# Structure of the side chain of the $C_{29}$ dicarboxylic bile acid occurring in infants with coprostanic acidemia

**G. Janssen,[1] S. Toppet, and G. Parmentier**

Rega Institute and Department of Chemistry, University of Leuven, Belgium

**Abstract** The structure of the side chain of the 3α,7α,12α-trihydroxy-5β-$C_{29}$ dicarboxylic bile acid occurring in body fluids of infants with coprostanic acidemia was investigated by means of mass spectrometry and nuclear magnetic resonance spectroscopy. The findings identified this bile acid as 3α,7α,12α-trihydroxy-27a,27b-dihomo-5β-cholestane-26,27b-dioic acid (3α,7α,12α-trihydroxy-27-carboxymethyl-5β-cholestan-26-oic acid).—**Janssen, G., S. Toppet, and G. Parmentier.** Structure of the side chain of the $C_{29}$ dicarboxylic bile acid occurring in infants with coprostanic acidemia. *J. Lipid Res.* 1982. **23:** 456–465.

**Supplementary key words** mass spectrometry • $^{1}H$ nuclear magnetic resonance • $^{13}C$ nuclear magnetic resonance

Recently we reported the occurrence of a new bile acid in body fluids of three infants with neurological, dysmorphic, and cholestatic symptoms (1). This bile acid predominated in serum (28–41%) and was present in only a small amount in bile (~0.5%) and urine (~5%). Mass spectrometric analysis of the acetate and TMS ether of both the methyl and ethyl ester identified the bile acid as a $C_{29}$ dicarboxylic acid consisting of a saturated 5β-bile acid nucleus with α-oriented hydroxyl groups at C-3, C-7, and C-12, and a saturated ten-carbon side chain including the two carboxyl groups, which are located beyond C-21.

As was pointed out, β-sitosterol could possibly be a precursor of the $C_{29}$ bile acid. Hydroxylation and subsequent oxidation of two terminal methyl groups of the side chain of β-sitosterol should, together with the elaboration and hydroxylation of the bile acid nucleus, lead to either 3α,7α,12α-trihydroxy-5β-stigmastane-26,27- or 26,29-dioic acid. However, gas–liquid chromatographic comparison, as dimethyl ester acetate, of the $C_{29}$ bile acid with the synthetic sample of either stigmastanedioic acid pointed to nonidentity (2).

In this report we describe the elucidation of the structure of the side chain of the $C_{29}$ bile acid.

## MATERIALS AND METHODS

### Isolation of the $C_{29}$ bile acid

The bile salts were extracted from urine by stirring about 500 ml with 100 g of Amberlite XAD-2, 300–1000 μm (Serva, Heidelberg, West Germany) for 2 hr. After filtering and washing the resin with 350 ml of water, the bile salts were eluted with 1,250 ml of methanol–25% ammonium hydroxide 99:1 (v/v). The eluate was evaporated to dryness and the residue was dissolved in 375 ml of 72% ethanol (v/v). This solution was passed through a column of 50 g of Amberlyst A-15 in the $H^+$ form (Serva) and the column was rinsed with 500 ml of 72% ethanol (v/v). Effluent and washing were percolated through a column (12.5 cm × 3 cm) of 17.5 g of DEAP-Sephadex LH-20 (prepared as described by Almé et al. (3); the dry gel had a titrated diethylamine content of 1.6 meq/g). Neutral compounds and unconjugated monocarboxylic bile acids were successively eluted from the DEAP-Sephadex LH-20 column with 800 ml of 72% ethanol (v/v) and 325 ml of a solution of 0.1 M acetic acid in 72% ethanol (v/v). The fraction of glycine conjugates, containing the unconjugated $C_{29}$ bile acid, was subsequently eluted with 600 ml of 0.3 M ammonium acetate, pH 5.0. After concentration and acidification to pH 1.0 with 6 N HCl, the $C_{29}$ bile acid was extracted with diethyl ether and the ether solution was evaporated. From 2,185 ml of urine, 81 mg of crude yellow material was obtained. This residue was esterified with diazomethane and purified by chromatography on a column (3 cm × 45 cm) of 70 g of Sephadex LH-20

Abbreviations and trivial names: cholic acid, 3α,7α,12α-trihydroxy-5β-cholan-24-oic acid; $C_{29}$ bile acid, 3α,7α,12α-trihydroxy-27a,27b-dihomo-5β-cholestane-26,27b-dioic acid; TMS, trimethylsilyl; DEAP, diethylaminohydroxypropyl; GLC, gas–liquid chromatography; MS, mass spectrometry; NMR, nuclear magnetic resonance.

[1] Address reprint requests to: G. Janssen, Ph.D, Rega Institute—K. U. Leuven, Minderbroedersstraat 10, B-3000 Leuven, Belgium.

(Pharmacia, Uppsala, Sweden) with chloroform. The fractions containing the $C_{29}$ bile acid, determined by GLC after acetylation of an aliquot of each fraction (1), were combined and evaporated to dryness. The faintly colored oil so obtained contained 925 $\mu$g of $C_{29}$ dimethyl ester as determined by GLC (1). Final purification was accomplished by high-performance liquid chromatography of the methyl ester acetate on a stainless steel column (30 cm × 4 mm) packed with $\mu$Bondapak $C_{18}$ (Waters Associates, Milford, MA) and elution with acetonitrile–water, 60:40 (v/v). A Waters liquid chromatograph equipped with a model M 6000 pump and a U6K loop injector was used. The sample was dissolved in 0.25 ml of the mobile phase and for each run 10 $\mu$l was injected. Elution was performed at 2 ml/min and fractions of 5 ml were collected. GLC indicated that the $C_{29}$ dimethyl ester acetate was eluted in fraction 6 (from 25–30 ml of eluate). The fractions were combined and evaporated. The residue was dissolved in acetone, filtered, and the solvent was evaporated. Retention times and mass spectrum of the isolated substance were identical with those obtained for the $C_{29}$ dimethyl ester acetate in a total urine sample.

### Mass spectra

Mass spectra were recorded at 270°C in the GLC–MS mode as described (1). Mass spectra of the $C_{29}$ dimethyl ester TMS ether and acetate are shown in **Fig. 1.** Mass measurements were performed on the $C_{29}$ dimethyl ester acetate by either peak matching or computer processing at a resolving power (10% valley) of 9000 using perfluorokerosene as the mass standard with a AEI MS-902S mass spectrometer equipped with a VG 2020 data system; the sample was inserted directly into the ion source at 170°C. The mass spectrum of the $C_{29}$ dimethyl ester acetate obtained by direct insertion at 170°C displays, besides a small molecular ion and M-HOAc ion, the same fragments as the spectrum acquired using GLC–MS at 270°C. Compared to the latter, the relative abundance of the peaks in the mass spectrum recorded at 170°C is, apart from a strong increase of the intensity of the M-2HOAc and M-3HOAc fragments, only slightly enhanced in the high mass region and similarly decreased at the lower mass end.

