Biochem. J. (1972) **126**, 1161–1170
Printed in Great Britain

Comparative Studies of Bile Salts

A NEW TYPE OF BILE SALT FROM *ARAPAIMA GIGAS* (CUVIER) (FAMILY OSTEOGLOSSIDAE)

By G. A. D. HASLEWOOD and L. TÖKÉS
Guy's Hospital Medical School, London SE1 9RT, U.K., and Institute of Organic Chemistry, Syntex Research, Palo Alto, Calif. 94304, *U.S.A.*

(*Received* 12 *October* 1971)

1. *Arapaima gigas* bile salts were hydrolysed by alkali or cleaved with dioxan–trichloroacetic acid to give cholic acid, arapaimic acid, arapaimol-A and arapaimol-B. 2. I.r., n.m.r. and mass spectroscopy and $[\alpha]_D$ measurements indicated that arapaimic acid and arapaimol-A and -B are respectively 2α,3α,7α,12α-tetrahydroxy-5β,25ε-cholestan-26-oic acid, 5β,25*R*-cholestane-2β,3α,7α,12α,26-pentol and 5β-cholestane-2β,3α,7α,12α,26,27-hexol. 3. Partial synthesis of 2β,3α,7α,12α-tetrahydroxy-5α- and -5β-cholan-24-oic acid and their spectral examination fully confirmed these conclusions. 4. *A. gigas* bile salts show primitive features in that they comprise alcohol sulphates and a C_{27} acid; they are also specialized in showing 2β-hydroxylation.

The osteoglossid fishes comprise a number of mainly freshwater forms which Greenwood *et al.* (1966) classify as Division II of the teleosts and which, according to these authors, could not 'possibly have been involved in the ancestry of other teleosteans'. Through the kindness of Dr. P. H. Greenwood we have been able to examine the bile salts of *Arapaima gigas* (Cuvier) and we now report on our findings, partially described in a preliminary communication (Haslewood & Tökés, 1970).

Results

Bile salts of the arapaima on t.l.c. showed at least four chief spots, one corresponding in mobility to taurocholate. The i.r. spectrum suggested only that the bile salts might be steroids mainly or wholly of the 5β configuration.

The usual alkaline hydrolysis produced much sulphate ion but provided only small quantities of neutral material, together with an acidic fraction that included cholic acid (3α,7α,12α-trihydroxy-5β-cholan-24-oic acid), identified by t.l.c. and g.l.c., and 'arapaimic acid', whose chemistry we describe below.

Dioxan–trichloroacetic acid cleavage gave neutral material which was readily separated on Celite columns into two crystalline alcohols, 'arapaimol-A' and 'arapaimol-B', both of which gave an intense purple response in the Hammarsten (HCl) test. The i.r. spectra of these alcohols (Fig. 1*a*) were almost identical and showed all the peaks between 9 and 15 μm seen in the spectrum of arapaimic acid. These peaks do not correspond with those we recognize as being due to the steroid nucleus of cholic acid or its 5α-epimer, allocholic acid (Haslewood, 1967); in particular the arapaimol spectra showed peaks at about 11.04 and 11.25 μm, corresponding neither to the characteristic cholic acid (10.95 μm) nor to allocholic acid (11.20 μm) peaks in this spectral region. We were thus not able to decide from this evidence whether the new arapaima bile alcohols and acid had the 5α or 5β configuration.

The mass spectrum of arapaimol-A exhibited a molecular ion of 452 mass units, corresponding to an empirical formula $C_{27}H_{48}O_5$, and fragment ions of *m/e* 434, 416, 398 and 380. Acetylation of this compound gave a tetra- and a penta-acetate derivative (not crystallized), which showed molecular ions of 620 and 662 mass units respectively.

The mass spectrum of arapaimol-B was very similar to that of arapaimol-A but it showed no molecular ion. The peak of highest *m/e* was due to an M^+-H_2O fragment (*m/e* 450), which was accompanied by subsequent dehydration products (*m/e* 432, 414 and 396). The molecular weight of arapaimol-B (468, corresponding to $C_{27}H_{48}O_6$) was established by the mass spectra of non-crystalline penta- and hexa-acetate derivatives, which exhibited molecular-ion peaks at *m/e* 678 and 720 respectively.

These results suggested that arapaimol-A and -B are penta- and hexa-hydroxycholestane respectively. The similarity of the fragmentation patterns provided further information about the location of the hydroxyl groups. Both spectra showed a fragment ion of 305 mass units (accompanied by subsequent dehydration products of *m/e* 287, 269 and 251), which is formed by the loss of the C-17 side chain and a

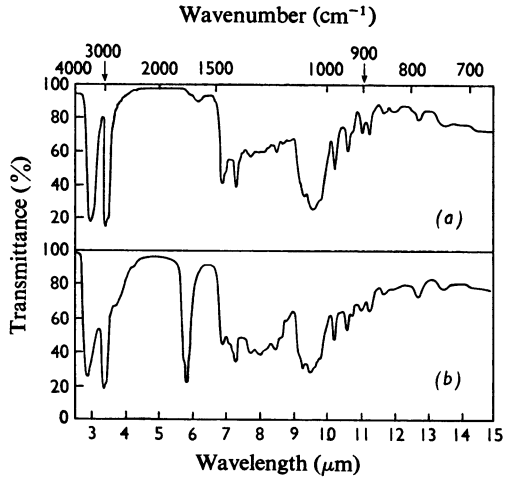


Fig. 1. *Infrared spectra*

I.r. spectra (KBr disc) of (*a*) arapaimol-A (IX) and (*b*) 2β,3α,7α,12α-tetrahydroxy-5β-cholan-24-oic acid (2β-hydroxycholic acid, VIII).

molecule of water from the steroid nucleus. This fragmentation pattern is characteristic for steroids bearing hydroxyl groups at C-12 (Dean & Aplin, 1966; Tökés, 1970) and clearly shows that arapaimol-A and -B contain four hydroxyl groups on the nucleus and one and two such groups respectively on the side chain. The peaks mentioned above were seen appropriately at 431, 371, 311 and 251 mass units in the spectra of the acetylated derivatives.

Arapaimol-A was recovered unchanged after treatment with dry acetone and hydrochloric acid; with lead tetra-acetate it consumed about 3 atoms of oxygen/mol. These results showed that the fourth hydroxyl group in the arapaimol ring system is vicinal to one of the others and in *trans* relationship to it.

Mass-spectral analysis of the rather non-volatile arapaimic acid was carried out on a sample which was methylated and then acetylated. The resulting tetra-acetyl methyl ester exhibited a barely detectable molecular-ion peak at *m/e* 648, corresponding to $C_{36}H_{56}O_{10}$. It also exhibited the *m/e* 431 fragment ion, which results from the loss of the side chain and a molecule of acetic acid, and its subsequent deacetylation products of *m/e* 371, 311 and 251. These results indicate that arapaimic acid is a tetrahydroxy-cholestanoic acid ($C_{27}H_{46}O_6$) with four hydroxyl groups on the steroid nucleus and a carboxylic acid function on the side chain.

