

(18) J. A. Gilpin and F. W. McLafferty, *Anal. Chem.*, **29**, 990 (1957).

(19) E. Heftmann, "Steroid Biochemistry," Academic Press, New York, N. Y., 1971, pp 55–59.

(20) M. N. Mitra and W. H. Elliot, *J. Org. Chem.*, **34**, 2170 (1969).

conjunction with an AEI MS-9 mass spectrometer. A 45 m × 0.05 mm i.d. capillary column coated with OV-17 operated at a He flow rate of about 2 ml/min was coupled directly to the source of the MS-9 without the use of an interface. This achieves maximum sample sensitivity with minimum loss of chromatographic resolution. The source pressure ion gauge read 1–2 × $10^{-5}$ Torr. A temperature of 330° was maintained at both the injection port and the column exit to the MS-9 source. The source was held at 240° using a temperature controller; the chromatograph was programmed from 100 to 340° at a rate of 4°/min using the total ion monitor of the MS-9 to obtain the chromatograph traces. Mass spectra were recorded at a magnet scan rate of 5 sec/decade at about 1000 resolution. A typical sample size injected using a 1-μl syringe was 0.3 μl. Peaks in the mass spectra of Figures 3, 5, and 6 smaller than 10% of the largest peak were not included unless of particular interest.

Gcms of the fluoroalcohol esters was performed using the same chromatograph but with a Nuclide 12-90-G mass spectrometer. The gcms conditions were similar to those described above for the MS-9 with the following differences: column length, 60 m; He flow rate, 5 ml/min, and pressure at the source ion gauge, 5 × $10^{-6}$ Torr. The mass spectrometer was set to cycle in the linear magnet current mode to gather spectra from *m/e* 140 to 750 by repetitive scanning at intervals of 8.7 sec. The typical peak width for the steroid esters in question is about 14.0 sec. Mass spectra were gathered at an electron accelerating voltage of 50 eV. All mass spectral data were recorded on a Honeywell Model 7600 tape recorder at a tape speed of 9.5 cm/sec. Selected spectra were then reproduced on a Honeywell Model 1508A Visicorder.

Both unresolved components of the mixture and column bleed interfered with the steroid mass spectra. A plot of molecular ion peak height as a function of the number of mass spectrometer scans for the petroleum fluoroalcohol ester gcms, defined as a mass chromatogram, is given in Figure 4. Scans 327, 336, 342, 362, and 368 were read and from them were subtracted scans 324, 339, 345, 364, and 370, respectively, to yield, after normalization, the partial spectra presented in Tables I and II. Average standard deviation of gas chromatographic retention time ratios of fluoroalcohol esters is 0.3% on duplicate runs. The partial spectra of the synthetic compounds were also obtained from gcms data but are more accurate because only column bleed background was present, and its subtraction could be done with better assurance that it was constant.

Exact mass measurements of synthetic labeled and unlabeled hydrocarbons and the intermediates leading to their synthesis were preferred to conventional elemental analyses. They were performed by the peak-matching technique referred to appropriate peaks of perfluorotributylamine using a quartz direct injection probe[21] for sample introduction. An average standard deviation of 4.6 ppm was observed. For preparative gas chromatography, a Wilkens Aerograph Model A-350 instrument was operated isothermally at 270°. The column was 6 mm o.d. × 6 m stainless steel packed with 3% OV-17 on Gas Chrom Q; He flow was 40 ml/min. The traps were 3 mm o.d. × 10 cm glass, air-cooled. Gas chromatographic coinjection of synthetic steroid hydrocarbons with samples derived from the petroleum (Figure 2) was made in a Model 204-1B Aerograph on a 45 m × 0.05 mm i.d. OV-17 column programmed from 180 to 300° at 6°/min, at 3–4 ml of He flow using 0.1 mg of sample.

The instrumentation for gel permeation chromatography was described previously.[13d] Optical rotations were measured in chloroform on a Cary Model 60 spectropolarimeter. Melting points were obtained on a Fisher Scientific Instrument melting point block and are uncorrected.

**Isolation, Conversion, and Separation of Natural Product Acids (Scheme I).** The isolation of carboxylic acid fraction D-4 from Midway–Sunset 31E, California crude, has been described previously.[13e] The earlier reported[13d] three-step procedure for conversion (for yields, see Scheme I) of carboxylic acids to hydrocarbons *via* alcohols was applied to this fraction by reducing half of the *p*-toluenesulfonate with lithium aluminum hydride and the other half with lithium aluminum deuteride.