### $^1$H NMR and $^{13}$C NMR spectra

$^1$H NMR spectra were recorded on a Varian XL-100 (continuous wave mode), a Varian XL-200 (Fourier transform mode), and a Bruker WM-500 (Fourier transform mode) apparatus. $^{13}$C NMR spectra were recorded on a Bruker WP-80 instrument operating at 20.1 MHz. The $^{13}$C NMR spectrum of the $C_{29}$ dimethyl ester acetate was recorded on a Bruker WM-250 apparatus operating at 62.8 MHz (determined on less than 1 mg of substance; 219,000 accumulations). The substances were dissolved in $CDCl_3$ in 5-mm sample tubes. For the $^1$H NMR spectra, tetramethylsilane was used as internal standard and for the $^{13}$C NMR spectra, the central peak of the solvent was taken as reference and placed at $\delta$ 77.3 ppm versus tetramethylsilane.

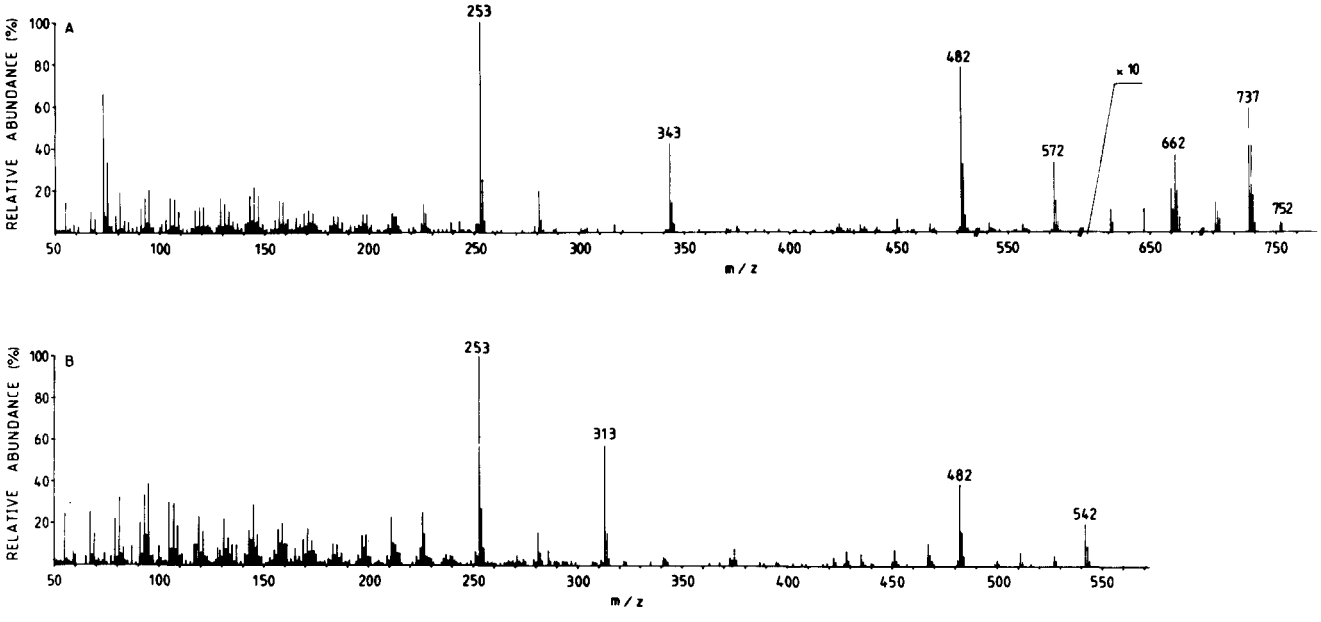



**Fig. 1.** Mass spectra of 3$\alpha$,7$\alpha$,12$\alpha$-trihydroxy-27a,27b-dihomo-5$\beta$-cholestane-26,27b-dioic acid (A) as dimethyl ester TMS ether and (B) as dimethyl ester acetate, recorded at 270°C using GLC–MS.

TABLE 1. Mass of the possible monomethyl ester McLafferty rearrangement ions arising from structure I (Scheme 1) related to the value of p and q and the substitution of the carbon atom $\alpha$ to the methyl ester group

| p, q | Carbon $\alpha$ to $-COOCH_3$ | | |
|---|---|---|---|
| | $-CH_2-$ | $>CH-$ | $-\overset{\vert}{C}-$ |
| 1 | 74 | | |
| 2 | 74 | 88 | |
| 3 | 74 | 88 | 102 |
| | | 102 | |
| 4 | 74 | 88 | 102 |
| | | 102 | 116 |
| | | 116 | |
| 5 | 74 | 88 | 102 |
| | | 102 | 116 |
| | | 116 | 130 |
| | | 130 | |

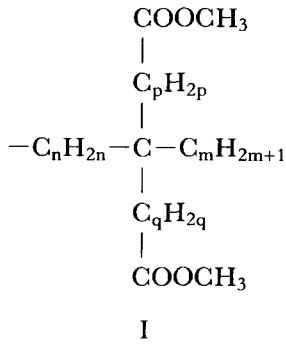
The Varian XL-200 NMR spectrum was run at the Laboratory of Organic Chemistry, University of Louvain-La-Neuve and the Bruker WM-500 and WM-250 NMR spectra were run in the Application Laboratory of Bruker GmbH, Karlsruhe, West Germany.

The carbons of the steroid nucleus were identified and assigned by comparison with the shift values reported by Leibfritz and Roberts for cholic acid (4). The tentative assignment of some side chain carbons of the $C_{27}$ and $C_{29}$ bile acids is based on the shift prediction calculated by Kelecom for saturated sterol side chains (5) and modified according to the value proposed by Levy and Nelson (6) for replacement of a methyl group by a COOR function.

## RESULTS AND DISCUSSION

The structure of side chains in steroids and tetracyclic triterpenoids can be established by stepwise degradation or by conversion to known substances. Besides, mass spectrometry (7, 8) and NMR spectroscopy provide useful structural information from functionalized side chains. Recently mass-analyzed ion kinetic energy spectrometry was shown to be applicable to the determination of the side chain structure (9).