The location of the side-chain hydroxyl groups in the arapaimols was established by n.m.r. analysis (Table 1). Integration of the peak area in the (δ) 3.5–4.5 p.p.m. region in the spectra of arapaimol-A and -B revealed the presence of six and eight protons respectively, which are on carbon atoms bearing a hydroxyl group. In the spectrum of arapaimol-A there are two secondary methyl resonances, one of which has the typical perturbed shape attributed to the C-21 protons, and a two-proton multiplet centred at 3.63 p.p.m. (C-26H), which became a doublet when deuterium oxide was added to the penta-deuteropyridine solution, indicating that the side-chain hydroxyl group is at C-26.

The n.m.r. spectrum of arapaimol-B showed only the C-21 secondary methyl resonance and a doublet centred at 4.07 p.p.m. for the four protons of two magnetically equivalent primary alcohols. Thus the two side-chain hydroxyl groups in arapaimol-B are at C-26 and C-27. Similarly, the presence of only two secondary methyl resonances in the n.m.r. spectrum of arapaimic acid, one of which is the typically perturbed C-21H doublet, indicates that this compound is a substituted cholestan-26-oic acid. These assignments were further substantiated by the n.m.r. results with the acetylated derivatives (see Table 1).

Comparison of the nuclear proton signals in the region below 3 p.p.m. and the chemical shifts of the angular methyl proton resonances in various solutions of the arapaimols, arapaimic acid and their acetylated derivatives (see Table 1) clearly confirms the deduction from i.r. spectra that the arapaimols and arapaimic acid have the same nuclear structure. Of the four secondary nuclear hydroxyl groups, two are equatorial (broad multiplets at 3.5–3.9 and 3.9–4.4 p.p.m. for the axial protons) and the other two are axial (narrower multiplets in the 4.0–4.3 p.p.m. region for the equatorial protons). The resonances at 3.5–3.9, 4.04–4.08 and 4.19–4.22 p.p.m. are suggestive of the presence of a 3(equatorial),-7α,12α-trihydroxy pattern. This observation is also supported by the acetylation results, which indicated that one of the axial hydroxyl groups exhibited the characteristically slower reactivity of the 12α-hydroxyl function.

The identity of the less-reactive hydroxyl group as being that at C-12α has been confirmed as follows: (*a*) acetylation of this group exhibited the expected deshielding effect on the C-18H resonance (Bhacca & Williams, 1964) and a strong shielding effect on the C-21H resonance (see, e.g., Tökés, 1970) in deuterochloroform solution; (*b*) chromic acid oxidation of arapaimol-A tetra-acetate gave a ketone which exhibited in its mass spectrum an intense peak at *m/e* 407 (accompanied by subsequent deacetylation products, *m/e* 347, 287 and 227), which corresponds to the characteristic ring-D cleavage of 12-oxo steroids (Djerassi & Tökés, 1966) with a hydrogen transfer to the charge-retaining fragment.

It thus became apparent that the arapaimols are most probably derivatives of either 3α,7α,12α-trihydroxy-5β-cholestane or of 3β,7α,12α-trihydroxy-5α-cholestane with an extra hydroxyl group on the nucleus. The former (5β) alternative was favoured because in pentadeuteropyridine the n.m.r. spectra of arapaimol-A, arapaimol-B and arapaimic acid exhibited a one-proton quartet resonance, centred at 3.15–3.25 p.p.m. (depending on the moisture or deuterium oxide content of the solvent), which is present (L. Tökés, unpublished work) in the spectra of 3α,7α-dihydroxy-5β-steroid derivatives (methyl cholate, 3.05 p.p.m.; 2β-hydroxycholic acid, 3.17 p.p.m.) but absent (L. Tökés, unpublished work) from the spectra of the 3α,7α-dihydroxy-5α-steroids 5α-cyprinol (5α-cholestane-3α,7α,12α,26,27-pentol), 5α-chimaerol (5α,25*S*-cholestane-3α,7α,12α,24ϵ,26-pentol), petromyzonol (5α-cholane-3α,7α,12α,24-tetrol), 2β-hydroxyallochoic acid (see below) and deoxymyxinol (5α,25ϵ-cholestane-3α,7α,26-triol). This resonance is apparently due to the C-4α proton, which is strongly deshielded by the solvated C-3α and C-7α hydroxyl groups in the 5β-configuration and coupled to the C-3β and C-5β protons (axial–axial coupling) and to the C-4β proton (geminal coupling) to about equal extent (*J* 13 Hz). Interpretation of the paramagnetic shift of the C-12α acetyl signal of the fully acetylated arapaimols in hexadeuterobenzene solution (see below) provided further evidence in favour of the 5β skeleton.

The ring-D cleavage fragment in the mass spectrum of the 12-oxo derivative made by chromic acid oxidation of arapaimol-A tetra-acetate was also useful in finding the location of the fourth nuclear hydroxyl function. Only the side-chain acetoxyl group was lost during this cleavage, which shows unequivocally that the remaining three acetate groups are in rings A, B or C. Ring C is excluded by the multiplicity of the C-12βH resonance, which should be a doublet in the presence of either an 11α-ol or 11β-ol, leaving the 1α, 2β, 4β and 6α positions as the only possible locations for an equatorial secondary hydroxyl group on a 5β skeleton. The 1α position is excluded by the lead tetra-acetate experiments and the observed quartets in the 3.15–3.25 p.p.m. region of the n.m.r. spectra of the arapaimols and arapaimic acid assigned to the 4α-H (see above) leave only the 2β and 6α positions in consideration. The resistance of arapaimol-A to acetonide formation indicated the 2β position and this idea was further supported by the following observation. In the n.m.r. spectra of the fully acetylated arapaimols in hexadeuterobenzene solution one of the acetyl resonances is at unusually low-field (2.04 and 2.05 p.p.m.). This signal is absent in the spectra of triacetyl methyl cholate and of the partially acetylated arapaimol derivatives (Table 1). This low-field signal is apparently due to a C-12α-acetoxy group that is sterically hindered by the extra acetate group preventing its complex-formation with the solvent. Such an interaction is more likely with a 2β substituent of a 5β-steroid than with a 6α substituent.

The reported 0.02–0.03 p.p.m. deshielding effect of a 2β-hydroxyl substituent on the C-19H resonance in pentadeuteropyridine solution in the crustecdysone series (Galbraith *et al.*, 1968) provides further support for the preferred 2β,3α,7α,12α-tetrahydroxy-5β-steroid nucleus. This value added to the chemical shift of the C-19 proton resonance of methyl cholate (see Table 1) gives 1.01 p.p.m., which is very close to the observed values for the arapaimols and arapaimic acid.

On the basis of all this evidence, the most likely structures for arapaimol-A, arapaimol-B and arapaimic acid are 5β-cholestane-2β,3α,7α,12α,-26-pentol (IX), 5β-cholestane-2β,3α,7α,12α,26,27-hexol (X) and 2β,3α,7α,12α-tetrahydroxy-5β-cholestan-26-oic acid (XI) respectively, (Schemes 1 and 2).