Solvents for chromatography were purified by distillation and chromatography on alumina. Milligram quantities of both labeled and unlabeled hydrocarbon fractions were chromatographed on purified[22] Grace Davison Chemical Co. 100–200 mesh activated Grade 923 desiccant silica gel using large columns (2.5 cm i.d. × 60 cm). A 1-l. petroleum ether eluate containing 60% of the charge was worked up and used for subsequent gel permeation chromatography. It was later found that a much smaller column (1.25 cm i.d. × 130 cm) operated at 14 psi of pressure with hexane suffices to separate charges of several hundred milligrams and allows the use of silica gel without prepurification simply by purging the column with 200 ml of hexane before charging. (In this case, the saturated hydrocarbons are eluted free of aromatics in the first 140 ml of eluate in 15–20 min.) Deuterium labeled and unlabeled samples contained significant amounts of aromatics and were both further separated by preparative (30-mg quantities) gel permeation chromatography[13d] in identical yields, avoiding contamination by cleaning the pump and extended purging of the columns with distilled toluene (Figure 1). The unlabeled fraction II-H and the labeled fraction II-D were subjected to gcms.

**20-Methyl-5α- and -5β-pregnane.** 3-Oxapregn-4-ene-20β-carboxaldehyde (the Upjohn Co., 3.29 g, 10 mmol) was reduced to hydrocarbon by a modified Wolff–Kishner reduction[23] in triethanolamine solvent. The reaction mixture was cooled to room temperature and mixed thoroughly with water (30 ml) and dichloromethane (30 ml). After separation, the aqueous phase was extracted twice with dichloromethane (15 ml each); the organic solutions were combined and washed twice with dilute aqueous HCl, once with dilute aqueous sodium bicarbonate, and once with water and dried with magnesium sulfate; evaporation of the solvent yielded 2.79 g (92% of theory) of a dark brown oil. Gas chromatography (3% OV-17 on Gas Chrom Q) showed the presence of three major species totaling about 90–95% of the sample. The mass spectrum indicated about 92% of mol wt 300 material (theory 300) and about 8% of mol wt 298 material.

This crude mixture of isomeric monoolefins (presumably contaminated by a small amount of diene) was hydrogenated. A 0.5-g (1.66 mmol) solution of the olefin in 5 ml of glacial acetic acid with 60 mg of suspended $PtO_2$ (84%) was subjected to 80 psi of hydrogen pressure for 3 hr at 22° and an additional 3 hr at 50° with theoretical hydrogen consumption. After filtering off the catalyst and azeotroping off the solvent with excess *n*-heptane, 0.5 g of a dark oil was obtained. Chromatography on silica gel using the above-described small column yielded 0.35 g of a mixture of 72% of 20-methyl-5β-pregnane and 28% of 20-methyl-5α-pregnane as determined by gas chromatography; the retention time ratio was found to be 5α/5β = 1.08 (lit.[24] 1.12). Isolation of each pure isomer was effected by repeated preparative gas chromatography and retrapping (to remove traces of the other isomer) and subsequent microsublimation[25] to obtain samples free from traces of silicone due to gc column bleed.

20-Methyl-5α-pregnane had the following physical characteristics: mp 113° (lit.[26] mp 111–112°); mol wt calcd for $C_{22}H_{38}$, 302.2973; found, 302.2966 (all molecular weights determined by high-resolution mass spectrometry); $[\alpha]D$ +7.3 ± 0.8° (lit.[26] $[\alpha]D$ +8.9°); partial mass spectrum in Figure 3B.

20-Methyl-5β-pregnane showed the following characteristics: mp 77–78°; mol wt calcd for $C_{22}H_{38}$, 302.2973; found, 302.2971; $[\alpha]D$ +8.3 ± 0.8°; *m/e* 151 equals 47% of *m/e* 217 at a source temperature of 124°; *m/e* 151:*m/e* 149 = 1.15. For gc coinjection with petroleum acid derived hydrocarbons, see Figures 2D and 2E.