As already pointed out, the side chain of the $C_{29}$ bile acid does reasonably consist of a saturated ten-carbon chain including the 20-methyl group (C-21) and two carboxyl groups that are located beyond C-21 (1). The elemental composition of the side chain, $C_{10}H_{17}O_4$, has been verified by exact mass measurements of the very intense ions M-2HOAc, $C_{33}H_{50}O_6$, at m/z 542 and M-3HOAc, $C_{31}H_{46}O_4$, at m/z 482 on the one hand and M-2HOAc-side chain, $C_{21}H_{29}O_2$, at m/z 313 and M-3HOAc-side chain, $C_{19}H_{25}$, at m/z 253 on the other hand, all formed on fragmentation of the dimethyl ester acetate (see Table 6 below). Hence, the C-22–C-29 part of the side chain consists of a dicarboxylic acid radical whose structure can be represented by I (as dimethyl ester, Scheme 1)



SCHEME 1

with m, n, p, and q = 0–5 and m + n + p + q = 5, in which the $C_xH_{2x}$ groups are methylene or polymethylene groups possibly substituted by straight or branched alkyl groups, and in which the alkyl group $C_mH_{2m+1}$ is straight or branched.

Mass spectra of methyl cholan-24-oates and cholestan-26-oates exhibit weak McLafferty rearrangement ions of the ester group at m/z 74 and m/z 88, respectively (10–12). Similarly, McLafferty rearrangement ions containing one or two methyl ester groups can originate from I if a $\gamma$-hydrogen is available. Ions with one ester group arise by $\beta$-cleavage in a branch bearing the ester group and their mass depends on the value of p and q and on the substitution of the carbon atom carrying the ester group (**Table 1**). Ions containing both ester groups occur at m/z 202, 188, 174, 160, 146, or 132 if p = 0 and n = 0, 1, 2, 3, 4, or 5, respectively. Furthermore, a McLafferty rearrangement can lead to elimination of an unsaturated ester and/or an olefin. Loss of 86, 100, 114, or 128 mass units corresponds to elimination of an unsaturated ester and occurs if p = 0 and q = 2, 3, 4, or 5, respectively. $\beta$-Fission expels as olefin (*a*) the $C_mH_{2m+1}$ group if p = 0 or (*b*) an alkyl branch on the carbon atom bearing the ester group.

The search for McLafferty rearrangement ions was achieved with the mass spectra of the $C_{29}$ bile acid derivatives recorded by GLC–MS of serum bile acid mixtures (1). Mass spectra of the TMS ether and acetate of the $C_{29}$ dimethyl ester did not show the loss of an unsaturated ester or olefin from either the molecular ion or the ions derived from it by successive losses of TMSOH and HOAc, respectively. Nor did these spectra display any distinct peaks of ester group-containing ions.

TABLE 2. Absolute and relative differences between the relative abundances (%) of the peaks at m/z values corresponding to theoretically possible McLafferty rearrangement ions from the $C_{29}$ dimethyl ester[a]

| | Methyl Cholate | | | | | | Methyl 3α,7α,12α-trihydroxy-5β-cholestan-26-oate | | | | | | $C_{29}$ Methyl Ethyl Ester | | | | $C_{29}$ Diethyl Ester | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Acetate (○) | | TMS (●) | | Dehydro[b] (⊘) | | Acetate (□) | | TMS (■) | | Dehydro[c] (⊠) | | Acetate (△) | | TMS (▲) | | TMS (▼) | |
| m/z | abs.[d] | rel.[e] | abs. | rel. | abs. | rel. | abs. | rel. | abs. | rel. | abs. | rel. | abs. | rel. | abs. | rel. | abs. | rel. |
| 74 | +0.3 | +7.7 | −5.7 | −36 | −2.9 | −22 | +3.0 | +250 | +1.4 | +16 | +8.9 | +685 | +1.7 | +41 | +1.3 | +13 | −6.3 | −62 |
| 88 | +0.3 | — | −0.1 | −17 | +1.2 | +1200 | −9.7 | −97 | −5.2 | −91 | −27.9 | −96 | −0.3 | −100 | +0.6 | +120 | −0.5 | −100 |
| 102 | +1.0 | — | −0.4 | −44 | −0.1 | −100 | +0.4 | +67 | −0.3 | −38 | −1.3 | −100 | −1.0 | −100 | −0.2 | −40 | −0.5 | −100 |
| 116 | −0.5 | −63 | −2.1 | −41 | −3.4 | −72 | −1.1 | −79 | +0.4 | +15 | −0.7 | −35 | +2.6 | +867 | +1.6 | +53 | −3.0 | −100 |
| 130 | −0.8 | −18 | −1.8 | −32 | −4.4 | −85 | −1.2 | −25 | +0.2 | +5.4 | −1.2 | −60 | −3.7 | −100 | +1.8 | +46 | −2.0 | −51 |
| 144 | | | | | | | | | | | | | −0.3 | −2.5 | +2.3 | +32 | −1.3 | −18 |
| 132 | −0.5 | −6.7 | −0.2 | −4.5 | −0.4 | −7.7 | −0.5 | −6.7 | +0.5 | +14 | +1.8 | +60 | −3.5 | −50 | +0.4 | +9.5 | −2.3 | −55 |
| 146 | +0.5 | +8.1 | +0.6 | +10 | −1.8 | −25 | +0.2 | +3.1 | +1.6 | +32 | +1.6 | +43 | +1.5 | +22 | −0.3 | −4.6 | −0.8 | −12 |
| 160 | +5.2 | +108 | +1.1 | +29 | +80 | +410 | +5.1 | +104 | +1.4 | +40 | +87 | +694 | −9.4 | −94 | +2.2 | +45 | +0.9 | +18 |
| 174 | +4.0 | +235 | +3.0 | +158 | −4.8 | −51 | +1.4 | +33 | +1.2 | +32 | −2.7 | −37 | +4.9 | +86 | +2.5 | +51 | −4.9 | −100 |
| 188 | −0.2 | −12 | −0.2 | −22 | +2.5 | +74 | +0.1 | +7.1 | −0.3 | −30 | +2.2 | +60 | +0.9 | +60 | +1.6 | +229 | +7.0 | +1000 |
| 202 | +0.9 | +300 | −0.4 | −44 | −1.4 | −41 | +0.6 | +100 | −0.2 | −29 | +0.2 | +11 | −1.2 | −100 | +0.6 | +120 | −0.5 | −100 |
| 216 | | | | | | | | | | | | | +1.4 | +140 | +0.6 | — | | |
| 232 | | | | | | | | | | | | | | | | | +1.2 | +171 |