We therefore made 2β,3α,7α,12α-tetrahydroxy-5β-cholan-24-oic acid (VIII, Scheme 2) as a model substance and, since 5α rather than 5β bile alcohols have previously been found to be characteristic of primitive osteichthyeans, we also prepared 2β,3α,7α,12α-tetrahydroxy-5α-cholan-24-oic acid (V, Scheme 1).

Methyl 7α,12α-diacetoxy-3-oxo-5α-cholan-24-oate (I) was easily brominated to give crystalline methyl 2α-bromo-7α,12α-diacetoxy-3-oxo-5α-cholan-24-oate (II), melting sharply and showing a main spot with a fainter contaminant on t.l.c. Its n.m.r. spectrum suggested that this material was contaminated with its 2β-bromo epimer. The mixture was reduced with sodium borohydride in methanol to a second mixture (III), which was boiled with zinc dust in acetic acid, giving a gum that presumably contained the substance (IV) (Fieser & Fieser, 1959, p. 251). Treatment with formic acid and hydrogen peroxide, followed by careful hydrolysis, gave material from which was isolated a crystalline acid with the properties expected of 2β,3α,7α,12α-tetrahydroxy-5α-cholan-24-oic acid (2β-hydroxyallocholic acid, V). This compound, m.p. 252°C (decomp.), $[\alpha]_D$ +26°, took up 1 atom of oxygen/mol with lead tetra-acetate and had the n.m.r. characteristics shown in Table 1; its i.r. spectrum was different from that of the arapaimols (see the Discussion section).

We made the 5β-epimer of (V) as shown in Scheme 2. Methyl 4β-bromo-7α,12α-diacetoxy-3-oxo-5β-cholan-24-oate (VI) on treatment with potassium acetate in acetic acid gave, as expected from the work of Satoh *et al.* (1966), a compound identified as methyl 2β,7α,12α-triacetoxy-3-oxo-5β-cholan-24-oate (VII). Sodium borohydride reduction of this or of the crude debrominated mixture, followed by hydrolysis, easily gave 2β,3α,7α,12α-tetrahydroxy-

Table 1. *Proton resonances of bile alcohols and acids from Arapaima gigas and of related model compounds*

The assignment of the C-7β and C-12β proton resonances of arapaimol-A, arapaimol-B, arapaimic acid, 2β-hydroxycholic acid, 2β-hydroxyallocholic acid and methyl cholate in the table is based on our observations that in pentadeuteropyridine solution the C-12βH in deoxycholic acid is represented by a triplet at 4.17p.p.m. and the C-7βH in deoxymyxinol by a narrow multiplet at 4.04p.p.m. We now wish to reverse the assignment of these protons in 5α-chimaerol (Tökés, 1970), i.e. the C-7βH is represented by the narrow multiplet at 4.08p.p.m. and the C-12βH resonance is overlapped by the C-3βH signal at 4.27p.p.m. (d), Doublet; (m), multiplet; (bm), broad multiplet; (nm), narrow multiplet; (t), triplet; $W_{\frac{1}{2}}$, width of peak at half-height; c, with the small amount of sample used, the spectrum was too weak for unequivocal assignments of these signals.

	Chemical shifts (δ, p.p.m.)									
Protons at ... / Compounds	C-18	C-19	C-21	C-26	C-27	C-2α	C-3β	C-7β	C-12β	—OCOCH$_3$
Arapaimol-A (IX) in pentadeuteropyridine	0.795	1.04	1.15 (d, J 5.5Hz)	3.63 (nm* $W_{\frac{1}{2}}$ 13Hz)	1.06 (d, J 6Hz)	3.9–4.3 (bm)†	3.5–3.9 (bm)	4.04 (m)	4.19 (m)	
penta-acetate in deuterochloroform	0.72	0.94	0.80 (d, J 6Hz)	3.86 (d, J 6.5Hz)	0.90 (d)	c	c	c	c	1.98, 2.02, 2.03, 2.07, 2.16
in hexadeuterobenzene	0.535 and 0.555		0.955 (d, J 6Hz)	3.88 (d, J 6Hz)	0.81 (d, J 6.5Hz)	c	c	c	c	1.64, 1.685, 1.71 (6H), 2.04
tetra-acetate in hexadeuterobenzene	0.51	0.63	0.91 (d, J 6Hz)	3.92 (d, J 6.5Hz)	0.82 (d, J 7Hz)	c	c	c	c	1.63–1.76
Arapaimol-B (X) in pentadeuteropyridine	0.78	1.045	1.18 (d, J 5.5Hz)	4.07 (d, J 5.5Hz)		3.95–4.35 (bm)†	3.5–3.9 (bm)	4.02–4.15‡	4.22 (m)	
hexa-acetate in deuterochloroform	0.72	0.94	0.81 (d, J 5.5Hz)	4.05 (d, J 6Hz)		c	c	c	c	1.97, 2.01, 2.03 (6H), 2.07, 2.16
in hexadeuterobenzene	0.54 and 0.555		0.95 (d, J 5.5Hz)	4.05 (d, J 6Hz)		c	c	c	c	1.65 (9H), 1.68, 1.71, 2.05
penta-acetate in deuterochloroform	0.66	0.95	~0.94 c	4.05 (d, J 6Hz)		c	c	c	c	1.98–2.05
in hexadeuterobenzene	0.51	0.63	0.91 (d, J 6Hz)	4.06 (d, J 6Hz)		c	c	c	3.73 (m, $W_{\frac{1}{2}}$ 6Hz)	1.63, 1.67 (6H) 1.72, 1.75
Arapaimic acid (XI) in pentadeuteropyridine	0.775	1.03	1.17 (d, J 5.5Hz)	1.29 (d, J 6.5Hz)		4.0–4.4 (bm)†	3.5–3.9 (bm)	4.08 (m)	4.20 (m)	
methyl ester tetra-acetate in deuterochloroform	0.72	0.94	0.80 (d, J 6Hz)	1.13 (d, J 6.5Hz)		c	c	c	c	1.98, 2.02, 2.08, 2.16
2β-Hydroxycholic acid (VIII) in pentadeuteropyridine	0.76	1.01	1.19 (d, J 5Hz)			4.0–4.4 (bm)†	3.5–3.9 (bm)	4.05 (m)	4.19 (m)	
methyl ester tetra-acetate in deuterochloroform	0.72	0.94	0.81 (d, J 5.5Hz)			4.5–5.2 (bm)		4.89 (m, $W_{\frac{1}{2}}$ 8Hz)	5.10 (t, $W_{\frac{1}{2}}$ 6Hz)	1.97, 2.01, 2.05, 2.14
in hexadeuterobenzene	0.46	0.53	0.82 (d, J 5Hz)			4.7–5.4 (bm)		4.87 (m, $W_{\frac{1}{2}}$ 8Hz)	5.22 (m)	1.64, 1.67, 1.70, 1.98
methyl ester triacetate in deuterochloroform	0.67	0.95	0.97			c	c	4.85 (m)	4.02 (m, $W_{\frac{1}{2}}$ 6Hz)	1.98, 2.00, 2.02
in hexadeuterobenzene	0.45	0.60	0.81 (d)			c	c	4.90 (m)	3.65 (m)	1.62, 1.69, 1.73