**5α-Pregnane-20ξ-carboxylic Acid Methyl Ester.** Chromium trioxide (78 mg) was dissolved in 3.3 ml of an acid mixture prepared from 0.87 ml of concentrated sulfuric acid and 3 ml of water. This oxidant solution[27] was added at 22° over the course of 1 hr to a suspension of 267 mg (0.77 mmol) of 5α-pregnan-3β-ol-20ξ-carboxylic acid, mp 278–279° (lit.[28] 274–276° (Steraloids, Inc.)), in 1.5 ml of acetone. The reaction mixture was diluted with 15 ml of water and extracted with dichloromethane which was washed with water, dried over magnesium sulfate, and evaporated to dryness. A second run with 471 mg (1.35 mmol) was made, and the combined products from the two runs were recrystallized from

(21) E. J. Gallegos, Invited Paper No. 30, Seventh World Petroleum Congress, Mexico City, Mexico, April 2–8, 1967.

(22) R. Hernandez, R. Hernandez, and L. R. Axelrod, *Anal. Chem.*, 33, 371 (1961).

(23) L. F. Fieser and M. Fieser, "Reagents for Organic Synthesis," Wiley, New York, N. Y., 1967, p 435.

(24) C. J. W. Brooks and L. Hanaineh, *Biochem. J.*, 87, 151 (1963).

(25) W. H. Melhuish, *Nature (London)*, 184, 2933 (1959).

(26) J. P. Dusza and W. Bergmann, *J. Org. Chem.*, 25, 79 (1960).

(27) See ref 23, p 142.

(28) W. Bergmann, D. H. Gould, and E. M. Low, *J. Org. Chem.*, 10, 570 (1945).

chloroform–cyclohexane to yield 466 mg (63% of theory) of 3-oxo-5α-pregnane-20ξ-carboxylic acid: mp 252°; mol wt calcd for $C_{22}H_{34}O_3$, 346.2508; found, 346.2518. The keto acid was reduced to the acid and converted to its methyl ester by the same techniques as used for the preparation of 5α-cholanic acid methyl ester (below). The crude ester (200 mg) was purified by silica gel chromatography on the above-described small column by eluting with cyclohexane followed by benzene. The cyclohexane eluate (700 ml) and the first 200 ml of benzene eluate were discarded. Pure ester product was isolated from the subsequent 400 ml of benzene eluate: mp 87.5–88°; mol wt calcd for $C_{23}H_{38}O_2$, 346.2872; found, 346.2891; ir carbonyl absorption at 1740 $cm^{-1}$. The gas chromatogram indicated the presence of more than 90% of one isomer.

**5α-Pregnane-20ξ-carboxylic Acid and 1,1,7-Trihydroperfluoroheptyl Ester.** A portion (6.6 mg) of the purified methyl ester (above) was treated with 5 ml of 2 *N* ethanolic KOH at 95° for 27 hr, neutralized with 1 *N* HCl at 0°, and extracted with ether six times. The ether solution was washed twice with water and evaporated to dryness. The 5.9 mg of crude acid was chromatographed on $SiO_2$ with hexane and benzene. Recovery was 4.0 mg of pure acid: mol wt calcd for $C_{22}H_{36}O_2$, 332.2775; found, 332.2728; ir carbonyl absorption at 1710 $cm^{-1}$. The acid was esterified with 1,1,7-trihydroperfluoroheptanol by a previously described method.[13g] The mass spectral fragmentation pattern of this ester as obtained by gcms is given in Table I.

**20ξ-Deuteriomethyl-5α-pregnane.** 5α-Pregnane-20ξ-carboxylic acid methyl ester, whose preparation is described above, was reduced with lithium aluminum hydride to its corresponding alcohol:[13d] mol wt calcd for $C_{22}H_{38}O$, 318.2922; found, 318.2946. The alcohol was converted to the tosylate (mol wt calcd for $C_{29}H_{44}O_3S$, 472.3009; found, 472.3017), and the latter was reduced with lithium aluminum deuteride to give 20ξ-deuteriomethyl-5α-pregnane, mp 112–113°; the corresponding unlabeled hydrocarbon synthesized by a different route (above) has mp 113°; mol wt calcd for $C_{22}H_{37}D$, 303.3036; found, 303.3046; partial mass spectrum in Figure 3D.