[a] The differences were calculated by referring the $C_{29}$ dimethyl ester to methyl cholate and methyl 3α,7α,12α-trihydroxy-5β-cholestan-26-oate and by referring the $C_{29}$ methyl ethyl ester and diethyl ester to the $C_{29}$ dimethyl ester. For example, the figures in the first column and the first row (an absolute increase of 0.3 corresponding to a relative increase of 7.7%) are obtained by referring the relative abundance (%) of m/z 74 in the mass spectrum of the $C_{29}$ dimethyl ester acetate (4.2) to that of m/z 74 in the mass spectrum of the acetate of methyl cholate (3.9). Except for the trioxo-$C_{29}$ dimethyl ester and the derivatives of the $C_{29}$ methyl ethyl and diethyl ester, the relative abundances (%) are the average of two values, corrected for background, from mass spectra recorded across the top of the gas chromatographic peak, monitored by the total ion current, in different chromatographic runs. In the mass spectra of the acetates and TMS ethers, m/z 253 was the base peak, except for the TMS ether of methyl cholate for which m/z 253 (relative abundance 91%; base peak m/z 73) was taken as base peak. Methyl 3,7,12-trioxo-5β-cholan-24-oate and 3,7,12-trioxo-5β-cholestan-26-oate exhibited the base peak at m/z 261. Mostly the absolute intensity of the base peak largely exceeded $1.7 \times 10^{-7}$ A electron multiplier current corresponding to 1 cm on the × 1 scale of the mass spectrum. Only in the case of the trioxo-$C_{29}$ dimethyl ester, the $C_{29}$ methyl ethyl ester acetate, and the $C_{29}$ diethyl ester TMS ether the absolute intensity was about half that value. The absolute intensity of the $C_{29}$ diethyl ester acetate was too low to be accurately measured. All mass spectra were recorded under the same GLC–MS conditions. The symbols are used in Fig. 2.

[b] Methyl 3,7,12-trioxo-5β-cholan-24-oate.

[c] Methyl 3,7,12-trioxo-5β-cholestan-26-oate.

[d] Absolute.

[e] Relative, expressed in percent.

Likewise, the mass spectrum of the trioxo-$C_{29}$ dimethyl ester did not show any of the molecule losses but, surprisingly, exhibited a base peak at m/z 160.

The low mass range in the spectra of bile acid methyl esters and their derivatives consists of ring system fragments spiked with a few, generally weak, characteristic ions. Hence, comparison of the relative abundance of the peaks at m/z values of the McLafferty rearrangement ions in the mass spectra of the derivatives of the $C_{29}$ dimethyl ester with those of the same derivatives of methyl cholate and methyl 3$\alpha$,7$\alpha$,12$\alpha$-trihydroxy-5$\beta$-cholestan-26-oate, both having the same nucleus as the $C_{29}$ bile acid, may reveal small, but significant differences. These differences may also be observed by comparing the $C_{29}$ methyl ethyl and diethyl ester with the dimethyl ester. The results of this scrutiny are summarized in **Table 2** and **Fig. 2.** Clearly, only those ions with the same mass number whose relative abundances differ sufficiently from one another, both in absolute and in relative value, are significant (i.e., the absolute and relative increase or decrease of the relative abundance). An increase of the relative abundance indicates that the

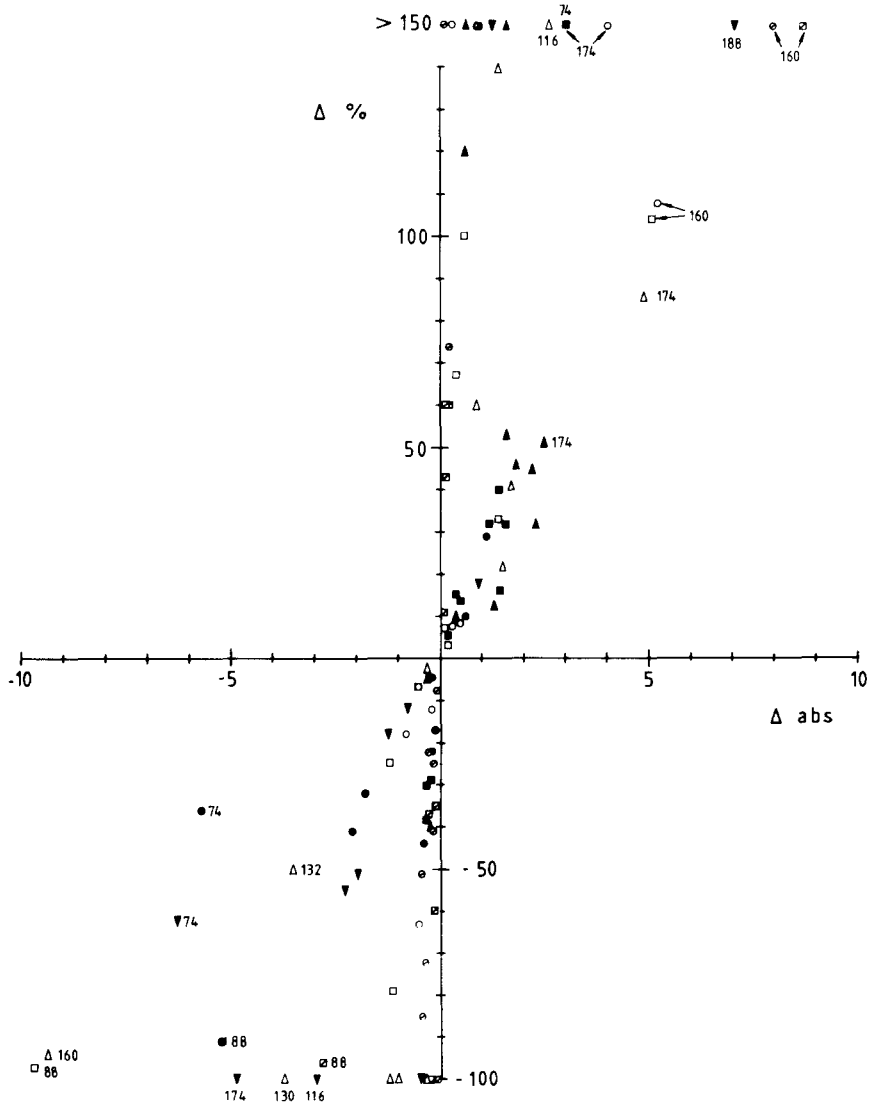



**Fig. 2.** Graph relating the absolute (abscissa) and relative (ordinate) differences summarized in Table 2. Symbols indicating the difference of the $C_{29}$ dimethyl ester derivatives and: methyl cholate acetate (○); methyl cholate TMS ether (●); methyl 3,7,12-trioxo-5$\beta$-cholan-24-oate (⊘); methyl 3$\alpha$,7$\alpha$,12$\alpha$-triacetoxy-5$\beta$-cholestan-26-oate (□); methyl 3$\alpha$,7$\alpha$,12$\alpha$-trihydroxy-5$\beta$-cholestan-26-oate TMS ether (■); methyl 3,7,12-trioxo-5$\beta$-cholestan-26-oate (⧄); $C_{29}$ methyl ethyl ester acetate (△); $C_{29}$ methyl ethyl ester TMS ether (▲); and $C_{29}$ diethyl ester TMS ether (▼). The absolute differences from comparison of the trioxo-compounds have to be multiplied by 10. The numbers indicate the mass of the ions exceeding a threshold intensity change. These ions are compiled in Table 3.

ion is present only in the compound which is being compared, whereas a decrease indicates its presence only in the reference compound.