2β-Hydroxyallocholic acid (V)	0.79	1.40	1.20 (d,	4.27	4.09 (m)	4.16 (m)	
in pentadeuteropyridine			*J* 5.5Hz)	(nm, $W_{\frac{1}{2}}$ 6Hz)			
methyl ester tetra-acetate	0.71	0.90	0.79 (d,	4.89		5.03 (t,	2.01, 2.06, 2.09
in deuterochloroform			*J* 5.5Hz)	(m, $W_{\frac{1}{2}}$ 9Hz, 3*H*)		$W_{\frac{1}{2}}$ 6Hz)	(6H)
in hexadeuterobenzene	0.485	0.695	0.79 (d,	4.95–5.15			1.51, 1.675
			J 5Hz)	(m, $W_{\frac{1}{2}}$ 14Hz, 4*H*)			(6H), 1.71
methyl ester triacetate	0.675	0.915	0.97 (d,	4.91		3.96 (t,	2.02, 2.04, 2.06
in deuterochloroform			*J* 5.5Hz)	(nm, $W_{\frac{1}{2}}$ 7Hz, 3*H*)		$W_{\frac{1}{2}}$ 6Hz)	
in hexadeuterobenzene	0.48	0.73	0.84 (d,	4.95–5.13		3.64 (m,	1.62, 1.65, 1.68
			J 5.5Hz)	(m, $W_{\frac{1}{2}}$ 11Hz, 3*H*)		$W_{\frac{1}{2}}$ 7Hz)	
Methyl cholate	0.78	0.98	1.14 (d,	3.45–3.85	4.05 (m,	4.18 (m,	
in pentadeuteropyridine			*J* 5.5Hz)	(bm)	$W_{\frac{1}{2}}$ 10Hz)	$W_{\frac{1}{2}}$ 9Hz)	

* This signal became a broad doublet (*J* 6.5Hz) when a drop of deuterium oxide was added.
† This signal was overlapped by the C-7β,C-12β,C-26 (and C-27H in X) resonances and hence it was identified mainly from the integral values.
‡ The multiplicity of this signal could not be established because of overlap by the C-26H and C-27H resonances.

5β-cholan-24-oic acid (2β-hydroxycholic acid, VIII), m.p. 228°C, $[\alpha]_D$ +27°, which took up about 3 atoms of oxygen/mol with lead tetra-acetate and the i.r. spectrum (Fig. 1*b*) of which left no doubt that it contained the arapaimol steroid nucleus.

The n.m.r. findings for 2β-hydroxycholic acid (VIII) and its methyl ester tri- and tetra-acetate (Table 1) in a variety of solvents are in good agreement with the results derived from the arapaimols and arapaimic acid and they are distinctly different from those of the 5α-epimers. Hence arapaimol-A and -B have the structures (IX) and (X) respectively. As argued above, mass and i.r. spectra (Table 1) clearly point to formula (XI) for arapaimic acid.

Experimental

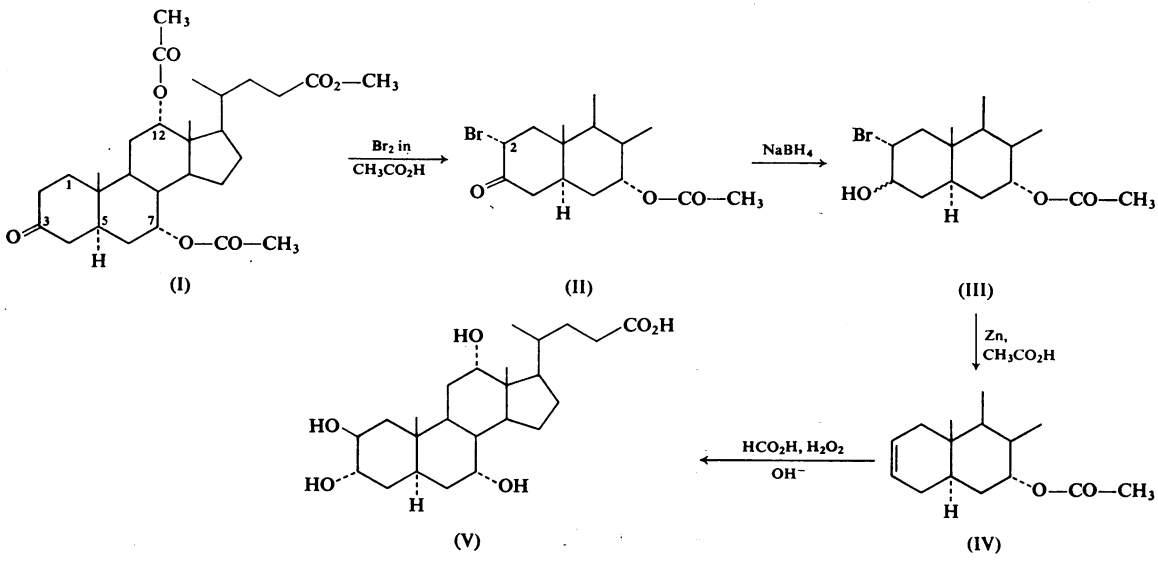
General

Isolation of bile salts, alkaline hydrolysis, esterification, separation on Celite and determination of i.r. spectra were as described by Haslewood (1967). Microanalysis (C and H) was done by F. B. Strauss, Oxford, U.K. Melting points are corrected. T.l.c. was as previously described (Haslewood, 1971) and g.l.c. (by I. G. Anderson) as described by Anderson & Haslewood (1970). N.m.r. spectra were taken at 100MHz with a Varian HA-100 spectrometer, with tetramethylsilane as internal reference. Mass spectra were determined with an Atlas CH-4 mass spectrometer, equipped with an EFO-4B ion source. The ionizing voltage was maintained at 70eV.

