Isolation and Crystal Structure of an Arsenic-containing Sugar Sulphate from the Kidney of the Giant Clam, *Tridacna maxima*. *X*-Ray Crystal Structure of (2*S*)-3-[5-Deoxy-5-(dimethylarsinoyl)-β-D-ribofuranosyloxy]-2-hydroxypropyl Hydrogen Sulphate

John S. Edmonds and Kevin A. Francesconi
Western Australian Marine Research Laboratories, P.O. Box 20, North Beach, Western Australia, 6020
Peter C. Healy
School of Science, Griffith University, Nathan, Queensland, 4111
Allan H. White
Department of Physical and Inorganic Chemistry, University of Western Australia, Nedlands, Western Australia, 6009

(2*S*)-3-[5-Deoxy-5-(dimethylarsinoyl)-β-D-ribofuranosyloxy]-2-hydroxypropyl hydrogen sulphate, (1a) and 3-[5-deoxy-5-(dimethylarsinoyl)-β-D-ribofuranosyloxy]propylene glycol, (1b), have been isolated from the kidney of the giant clam, *Tridacna maxima*, collected from Shark Bay, Western Australia. The structure of compound (1a) was determined by *X*-ray diffraction. The source of the arsenicals (1a) and (1b) is likely to be symbiotic, unicellular, green algae living in the clam tissues. As arsenic-containing sugars [including compound (1b)] have previously been isolated from *Ecklonia radiata* (a macroalga), it is possible that the production of such compounds is a general response of algae to oceanic arsenate.

Clams of the genus *Tridacna* support, within their tissues, symbiotic unicellular algae (zooxanthellae) [1] and, as a consequence, products of algal metabolism and senescence might be expected in their large and accumulatory kidneys. Benson and Summons [2] have studied the arsenicals present in the kidney of the giant clam *Tridacna maxima* taken from the Great Barrier Reef, Australia, and they accounted for over 90% of the arsenic present in terms of trimethylarsoniolactate, $Me_3As^+CH_2CH(OH)CO_2^-$, and its derivatives. Trimethylarsoniolactate and compounds based upon it had previously been reported [3] as being produced by unicellular marine algae exposed to radio-labelled arsenate (see below). We now report the isolation of the novel (2*S*)-3[5-deoxy-5-(dimethylarsinoyl)-β-D-ribofuranosyloxy]-2-hydroxypropyl hydrogen sulphate, (1a), from the kidney of *Tridacna maxima*, collected from Shark Bay, Western Australia. The structure of this compound was determined by single-crystal *X*-ray diffraction.

Bulked kidneys (0.02% As) from 25 clams were extracted with methanol and water and compound (1a) and the related glycoside 3-[5-deoxy-5-(dimethylarsinoyl)-β-D-ribofuranosyloxy]propylene glycol † (1b) (previously isolated [4] from the brown kelp *Ecklonia radiata*) were isolated from the extract by chromatography on Sephadex LH-20 and ion-exchange Sephadex resins and by preparative layer chromatography (p.l.c.). Compounds (1a) and (1b) together accounted for over 80% of the arsenic present in the clam kidneys.