**Table 3** comprises the ions which either increase or decrease with an absolute increment ≥ 2.5 for the acetates and TMS ethers, ≥ 25 for the trioxo-compound, and with a relative increment ≥ 50%. These ions are indicated by their mass in Fig. 2. The increase frequency of the m/z 160 ion, shifted to m/z 174 and m/z 188 in the methyl ethyl and diethyl ester, respectively, is very significant. Besides, the shift to m/z 174 is accompanied by a decrease of m/z 160. Obviously, the occurrence of a dimethyl ester ion at m/z 174, showing only twice an increase and once a decrease, is much less probable. The same certainly applies to m/z 132. The frequent appearance of the m/z 160 ion almost excludes the structures that cannot give rise to a dimethyl ester ion.

The appearance of the m/z 74 ion in Table 3, increased versus methyl triacetoxy-cholestanoate and decreased versus the $C_{29}$ diethyl ester, advocates the presence of the $-CH_2COOCH_3$ group. This is further supported by the absence of a decrease of the m/z 74 ion in comparison with two derivatives of methyl cholate. Its decrease versus methyl cholate TMS ether is caused by the much higher abundance of the TMS cation at m/z 73 and hence of its isotopic abundance at m/z 74 in the mass spectrum of the latter. The very distinct decrease of the monomethyl ester ion at m/z 88 versus the methyl cholestanoates suggests the absence of a 1-methoxycarbonyl-ethyl group. A McLafferty rearrangement ion at m/z 116 in the decrease column should contain a methyl ester group, whereas that in the increase column should comprise an ethyl ester group which corresponds to a methyl ester ion of 102 mass units. However, neither of these monomethyl ester ions can occur together with a dimethyl ester ion at m/z 160, and only the m/z 102 ion can be formed together with m/z 174; its occurrence is thus improbable. Likewise, a monomethyl ester ion at m/z 130 cannot arise together with m/z 160 and m/z 174. Hence, with each of the latter dimethyl ester ions, the intensity change of the ions at m/z 116 and m/z 130 is meaningless.

In conclusion, the McLafferty rearrangement of the side chain most probably yields the dimethyl ester ion at m/z 160 (n = 3, p = 0 structure I, Scheme 1) whether with concurrent formation of the monomethyl ester ion at m/z 74 or not. This reduces the number of possible structures to 16, 10 of which can also give rise to m/z 74.

Obviously further investigation of the structure of the side chain necessitated the isolation of the $C_{29}$ bile acid. For lack of a sufficient amount of serum that contained a considerable relative quantity, the isolation was accomplished from urine. The identity between the $C_{29}$ bile acid in serum and in urine was established by the same gas–liquid chromatographic retention times and identical mass spectrum of their dimethyl ester acetate (1). As in serum, the $C_{29}$ bile acid was largely unconjugated in urine. Since straight chain dicarboxylic acids, apart from malonic acid, have $pK_a$[1] values of 4.2–4.6, the $C_{29}$ bile acid might elute from DEAP-Sephadex LH-20 with the glycine conjugates, which have $pK_a$ values of 3.8–4.8.

Urine was desalted with Amberlite XAD-2 whereupon the bile salts were converted to the $H^+$ form on Amberlyst A-15 and separated in groups on DEAP-Sephadex LH-20. The fraction of glycine conjugates was

TABLE 3. Compilation of fragment ions corresponding in mass to theoretically possible McLafferty rearrangement ions from the $C_{29}$ dimethyl ester and exceeding a threshold intensity change on comparison of the $C_{29}$ dimethyl ester with methyl cholate, methyl 3α,7α,12α-trihydroxy-5β-cholestan-26-oate, the $C_{29}$ methyl ethyl ester, and the $C_{29}$ diethyl ester.

| | | Increase | | | | Decrease | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Monoester Ion | | Diester Ion | | Monoester Ion | | | | Diester Ion | | | |
| | | m/z | | m/z | | m/z | | | | m/z | | | |
| $C_{29}$ dimethyl ester vs. methyl cholate | acetate | | | 160 | 174 | | | | | | | | |
| | TMS ether | | | | 174 | (74) | | | | | | | |
| | dehydro[a] | | | 160 | | | | | | | | | |
| $C_{29}$ dimethyl ester vs. methyl 3α,7α,12α-trihydroxy-5β-cholestan-26-oate | acetate | 74 | | 160 | | | 88 | | | | | | |
| | TMS ether | | | | | | 88 | | | | | | |
| | dehydro[b] | | | 160 | | | 88 | | | | | | |
| $C_{29}$ methyl ethyl ester vs. $C_{29}$ dimethyl ester | acetate | | 116 | 174 | | | | | 130 | 132 | 160 | | |
| | TMS ether | | | 174 | | | | | | | | | |
| $C_{29}$ diethyl ester vs. $C_{29}$ dimethyl ester | TMS ether | | | 188 | | 74 | | 116 | | | | | 174 |

[a] Methyl 3,7,12-trioxo-5β-cholan-24-oate.
[b] Methyl 3,7,12-trioxo-5β-cholestan-26-oate.

TABLE 4. $^1$H NMR data of methyl cholate acetate ($C_{24}$-OAc), methyl 3$\alpha$,7$\alpha$,12$\alpha$-trihydroxy-5$\beta$-cholestan-26-oate ($C_{27}$-OH) and its acetate ($C_{27}$-OAc), and the $C_{29}$ dimethyl ester acetate ($C_{29}$-OAc)[a]