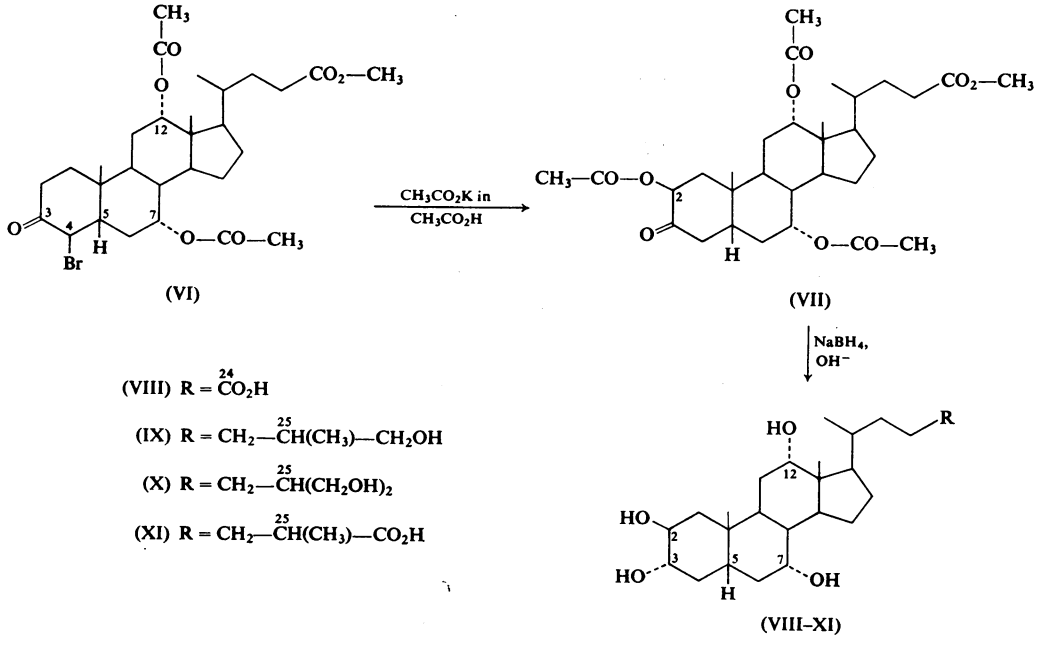
Bile alcohols and acids from A. gigas

Bile salts (4.6g from one fish weighing about 60kg) on t.l.c. in the solvent system 3-butyl acetate–acetic acid–water (5:5:2, by vol.) gave four principal spots, one of which had the R_F value (about 0.33) of taurocholate and the others R_F values about 0.47, 0.41 and 0.28. Similar results were given by another sample of *A. gigas* bile salts, but two other osteoglossids showed only the three least-polar spots.

Alkaline hydrolysis. Bile salts (463mg) in 0.5M-NaOH (5ml) were hydrolysed at 193°C for 17h. The resulting mixture was clear and filtrable; it was acidified with 2M-HCl (excess) and, after refrigeration, the precipitate was collected. The acidic liquors with 0.5M-$BaCl_2$ (5ml) gave $BaSO_4$ (77mg). The organic material dissolved readily in aq. NH_3 soln., from which it was not precipitated by NaCl (excess) alone but on the further addition of 2M-HCl (excess). After refrigeration, the precipitate was collected, washed with water and redissolved in ethanol–acetone. Evaporation of the solvents left a residue, which was extracted with ethyl acetate; the extract was filtered and evaporated to give ethyl acetate-soluble material



Scheme 1.



(VIII) R = $^{24}CO_2H$

(IX) R = $CH_2—\overset{25}{C}H(CH_3)—CH_2OH$

(X) R = $CH_2—\overset{25}{C}H(CH_2OH)_2$

(XI) R = $CH_2—\overset{25}{C}H(CH_3)—CO_2H$

Scheme 2.

(212mg). With aq. NH_3 soln. this gave a turbid solution, which was extracted three times with ethyl acetate and enough ether to break emulsions. Evaporation of this extract gave 'neutrals' (129mg), and 'acids' were recovered by similar extraction after acidification of the aqueous ammoniacal portion with

HCl and NaCl (excess). The 'acids' (40mg) with a little acetone gave white crystals (9mg), m.p. 234–246°C (decomp.), which on recrystallization from acetone gave *arapaimic acid* (XI) as colourless prisms, m.p. 244–248°C (decomp.), giving a purple colour in the Hammarsten (HCl) test. On t.l.c. in the system isopropyl acetate–acetic acid–water (15:3:1, by vol.) this acid gave a single spot with R_F about 0.34, when cholic acid had R_F about 0.41. Its i.r. spectrum was closely similar to that shown in Fig. 1(*a*) in the region 9.0–15.0μm but differed between 6.5 and 9.0μm and showed a strong carbonyl band (attributed to CO_2H) at about (max.) 5.86μm. For mass-spectral analysis this acid was methylated with diazomethane and acetylated with isopropenyl acetate. The resulting tetra-acetoxy methyl ester exhibited the following diagnostically important peaks: *m/e* 648 (M^+), 528 ($M^+-2CH_3CO_2H$), 468 ($M^+-3CH_3CO_2H$), 408 ($M^+-4CH_3CO_2H$), 431 [(M^+-side chain = *a*)$-CH_3CO_2H$], 371 (*a*$-2CH_3CO_2H$), 311 (*a*$-3CH_3CO_2H$) and 251 (*a*$-4CH_3CO_2H$). The n.m.r. data are shown in Table 1.

The acetone mother liquors from arapaimic acid were evaporated. The residues on t.l.c., as described above, showed spots corresponding to both cholic acid and arapaimic acid and, after methylation and conversion into trimethylsilyl derivatives, the same acids were evident on g.l.c. The retention time of methyl trimethylsilyl arapaimate was 4.07, relative to methyl 5β-cholan-24-oate (1.0). The above 'neutrals' were not crystallized, but ethyl acetate-insoluble material from them showed the same i.r. spectrum as the arapaimols (Fig. 1*a*).

A similar alkaline hydrolysis of bile salts (300mg), but with the addition of '100%' hydrazine hydrate (about 10%, v/v), gave a less-coloured product which similarly yielded $BaSO_4$ (41mg), 'acids' (68mg) and 'neutrals' (56mg). The 'acids', with acetone, gave crude arapaimic acid (13mg), m.p. 215–245°C (decomp.).