**5α-Cholanic Acid Methyl Ester.** 3,6-Dioxo-5α-cholanic acid (Mann Research Laboratories, mp 203–206° (lit.[29] 208–210°), 0.377 g, 0.97 mmol) was reduced by a modified Wolff–Kishner reaction[23] using triethanolamine as the solvent. A solution of the reaction mixture in dichloromethane was washed three times with dilute aqueous hydrochloric acid and with water and dried (magnesium sulfate). Evaporation of the solvent yielded crude 5α-cholanic acid (318 mg, 91% of theory). The acid was dissolved in a mixture of methanol (15 ml) and boron trifluoride etherate (3 ml) and heated under reflux for 2 hr. A dichloromethane solution of the reaction mixture was washed with aqueous sodium bicarbonate and water and dried over magnesium sulfate, and the solvent evaporated, yielding 5α-cholanic acid methyl ester (336 mg, 101% of theory); the latter was purified by chromatography on silica gel as described for the 5α-pregnane-20ξ-carboxylic acid methyl ester (above). Gas chromatographic coinjection of this "5α" ester with authentic 5β-cholanic acid methyl ester revealed that this product consisted of 77% of the 5α and 23% of the 5β isomer (retention time 5α/5β = 1.04); mol wt calcd for $C_{25}H_{42}O_2$, 374.3185; found, 374.3157. The 5β isomer probably was already present in the starting diketo acid, as indicated by a gas chromatogram of its methyl ester. This would not be surprising because the 5α-diketo acid is commercially prepared from 5β bile acids. An alternative would be isomerization of some 5α to 5β during the Wolff–Kishner reaction.

**5α-Cholanic Acid and 1,1,7-Trihydroperfluoroheptyl Ester.** The acid (mol wt calcd for $C_{24}H_{40}O_2$, 360.3028; found, 360.3035) was prepared by hydrolysis of the methyl ester. The fluoroalcohol ester was prepared from the acid as described previously.[13g] Gcms gave the fragmentation pattern shown in Table II; gc retention times: synthetic 5α- and 5β-fluoroalcohol esters 1.019 *vs.* 1.013 for natural product derivatives E/D.

**5α-Cholane and 24-Deuterio-5α-cholane.** The mixture of 5α- and 5β-cholanic acid methyl esters (described above) was converted to the corresponding mixture of stereoisomeric cholanols and tosylates by previously described methods;[13d] lithium aluminum deuteride was used for part of the tosylate to give (after preparative gc and sublimation) pure 24-deuterio-5α-cholane. 5α-Cholane showed the following characteristics: mp 73–74°; mol wt calcd for $C_{24}H_{42}$, 330.3286; found, 330.3296; $[\alpha]D$ 22.0 ± 2.0°; partial mass spectrum in Figure 5B; gc coinjection with petroleum-based hydrocarbon in Figure 2B. 24-Deuterio-5α-cholane showed the following characteristics: mp 73–74°; mol wt calcd for $C_{24}H_{41}D$, 331.3349; found, 331.3359; partial mass spectrum in Figure 5D.

**24-Deuterio-5β-cholane.** 24-Deuterio-5β-cholane was prepared by reduction of the previously described[13d] corresponding tosylate with lithium aluminum deuteride and purified as above: mp 92–92.5° (lit.[30] mp of 5β-cholane, 90°); mol wt calcd for $C_{24}H_{41}O$, 331.3349; found, 331.3299; partial mass spectrum in Figure 6D; gc coinjection of 5β-cholane with petroleum acid derived hydrocarbon in Figure 2C; gc retention time for 5α/5β cholane, 1.045.

**Acknowledgment.** We gratefully acknowledge Professor C. Djerassi and Dr. R. Hale, Stanford University, and also Professor M. Calvin, University of California, Berkeley, for stimulating discussions. We thank Dr. W. D. Redfield for optical rotation measurements, Dr. M. J. R. Cantow for gel permeation chromatography, Dr. W. G. Howells for some preliminary separations, and Mr. R. G. Prince for technical assistance. Grateful acknowledgment for support of this work and for the permission to publish is made to the management of Chevron Oil Field Research Co., La Habra, and Chevron Research Co., Richmond, Calif. This paper was presented in part at the Gordon Research Conference on Organic Geochemistry, Holderness, N. H., Aug 1970.

(29) H. Ishii, Y. Nozaki, T. Okumura, and D. Satoh, *Yakugaku Zasshi*, **81**, 805 (1961).

(30) R. T. Blickenstaff and F. C. Chang, *J. Amer. Chem. Soc.*, **80**, 2726 (1958).