| Proton | $C_{24}$-OAc | $C_{27}$-OH | $C_{27}$-OAc | $C_{29}$-OAc |
|---|---|---|---|---|
| H-12 | 5.09 (t, 6 Hz)[b] | 3.98 (t, 6 Hz) | 5.10 (t, 6 Hz) | 5.10 (t, 6 Hz) |
| H-7 | 4.92 (q, 9 Hz) | 3.85 (q, 9 Hz) | 4.92 (q, 9 Hz) | 4.92 (q, 9.5 Hz) |
| H-3 | 4.60 (t.t, 31 Hz) | 3.45 (t.t, 31 Hz) | 4.60 (t.t, 31 Hz) | 4.60 (t.t, 31 Hz) |
| $CH_3CO$ | 2.13, 2.08, 2.03 (s) | | 2.13, 2.09, 2.05 (s) | 2.14, 2.10, 2.06 (s) |
| HO | | 2.35 (broad) | | |
| $CH_3$-21 | 0.82 (d, 6 Hz) | 0.95 (d, 6.5 Hz) | 0.79 (d, 6.5 Hz) | 0.79 (d, 6.5 Hz) |
| $CH_3$-19 | 0.92 (s) | 0.89 (s) | 0.92 (s) | 0.92 (s) |
| $CH_3$-18 | 0.74 (s) | 0.68 (s) | 0.72 (s) | 0.72 (s) |
| $CH_3OOCCH$-R | R = H, 2.32 (m) | R = $CH_3$, 1.14 ($CH_3$, d, 7 Hz) 2.44 (sx, 35 Hz) | R = $CH_3$, 1.14 ($CH_3$, d, 7 Hz) 2.44 (sx, 35 Hz) | R = H, 2.29, 2.35 R = $CH_2$, 2.40 |
| $COOCH_3$ | 3.66 (s) | 3.67 (s) | 3.67 (s) | 3.67, 3.675 (s) |

[a] In $CDCl_3$ at 100 MHz: $C_{24}$-OAc; 200 MHz: $C_{27}$-OAc; 500 MHz: $C_{27}$-OH and $C_{29}$-OAc.
[b] Chemical shifts are in parts per million ($\delta$). Figures in parentheses: multiplicity, sum of the vicinal couplings expressed in hertz (Hz); s, singlet; d, doublet; t, triplet; q, quartet; sx, sextet; m, multiplet.

acidified and the $C_{29}$ bile acid was extracted. After esterification, the $C_{29}$ dimethyl ester was eluted from Sephadex LH-20 with chloroform and purification was completed by reversed phase chromatography of the dimethyl ester acetate. As a result of the small amount of substance, chemical characterization was not possible, but NMR spectroscopy might yield valuable structural information.

A $^1$H spectrum of the $C_{29}$ dimethyl ester acetate was run at 500 MHz and a $^{13}$C spectrum was run at 62.8 MHz. Comparison of the chemical shifts with those of the acetate of methyl cholate and methyl 3$\alpha$,7$\alpha$,12$\alpha$-triacetoxy-5$\beta$-cholestan-26-oate confirm the presence of the 5$\beta$-cholane ring system up to and including C-21 and $\alpha$-substitution of the acetoxyl groups at C-3, C-7, and C-12 (**Tables 4 and 5**). The $^1$H chemical shifts of the C-18 and C-19 methyl groups at $\delta \sim 0.70$ and $\delta \sim 0.90$ ppm, respectively, are typical for 5$\beta$-bile acids (13) and more generally for steroids in the 5$\beta$,14$\alpha$-configuration

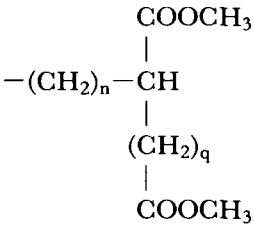
TABLE 5. $^{13}$C Chemical shifts of the acetate of methyl cholate ($C_{24}$-OAc), of methyl 3$\alpha$,7$\alpha$,12$\alpha$-trihydroxy-5$\beta$-cholestan-26-oate ($C_{27}$-OAc), and of the $C_{29}$ dimethyl ester ($C_{29}$-OAc)[a]

| | $C_{24}$-OAc | | $C_{27}$-OAc | $C_{29}$-OAc |
|---|---|---|---|---|
| | Solvent | | Solvent | Solvent |
| Carbon | Dioxane[b] | $CDCl_3$[c] | $CDCl_3$[c] | $CDCl_3$[c] |
| C-18 | 12.8 | 12.4 | 12.5 | 12.5 |
| C-21 | 18.5 | 17.7 | 18.1 | 18.2 |
| C-19 | 23.1 | 22.7 | 22.8 | 22.8 |
| C-15 | 23.5 | 23.0 | 23.1 | 23.1 |
| C-5 | 41.5 | 41.2 | 41.3 | 41.3 |
| C-14 | 43.9 | 43.6 | 43.7 | 43.7 |
| C-13 | 45.7 | 45.3 | 45.4 | 45.4 |
| C-17 | 48.4 | 47.6 | 48.0 | 48.0 |
| C-7 | 71.2 | 71.0 | 71.1 | 71.1 |
| C-3 | 74.3 | 74.3 | 74.4 | 74.4 |
| C-12 | 75.6 | 75.6 | 75.8 | 75.8 |
| $CH_3$-$\underline{C}OO$ | 169.3 | 170.7 | 170.8, 170.7 | 170.8, 170.7 |
| $\underline{C}H_3$-COO | 20.8 | 21.6, 21.7 | 21.6, 21.8 | 21.7, 21.9 |
| $\underline{C}OOCH_3$ | 175.4 | 174.8 | 177.7 | 176.4, 173.7 |
| C-23[d] | | | 24.0 | 23.5 |
| C-25[d] | | | | 45.0 |
| C-27[d] | | | 17.2, 17.5 (R, S) | |

[a] Chemical shifts are in parts per million ($\delta$). $C_{24}$-OAc and $C_{27}$-OAc recorded at 20.1 MHz; $C_{29}$-OAc at 62.8 MHz.
[b] See reference (4).
[c] This work; $\delta$ values between 25 and 40 ppm are not reported, assignment being difficult in that crowded region.
[d] Tentative, see text.

(14). The shift of the C-21 methyl group is less typical since its value in the 3$\alpha$,7$\alpha$,12$\alpha$-trihydroxy-5$\beta$-bile acids ($\delta \sim 0.95$ ppm) differs from that in their acetates ($\delta \sim 0.80$ ppm).

Furthermore these spectra provide the following information. First, the molecule contains no other methyl groups than those attached to C-10, C-13, and C-20. This means that in the general formula I (Scheme 1) the $C_xH_{2x}$ groups are straight, $(CH_2)_x$, and m = 0. This fact eliminates all but eleven structures. Second, the two methoxycarbonyl groups are different as well in $^1H$ as in $^{13}C$ NMR. The $^{13}C$ $\delta$ values of these groups differ by about 3 ppm. This value is related to the number of $\beta$-alkyl substituents, the absorption range being $\delta$ 175–178 for $>CH\underline{C}OOCH_3$ and $\delta$ 172–175 for $-CH_2\underline{C}OOCH_3$ (Ref. 15 and Table 5: compare the chemical shift of the carbonyl carbons in methyl cholate acetate and methyl 3$\alpha$,7$\alpha$,12$\alpha$-triacetoxy-5$\beta$-cholestan-26-oate recorded in $CDCl_3$). This indicates the presence of the $>CH\underline{C}OOCH_3$ ($\delta$, 176.4) and $-CH_2\underline{C}OOCH_3$ ($\delta$, 173.7) structure and puts p = 0 and q $\geq$ 1.