Dioxan–trichloroacetic acid cleavage. A. gigas bile salts (506mg) in acetic acid (5ml) with acetic anhydride (5ml) were heated on a boiling-water bath under reflux for 1h. Removal of solvent at atmospheric pressure, then *in vacuo*, gave a residue which was dissolved in 10ml of a 40% (w/w) solution of dry trichloroacetic acid in dry dioxan and kept in the dark with occasional mixing for 11 days. The mixture was diluted with water and NaCl (excess) and ether-extracted thrice. The aqueous portion with 2M-HCl (excess) and 0.5M-$BaCl_2$ (5ml) gave no precipitate at first, but after heating on a boiling-water bath for 3h yielded (presumably) $BaSO_4$ (75mg). The combined ethereal extract was washed with aq. NH_3 soln. and water, dried (over Na_2SO_4) and evaporated, leaving a residue which was hydrolysed by boiling under reflux for 1.5h in methanol (5ml) with 1M-NaOH (1ml). Solvent was removed in a stream of N_2 and the residue diluted with water, refrigerated, collected and washed with water, giving a neutral product (100mg). This material (89mg), whose i.r. spectrum suggested the presence of acetyl residues, was again hydrolysed, as described above, to give a product (74mg) whose i.r. spectrum was that of the arapaimols (Fig. 1*a*). The above re-hydrolysed product (70mg) was separated on Celite (10g) in the system benzene–light petroleum (b.p. 80–100°C)–ethanol–water (6:1:5:2, by vol.). Moving phase (84 fractions of 5ml each) eluted material in three principal peaks as follows (peak no., ml of moving phase, mg eluted): I, 5–45, 7; II, 50–95, 33; III, 150–245, 19. Material between the peaks (5mg) and substance (4mg) afterwards stripped from the column with methanol (100ml) account, with the above peaks, for a total recovery of 68mg. Material from peak I was an oil; that from the apex of peak II (5.4mg), and eluted by 60–65ml of moving phase, formed white prisms of (probably) hydrated *arapaimol-A* (IX), m.p. 216–220°C, $[\alpha]_D^{25}$ +33.9±2° (*c* 1.2 in ethanol); i.r. spectrum, Fig. 1(*a*) [Found, after drying *in vacuo* at 100°C to constant wt. (loss of wt., 9.1%): C, 69.8; H, 10.7. $C_{27}H_{48}O_5,H_2O$ requires C, 69.0; H, 10.6%]. Arapaimol-A had R_F about 0.46 when 5α-cyprinol had R_F 0.37 on t.l.c. in the system di-isopropyl ether–acetic acid–water (12:6:1, by vol.). On g.l.c. it had, as its trimethylsilyl derivative, a retention time of 1.93, relative to methyl trimethylsilyl cholate. Arapaimol-A (2mg) in dry acetone (0.3ml) was recovered unchanged after treatment for 22h with a saturated solution of HCl in dry acetone (0.1ml). With 0.02M-lead tetra-acetate in acetic acid (1ml) at 50°C for 17h, arapaimol-A (0.94mg) took up 2.8 atoms of oxygen/mol. Combined material from the remainder of peak II, with ether, gave crystals (28.6mg) of slightly less pure arapaimol-A, m.p. 214–220°C, which on t.l.c. as above showed a faint second spot just behind the main spot and on g.l.c. a minor component (unidentified) that had, as its trimethylsilyl derivative, a retention time of 1.43 relative to methyl trimethylsilyl cholate. Peak III material, from aq. ethanol, gave fine white needles of *arapaimol-B* (X), m.p. 228–230°C, $[\alpha]_D^{25}$ +29.6±2° (*c* 1.1 in ethanol) [Found on material dried to constant wt. *in vacuo* (loss of wt., 12.5%): C, 69.9; H, 10.6. $C_{27}H_{48}O_6$ requires C, 69.2; H, 10.3%]. Arapaimol-B had R_F 0.21 on t.l.c. as described above, and on g.l.c. its trimethylsilyl derivative had a retention time of 3.11 relative to methyl trimethylsilyl cholate. Both arapaimols were acetylated by dissolving in acetic anhydride (1ml) and dry pyridine (1ml) and storing the solution at room temperature for 2 days in the dark. Dilution with water, ether extraction, washing of the ether phase with dilute $NaHCO_3$ soln. and plenty of water, drying and evaporation of the solvent gave a glassy residue, which was separated by t.l.c. on

Rhodamine-G-impregnated silica gel G plates (0.25mm thickness) in the system ether–benzene (1:1, v/v). The major zones, detected by u.v. light, were eluted with ether and the resulting gums shown by n.m.r. and mass spectra (see the Results section) to be arapaimol-A tetra- and penta-acetate and arapaimol-B penta- and hexa-acetate.

Arapaimol-A tetra-acetate (1.8mg) was oxidized with 2.7M-chromic acid (excess) in acetone solution by storing at room temperature for 1h. Methanol (a few drops) was added and then the reaction mixture was diluted with ether, which was washed with water and dried (over Na_2SO_4). Evaporation gave a glassy monoketone (1.1mg), shown (above) to be 2β,3α,7α,26-tetra-acetoxy-5β-cholestan-12-one, which exhibited the following diagnostic peaks in its mass spectrum: *m/e* 618 (M^+), 558, 498, 438, 407, 347, 287 and 227.

Preparation of 2β-hydroxyallocholic acid (V, Scheme 1)

Methyl 2α-bromo-7α,12α-diacetoxy-3-oxo-5α-cholan-24-oate (II). Methyl 3α,7α-dihydroxy-3-oxo-5α-cholan-24-oate (I, 202mg), made from methyl cholate by the method of Mitra & Elliott (1968), was acetylated by the perchloric acid method (Anderson & Haslewood, 1970) and the crude product (240mg) in acetic acid (2ml) was treated with 0.48ml of a 5% (v/v) solution of bromine in acetic acid. Bromine colour disappeared almost immediately and HBr was evolved. The product was at once diluted with water and the solid precipitate collected, washed with water and crystallized from aq. methanol to give (mainly) methyl 2α-bromo-7α,12α-diacetoxy-3-oxo-5α-cholan-24-oate (II), which, according to its n.m.r. spectrum was contaminated with its 2β-bromo epimer, as colourless prisms (205mg), m.p. 219–220°C (decomp.), n.m.r. (δ, p.p.m. in C^2HCl_3): 0.74 (C-18H), 1.08 (C-19H), 2.08, 2.09 ($2\times OCOCH_3$) and 3.67 (CO_2CH_3); mass spectrum: *m/e* 582 and 584 (M^+ with ^{79}Br and ^{81}Br) (Found after alkaline hydrolysis: Br, 13.7. $C_{29}H_{43}O_7Br$ requires 13.7%). This material on t.l.c. in the system benzene–acetone (9:1, v/v) gave a chief spot with R_F 0.39 and a second faint spot, R_F 0.57.

2β-Hydroxyallocholic acid (V). The above bromo compound (100mg) was suspended in methanol (2ml) and treated gradually, with cooling, with powdered $NaBH_4$ (25mg). The mixture became clear. After 30min 2M-HCl (excess) was carefully added and the precipitate, after refrigeration, was collected, washed with water and dried by evaporation *in vacuo* with ethanol. The yield was 8mg of material, giving on t.l.c., as for the starting bromo compound, two dark spots with R_F values of 0.22 and 0.30 and no spot corresponding to compound (II). A solution of this product (98mg) in acetic acid (2ml) was boiled under reflux for 1h after the addition of zinc dust (analytical quality, 0.4g). The mixture was diluted with water and extracted twice with ether. Br^- (97% of the theoretical amount) was estimated in the aqueous phase. The combined ether extracts were washed with aq. NH_3 soln. and water, dried (over Na_2SO_4) and evaporated, leaving a colourless gum (84mg), which was dissolved in formic acid (2ml of '98/100%', analytical quality). Hydrogen peroxide (0.4ml of '100vol.') was added and, after mixing, the material was left for 17h in the dark. It was then diluted with water and NaCl (excess) and after refrigeration the solid was collected, washed with water and dried by azeotropic distillation *in vacuo* with ethanol–benzene. The product (87mg) was treated with a portion (4ml) of a mixture prepared by boiling for 20min under reflux, in a stream of N_2, methanol (5ml), 1M-NaOH (5ml) and hydrazine hydrate (0.5ml of '99/100%'). Hydrolysis was continued under reflux in an atmosphere of N_2 for 1h, after which solvent was removed in N_2 and the residue treated with 2M-HCl (about 5ml) and NaCl (excess). After cooling, the precipitated acid was collected, washed with water and dried by evaporation with benzene–ethanol, leaving a residue that was dissolved in acetone. Evaporation of the filtered solution left a product (57mg), which with a little acetone gave crude crystals (21mg), m.p. 239–244°C (decomp.). Recrystallization from acetone gave *2β,3α,7α,12α-tetrahydroxy-5α-cholan-24-oic acid* (2β-hydroxyallocholic acid, V) as white crystalline globules, m.p. 250–252°C (decomp.), $[\alpha]_D^{25}$ +26.5± 1.6° (*c* 0.15 in ethanol). This material was apparently solvated (Found: C, 66.5; H, 9.2. $C_{24}H_{40}O_6,\frac{1}{2}H_2O$ requires C, 66.5; H, 9.5%).