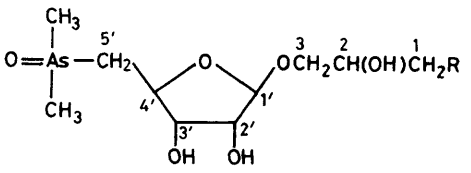
The arsenic-containing sugars (1b) and (1c) have previously been isolated [4] from brown kelp, *Ecklonia radiata*, growing in the near-shore waters of southwestern Australia. Trimethylarsonioacetate (arsenobetaine), $Me_3As^+CH_2CO_2^-$, has been found in the western rock lobster, *Panulirus cygnus*,[5] and the school whiting, *Sillago bassensis*,[6] associated with the ecosystem supporting the *Ecklonia* kelp. That the arsenic-containing sugars are precursors of arsenobetaine has been strongly suggested by the production of 2-dimethylarsinoylethanol, $O{=}As(Me)_2CH_2CH_2OH$, by anaerobic incubation of *Ecklonia*.[7] Although the ubiquity of arsenic-containing sugars in marine algae is yet to be established, arsenobetaine has been identified in marine animals where *Ecklonia* cannot have been involved.[5,8,9] In particular, arsenobetaine has been identified [5,9] in oceanic sharks where phytoplankton rather than macroalgae are likely to be the basis of the food web. These observations imply either that the biosynthesis of arsenic-containing sugars is a general response by algae (phytoplankton as well as macroalgae) to the arsenate to which they are subjected, or that arsenobetaine is a universal end-product of arsenic metabolised through different intermediates. The demonstration [3] that unicellular marine algae exposed to radio-labelled arsenate elaborated the previously unreported trimethylarsoniolactate, and compounds based upon it, tended to support the latter alternative as these compounds could also be envisaged as precursors of arsenobetaine. Support for the view that unicellular algae produce trimethylarsoniolactate and its derivatives was provided by Benson and Summons [2] in their study of the arsenicals in the kidney of *Tridacna maxima* from the Great Barrier Reef. Those authors did not, however, reveal how the arsenicals were identified. The work reported in this paper suggests that reports [2,3] of elaboration of trimethylarsoniolactate and its derivatives by unicellular marine algae require confirmation.

† Systematic name: 2,3-dihydroxypropyl 5-deoxy-5-(dimethylarsinoyl)-β-D-ribofuranoside.

Experimental

M.p.s was determined on a Kofler hot-stage apparatus. ^{1}H- and ^{13}C-N.m.r. spectra were recorded on Bruker instruments at 80 and 400 MHz (^{1}H) and 20.1 MHz (^{13}C). Tris refers to aminotris(hydroxymethyl)methane. All cellulose preparative- and thin-layer chromatographic (t.l.c.) plates were developed with butan-1-ol–acetic acid–water (60 : 15 : 25). After column chromatography on Sephadex DEAE, tris buffer was removed from eluted material, in each case, by passage through a Sephadex LH-20–water column. Arsenic was determined and located in chromatographic fractions by atomic absorption spectrophotometry.

Isolation of Arsenic Compounds from Kidneys.—Giant clams (*Tridacna maxima*) were collected from South Passage, Shark Bay, Western Australia and were deep-frozen during transit. After being thawed, the kidneys and adductor muscle were excised from 25 clams. Kidneys (380 g) were extracted with methanol (2 × 1.2 l) and then with water (4 × 200 ml) and the combined extracts were filtered to give a brown, powdery residue (52 g; 3 mg As) which was not further examined, and a



(1)

a; R = OSO_3H
b; R = OH
c; R = SO_3H

Systematic numbering scheme for compounds (1a—c)

Sephadex DEAE column (500 × 60 mm) equilibrated with 0.05M Tris buffer at pH 8.8. Isocratic elution separated the arsenic-containing fraction into two distinct fractions after elution with 9—11 and 17—20 l of eluant. The first of these (880 mg; 4 mg As) was shown to be a mixture of 3-arsenic-containing compounds [t.l.c.; cellulose (Merck; 0.1 mm pre-coated layers)] with arsenicals located at R_F 0.24, 0.32, and 0.39, and these were not further examined. The second fraction (950 mg; 16 mg As) was split into three equal portions and each portion was subjected to further chromatography on Sephadex DEAE (860 × 30 mm column; isocratic elution at pH 7.2 with 0.05M Tris buffer). After recombination of fractions containing arsenicals of interest, g.p.c. (Sephadex LH-20–methanol; 1 000 × 26 mm column) yielded a syrup (150 mg; 11.5 mg As).