II n = 0, q = 5
III n = 1, q = 4
IV n = 2, q = 3
V n = 3, q = 2
VI n = 4, q = 1

SCHEME 2

Based on these NMR data, five structures are still possible (formulas II–VI, Scheme 2) two of which can be excluded by careful examination of the pattern displayed by the three protons absorbing between $\delta$ 2.2 and 2.5 ppm, i.e., $>C\underline{H}COOCH_3$ and $-C\underline{H}_2COOCH_3$ (**Fig. 3**). The methine proton ($\delta$ 2.4 ppm) and the methylene protons ($\delta$ 2.29 and 2.35 ppm) constitute two isolated, slightly overlapping, spin systems and this makes their analysis possible at 500 MHz. The methine proton in $>CHCOOCH_3$ is a *12 lines* multiplet due to coupling with four neighboring protons: d (9 Hz) × d (8 Hz) × t (5.5 Hz). The sum of the four vicinal coupling constants is 28 Hz. Structure II in which that proton has only three neighbors is eliminated.

The methylene protons in $-CH_2COOCH_3$ are nonequivalent ($\Delta\delta$ 0.06 ppm) due to the proximity of the asymmetric carbon. They give rise to a *16 lines* pattern (15 of which are separated) and constitute the AB part

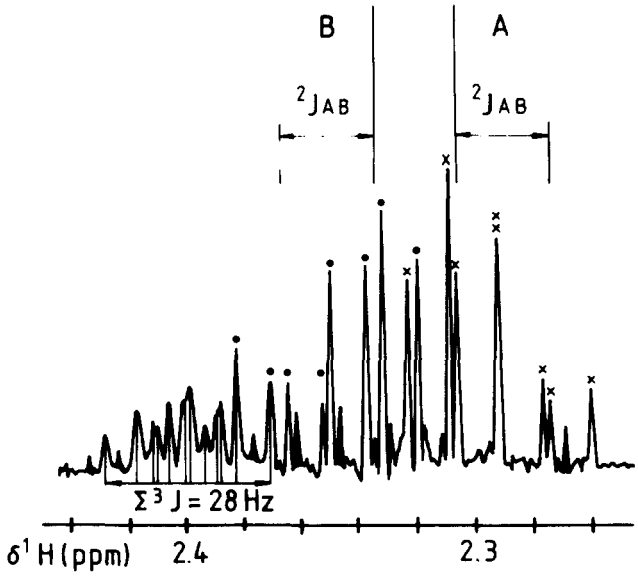



**Fig. 3.** Partial $^1H$ NMR spectrum (500 MHz with resolution enhancement) of the $C_{29}$ dimethyl ester acetate. The methine proton, $>C\underline{H}COOCH_3$, is indicated by vertical lines and the methylene protons, $-C\underline{H}_2COOCH_3$, by × and • for A and B protons, respectively. The two low-field lines of proton B are superposed on the two up-field lines of the methine proton. Impurities are shaded.

of an ABXY spin system: $^2J_{AB} = -16$ Hz, $^3J_{AX} + {}^3J_{AY} = 15.5$ Hz and $^3J_{BX} + {}^3J_{BY} = 14.5$ Hz. The high value of the coupling constant between the protons absorbing at 2.29 and 2.35 ppm (16 Hz) is never reached by a vicinal coupling between aliphatic protons and identifies the geminal methylene protons. Moreover, this value is entirely consistent with values reported for $^2J$ of a methylene group attached to one unsaturated carbon atom (16). These geminal methylene protons are vicinally coupled to two hydrogen nuclei. Therefore structure VI, in which the protons of the methylene group attached to the ester group have only one vicinal proton, is also excluded. Moreover, in the latter structure the three protons of the $-C\underline{H}(COOCH_3)C\underline{H}_2COOCH_3$ radical are expected to display a strongly coupled ABCXY pattern which is not observed.

The anisochronism of the methylene protons in $-CH_2COOCH_3$ is in favor of structure V, but structures III and IV cannot definitely be excluded by NMR spectroscopy.

The recognition of absence of a methyl group in the C-22–C-29 part of the side chain and of occurrence of the m/z 160 ion in the mass spectra of the derivatives of the $C_{29}$ dimethyl ester also point to structure V.

This evidence was verified by determination of the elemental composition of the peaks at m/z 160, 174, and 188, corresponding to the McLafferty rearrangement ions which, based on NMR data, were still possible. Likewise, the exact mass of peaks at m/z values of the ions possibly originating from the dimethyl ester ions by

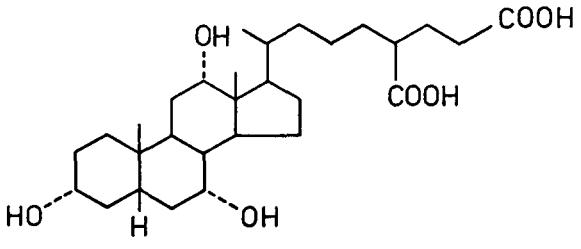
TABLE 6. Calculated and found masses of fragments occurring in the mass spectrum of the $C_{29}$ dimethyl ester acetate: 1. the mass of the peaks at m/z values corresponding to theoretically possible McLafferty rearrangement ions and their daughter ions (determined by peak matching) and 2. the mass of some major ions (determined by computer processing)