This acid gave an intense purple response in the Hammarsten (HCl) test; on t.l.c. in the system isopropyl acetate–acetic acid–water (15:2:1, by vol.) it gave a spot just behind that given by its 5β-isomer (VIII, see below). A sample (1.3mg) was dissolved in dry acetone (0.6ml) by heating in a stoppered flask. A saturated solution of HCl in dry acetone (0.3ml) was added and the stoppered mixture left for 20h. Solvent was removed and the residue found, by t.l.c. and i.r. spectroscopy, to consist of 2β-hydroxyallocholic acid. With lead tetra-acetate (1ml of 0.02M in acetic acid) at 50°C for 16.5h, 2β-hydroxyallocholic acid (0.4mg) took up almost exactly 1 atom of oxygen/mol. For n.m.r. spectral examination, the acid was methylated and acetylated in pyridine–acetic anhydride (1:1, v/v) for 2 days. Separation by t.l.c. in the system ether–benzene (2:3, v/v) gave non-crystalline methyl ester tri- and tetra-acetate, whose n.m.r. characteristics, together with those of compound (V), are shown in Table 1. The C-19H resonance of compound (V) is much further down-field than those of the arapaimols and the paramagnetic shift confirms the presence of the

2β-OH group, which has a strong deshielding effect on the C-19 protons in 5α-steroids. The mass spectrum of compound (V) exhibited the following diagnostic peaks: *m/e* 406 (M^+-H_2O), 388 (M^+-2H_2O), 370 (M^+-3H_2O), 305 [(M^+−side chain = *a*)−H_2O], 287 (*a*−$2H_2O$), 269 (*a*−$3H_2O$) and 251 (*a*−$4H_2O$).

Preparation of 2β-hydroxycholic acid (VIII, Scheme 2)

Methyl 7α,12α-diacetoxy-3-oxo-5β-cholan-24-oate (10g) in acetic acid (100ml) was brominated by adding bromine (1ml). The product, precipitated at once by pouring into water (excess), was crystallized from methanol to give *methyl 4β-bromo-7α,12α-diacetoxy-3-oxo-5β-cholan-24-oate (VI)*, m.p. 208–210°C (Found, by alkaline hydrolysis or debromination as described below: Br, 13.7. $C_{29}H_{43}O_7Br$ requires Br, 13.7%). A mixture of the above compound (VI, 0.2g) with freshly fused anhydrous potassium acetate (0.4g) and acetic acid (2ml) was boiled under reflux for 0.5h. The cooled mixture was diluted with water and Na_2SO_4 (excess) and, after refrigeration, the solid precipitate was collected and washed with water. Titration of the aqueous liquors showed that debromination was complete. The crude organic solid was dried by evaporation *in vacuo* with ethanol–benzene and dissolved in methanol (2ml). Sodium borohydride (0.1g) was gradually added, with cooling, and the mixture left for 1h. It was then diluted with 2M-HCl and NaCl (excess). After refrigeration, the precipitate was collected, washed with water and hydrolysed by boiling under reflux for 1h with methanol (2.5ml) and 1M-NaOH (2.5ml). Solvent was removed under N_2 and the residue dissolved in water (4ml). 2M-HCl and NaCl (excess) were added and the precipitated acid, after refrigeration, was collected, washed with water and dissolved in ethanol. Evaporation left a residue, which was dissolved in acetone. The filtered solution was evaporated; the residue with a little fresh cold acetone slowly gave white crystals (25mg), which on recrystallization from acetone gave *2β,3α,7α,12α-tetrahydroxy-5β-cholan-24-oic acid* (2β-hydroxycholic acid, VIII) as short colourless needles, m.p. 227–228°C (decomp.); $[\alpha]_D^{18}$ +27.2±2° (*c* 1.03 in ethanol); i.r. spectrum, Fig. 1(*b*); mass spectrum: *m/e* 424 (M^+), 406 (M^+-H_2O), 388 (M^+-2H_2O), 370 (M^+-3H_2O), 352 (M^+-4H_2O), 305 [(M^+−side chain = *a*)−H_2O], 287 (*a*−$2H_2O$), 269 (*a*−$3H_2O$) and 251 (*a*−$4H_2O$) (Found on a sample dried to constant wt. at 100°C: C, 67.2; H, 9.5. $C_{24}H_{40}O_6$ requires C, 67.9; H, 9.4%). This acid gave an intense purple response in the Hammarsten test; on t.l.c. in the system isopropyl acetate–acetic acid–water (15:3:1, by vol.) it had R_F 0.31 when cholic acid had R_F 0.53. On g.l.c. it had, as the methyl trimethylsilyl derivative, a retention time of 2.48 relative to methyl 5β-cholan-24-oate. In the conditions described above for compound (V) it gave no acetonide and with lead tetra-acetate took up about 2.8 atoms of oxygen/mol.

A sample was converted into (non-crystalline) methyl ester tri- and tetra-acetate as described above for compound (V) and the n.m.r. results for these and for compound (VIII) itself are shown in Table 1.

In another experiment, the crude debrominated product from compound (VI), prepared as described above and twice crystallized from aq. methanol by Dr. A. R. Tammar, gave tiny white needles, which from benzene–light petroleum (b.p. 40–60°C) yielded crystals of *methyl 2β,7α,12α-triacetoxy-3-oxo-5β-cholan-24-oate* (VII), m.p. 168–170°C; n.m.r. (δ, p.p.m. in C^2HCl_3) 0.76 (C-18H), 1.03 (C-19H), 2.08, 2.10, 2.12 (3×$OCOCH_3$) and 3.65 (CO_2CH_3); mass spectrum: *m/e* 562 (M^+) (Found after drying to constant wt. at 100°C *in vacuo*: C, 66.9; H, 8.6. $C_{31}H_{46}O_9$ requires C, 66.2; H, 8.2%). After reduction with $NaBH_4$ and hydrolysis, as above, this substance gave 2β-hydroxycholic acid. The important C-2αH resonance in the n.m.r. spectrum of compound (VII) was overlapped in deuterochloroform solution by the C-7βH and C-12βH resonances but it was isolated, with its centre at δ = 5.52 p.p.m., when the spectrum was taken in hexadeuterobenzene. The observed C-1β,C-2α diaxial proton coupling (*J* 14.5Hz) and the C-1α,C-2α equatorial–axial coupling (*J* 6Hz) establish unequivocally that the extra acetoxy group, introduced by this reaction sequence, is at the C-2β position.

Discussion

Chemical

The structure of the arapaimols and arapaimic acid is satisfactorily established. Stereochemistry of arapaimol-A at C-25 may be deduced as follows (see Anderson & Haslewood, 1970). Cholic acid has M_D +150° (M_D is defined as $[M]_D/100$); 2β-hydroxycholic acid, M_D +115°. Hence ΔM_D for 2β-OH is −35°. The C-25*R* and C-25*S* 5β-cholestane-3α,7α,12α,26-tetrols have M_D +157° and +108° respectively; the calculated values for the 2β-hydroxy derivatives (arapaimol-A) are +122° and +73°. Since arapaimol-A has M_D +153° it is the C-25*R* compound and agrees with 5α- and 5β-chimaerol in its stereochemistry at this centre.

Scheme 1. It was a surprise to find that a substantial amount of the 2β-epimer resulted from the bromination of compound (I); 5α-cholestan-3-one is thought to yield only the 2α-bromo compound (Fieser & Fieser, 1959, p. 281). Possibly insufficient time was allowed after bromination for the HBr

formed to complete conversion into the thermodynamically more stable 2α product. Although 2β-hydroxyallocholic acid (V) was obtained only in low yield, there seems little doubt about its constitution. The i.r. spectrum showed prominent bands at about 9.20, 9.65, 11.21 and 11.47μm. As compared with the allocholic acid spectrum (Haslewood, 1967) the double peak in the region 9.5–10.0μm is replaced by a single band at about 9.65μm with a shoulder at 9.90μm and an additional band appears at about 11.47μm. The reactivity with lead tetraacetate is in general agreement with that found for 5α-spirostan-2β,3α-diol (Djerassi & Ehrlich, 1954) and the failure to form an acetonide confirms that the hydroxyl groups at C-2 and C-3 are in *trans* relationship to each other. 2β-Hydroxyallocholic acid may be useful as a reference compound if 2β-hydroxylation is found in 5α bile salts.

Scheme 2. The i.r. (Fig. 1) and n.m.r. (Table 1) spectra leave no doubt about the identity of the steroid nucleus of the arapaimols and arapaimic acid with that of 2β-hydroxycholic acid and there can be little doubt also about the correctness of structure (VIII) for the latter. As compared with the i.r. spectrum of cholic acid, that of compound (VIII) shows greatly enhanced absorption in the region 9–10μm, with maxima at about 9.3 and 9.6μm as well as the features mentioned in the Results section. The failure to form an acetonide is consistent with structure (VIII); the great reactivity with lead tetraacetate is specially noteworthy.

The arapaimol (2β-hydroxycholic acid) nucleus is unstable to oxidation and also to acid, and this probably explains the poor yields both of arapaimol and of arapaimic acid obtained after alkaline hydrolysis and dioxan–trichloroacetic acid cleavage; little improvement resulted from the prior addition of hydrazine hydrate in the hydrolysis. Thus investigation of bile salts of this type is likely to be difficult unless it becomes possible, for example, to isolate conjugates containing the arapaimol nucleus, which could then be examined spectroscopically, or unless mild (e.g. enzymic) methods of hydrolysis can be found.

Biological

The bile salts of *A. gigas* are both primitive and specialized; primitive because they consist largely of bile alcohol sulphates and a C_{27} acid, and specialized in showing 2β-hydroxylation. How far these features are shared by other osteoglossids remains to be discovered, but t.l.c. of bile salts from a few other representatives suggested similarities with *A. gigas*. Thus it seems likely that the chemical nature of osteoglossid bile salts is entirely in accord with the systematic position of the fishes as delineated by Greenwood *et al.* (1966) on general taxonomic grounds. The presence of taurocholate is to be expected in teleosts generally, on the basis of all previous findings. If the special chemical feature, 2β-hydroxylation, is primary (i.e. caused by liver enzymes) it ought to be of taxonomic value and it should prove to be readily detectable.

The discovery of yet another chemical pattern in bile salts emphasizes again the remarkable nature of this character; wide variations are clearly compatible with physiological function and offer attractive new possibilities for the study of molecular evolution.

We thank Dr. P. H. Greenwood, British Museum (Natural History) for the bile of *Arapaima* and other osteoglossids, and Professor W. H. Elliott, Department of Biochemistry, School of Medicine, St. Louis University, for a generous gift of methyl 7α,12α-diacetoxy-3-oxo-5α-cholan-24-oate and advice on its preparation. We also thank Dr. A. R. Tammar for methyl 2β,7α,12α-triacetoxy-3-oxo-5β-cholan-24-oate, Dr. I. G. Anderson for g.l.c. and, at Syntex Research, Dr. M. Maddox and Mrs. J. Nelson for recording the n.m.r. spectra, Mr. J. W. Smith for the mass-spectral data and Mr. B. Amos for some of the t.l.c. separations. This is contribution no. 392 from the Institute of Organic Chemistry, Syntex Research.

References

Anderson, I. G. & Haslewood, G. A. D. (1970) *Biochem J.* **116**, 581

Bhacca, N. S. & Williams, D. H. (1964) *Applications of NMR Spectroscopy in Organic Chemistry*, Chapter 2, Holden-Day Inc., San Francisco

Dean, P. D. G. & Aplin, R. T. (1966) *Steroids* **8**, 565

Djerassi, C. & Ehrlich, R. (1954) *J. Org. Chem.* **19**, 1351

Djerassi, C. & Tökés, L. (1966) *J. Amer. Chem. Soc.* **88**, 536

Fieser, L. F. & Fieser, M. (1959) *Steroids*, Reinhold Publishing Corp., New York

Galbraith, M. N., Horn, D. H. S. & Middleton, E. J. (1968) *J. Chem. Soc. D* 83

Greenwood, P. H., Rosen, D. E., Weitzmann, S. H. & Myers, G. S. (1966) *Bull. Amer. Mus. Natur. Hist.* **131**, 341

Haslewood, G. A. D. (1967) *Bile Salts*, Chapter 3, Methuen and Co. Ltd., London

Haslewood, G. A. D. (1971) *Biochem. J.* **123**, 15

Haslewood, G. A. D. & Tökés, L. (1970) *Biochem. J.* **121**, 8P

Mitra, M. N. & Elliott, W. (1968) *J. Org. Chem.* **33**, 175

Satoh, Y., Mukoh, M., Ogaki, Y., Takahashi, T., Kimura, T., Aoki, H. & Hagitani, A. (1966) *Bull. Chem. Soc. Jap.* **89**, 855

Tökés, L. (1970) *Biochem. J.* **116**, 585