| m/z | Relative Abundance % | Structure | Elemental Composition | Mass | |
|---|---|---|---|---|---|
| | | | | Calcd. | Found[a] |
| 74 | 4.2 | $[HOC(OCH_3)=CH_2]^{+\cdot}$ | $C_3H_6O_2$ | 74.0368 | 74.0363<br>74.0152 |
| 160 | 10 | $[HOC(OCH_3)=CH(CH_2)_2COOCH_3]^{+\cdot}$ | $C_7H_{12}O_4$ | 160.0736 | 160.0730<br>160.1232 |
| 129 | 7.4 | $[160\text{-}CH_3O]^+$ | $C_6H_9O_3$ | 129.0552 | 129.0691 |
| 128 | 7.9 | $[160\text{-}CH_3OH]^{+\cdot}$ | $C_6H_8O_3$ | 128.0473 | 128.0497<br>128.0622 |
| 174 | 5.7 | $[HOC(OCH_3)=CH(CH_2)_3COOCH_3]^{+\cdot}$ | $C_8H_{14}O_4$ | 174.0892 | 173.9877<br>174.1368 |
| 143 | 16 | $[174\text{-}CH_3O]^+$ | $C_7H_{11}O_3$ | 143.0708 | 142.9941<br>143.0872 |
| 142 | 4.7 | $[174\text{-}CH_3OH]^{+\cdot}$ | $C_7H_{10}O_3$ | 142.0630 | 142.0773 |
| 188 | 1.5 | $[HOC(OCH_3)=CH(CH_2)_4COOCH_3]^{+\cdot}$ | $C_9H_{16}O_4$ | 188.1049 | 188.1538 |
| 157 | 17 | $[188\text{-}CH_3O]^+$ | $C_8H_{13}O_3$ | 157.0865 | 157.1016 |
| 156 | 5.3 | $[188\text{-}CH_3OH]^{+\cdot}$ | $C_8H_{12}O_3$ | 156.0786 | 156.0939 |
| 542 | 21 | $[M\text{-}2HOAc]^{+\cdot}$ | $C_{33}H_{50}O_6$ | 542.3607 | 542.3603 |
| 482 | 39 | $[M\text{-}3HOAc]^{+\cdot}$ | $C_{31}H_{46}O_4$ | 482.3396 | 482.3387 |
| 313 | 58 | [M-2HOAc-side chain]$^+$ | $C_{21}H_{29}O_2$ | 313.2167 | 313.2163 |
| 253 | 100 | [M-3HOAc-side chain]$^+$ | $C_{19}H_{25}$ | 253.1956 | 253.1951 |

[a] The presence of two accurate masses at the same nominal mass indicates that the low resolution ion is a composite one.

loss of a methoxyl radical or methanol was determined. Finally, the exact mass of the m/z 74 rearrangement ion, which now must be present, was verified.

The results of these measurements, obtained with the purified $C_{29}$ dimethyl ester acetate, are compiled in **Table 6.** Clearly, only the mass of the dimethyl ester ion at m/z 160 and of the concurrent monomethyl ester ion at m/z 74 match the respective calculated values. Hence, the $C_{29}$ bile acid is formulated as 3$\alpha$,7$\alpha$,12$\alpha$-trihydroxy-27a,27b-dihomo-5$\beta$-cholestane-26,27b-dioic acid (3$\alpha$,7$\alpha$,12$\alpha$-trihydroxy-27-carboxymethyl-5$\beta$-cholestan-26-oic acid) (VII).




*VII*

The configuration at C-25 remains to be determined.

We are very grateful to Dr. W. Hull, Bruker GmbH, Karlsruhe, West Germany and to Mrs. A. Mockel, Laboratory of Organic Chemistry, University of Louvain-La-Neuve for recording the high field $^1H$ and $^{13}C$ NMR spectra, and to Mr. R. De Boer, Department of Chemistry, University of Leuven, for the mass measurements.

*Manuscript received 28 May 1981 and in revised form 29 October 1981.*

## REFERENCES

1. Parmentier, G. G., G. A. Janssen, E. A. Eggermont, and H. J. Eyssen. 1979. $C_{27}$ Bile acids in infants with coprostanic acidemia and occurrence of a 3$\alpha$,7$\alpha$,12$\alpha$-trihydroxy-5$\beta$-$C_{29}$ dicarboxylic bile acid as a major component in their serum. *Eur. J. Biochem.* **102:** 173–183.
2. Janssen, G., and G. Parmentier. 1981. A further study of the bile acids in infants with coprostanic acidemia. *Steroids.* **37:** 81–89.
3. Almé, B., A. Bremmelgaard, J. Sjövall, and P. Thomassen. 1977. Analysis of metabolic profiles of bile acids in urine using a lipophilic anion exchanger and computerized gas–liquid chromatography–mass spectrometry. *J. Lipid Res.* **18:** 339–362.
4. Leibfritz, D., and J. D. Roberts. 1973. Nuclear magnetic resonance spectroscopy. Carbon-13 spectra of cholic acids and hydrocarbons included in sodium desoxycholate solutions. *J. Am. Chem. Soc.* **95:** 4996–5003.

5. Kelecom, A. 1980. Use of calculated carbon chemical shift tables of saturated C-7 to C-11 sterol side chains for structure elucidation. *Bull. Soc. Chim. Belg.* **89:** 343–352.
6. Levy, G. C., and G. L. Nelson. 1972. Carbon-13 Nuclear Magnetic Resonance for Organic Chemists. Wiley-Interscience, New York. 47.
7. Budzikiewicz, H. 1972. Steroids. *In* Biochemical Applications of Mass Spectrometry. G. R. Waller, editor. Wiley-Interscience, New York. 251–289.
8. Enzell, C. R., R. A. Appleton, and I. Wahlberg. 1972. Terpenes and terpenoids. *In* Biochemical Applications of Mass Spectrometry. G. R. Waller, editor. Wiley-Interscience, New York. 351–385.
9. Bozorgzadeh, M. H., A. G. Brenton, J. L. Wiebers, and J. H. Beynon. 1979. Isolation and structural analysis by mass analysed ion kinetic energy spectroscopy of diverse sidechains from steroids of similar tetracyclic carbon skeletons. *Biomed. Mass Spectrom.* **6:** 340–344.
10. Bergström, S., R. Ryhage, and E. Stenhagen. 1961. Mass spectrometric studies on sterols and bile acids. *Sven. Kem. Tidskr.* **73:** 566–581.
11. Sjövall, J., P. Eneroth, and R. Ryhage. 1971. Mass spectra of bile acids. *In* The Bile Acids. Chemistry, Physiology and Metabolism. Vol. 1. P. P. Nair and D. Kritchevsky, editors. Plenum Press, New York. 209–248.
12. Elliott, W. H. 1972. Bile acids. *In* Biochemical Applications of Mass Spectrometry. G. R. Waller, editor. Wiley-Interscience, New York. 291–312.
13. Small, D. M. 1971. The physical chemistry of cholanic acids. *In* The Bile Acids. Chemistry, Physiology and Metabolism. Vol. 1. P. P. Nair and D. Kritchevsky, editors. Plenum Press, New York. 262–265.
14. Bhacca, N. S., and D. H. Williams. 1964. Applications of NMR Spectroscopy in Organic Chemistry. Holden-Day, Inc., San Francisco. 17.
15. Couperus, P. A., A. D. M. Clague, and J. P. C. M. Van Dongen. 1978. Carbon-13 chemical shifts of some model carboxylic acids and esters. *Org. Magn. Res.* **11:** 590–597.
16. Bothner-By, A. A. 1965. Geminal and vicinal proton-proton coupling constants in organic compounds. *In* Advances in Magnetic Resonance. Vol. 1. J. S. Waugh, editor. Academic Press, New York. 215.