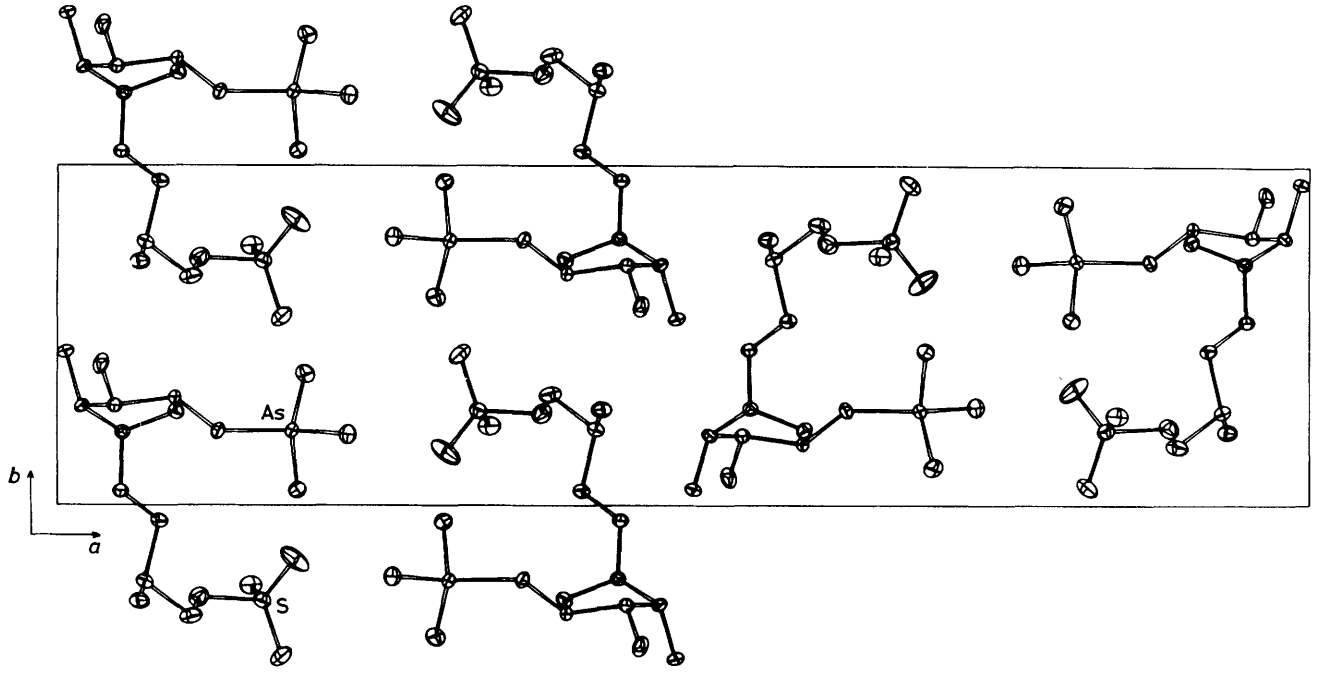


Figure 1. View of the unit cell of compound (1a) projected down *c*; 20% thermal ellipsoids are shown for the non-hydrogen atoms

filtrate which, on evaporation, yielded a dark, brown gum (63 g; 72.8 mg As). The gum was partitioned between water and diethyl ether; concentration of the organic layer gave a pale, yellow oil (2.18 g; 0.3 mg As) which was not further examined. The aqueous layer, in three portions, was subjected to gel-permeation chromatography (g.p.c.) on Sephadex LH-20 (389 g resin; 540 × 60 mm column; water as eluant). The arsenic-containing fractions were combined, adjusted to pH 7.6 with 0.05M Tris buffer, and applied to a column of Sephadex DEAE A25 (450 × 30 mm; equilibrated with 0.05M Tris at pH 7.6). Isocratic elution * separated the arsenicals into an acidic fraction (5.70 g; 31 mg As) and a non-acidic fraction (5.59 g; 14.6 mg As).

Isolation of acidic arsenic compound. The entire acidic fraction was subjected to repeated g.p.c. on a Sephadex LH-20–water column (820 × 50 mm) and was then applied to a Sephadex DEAE A25 column (250 × 26 mm) equilibrated at pH 8.8 with 0.05M Tris buffer. Isocratic elution yielded the arsenic-containing material as a pale, yellow, clear syrup (3.1 g; 20 mg As). This material was then dissolved in 0.05M Tris buffer (50 ml; pH 8.8) and the solution was applied to a

* Elution with buffer at the same molarity and pH as used to equilibrate the column.

This material was then sequentially subjected to p.l.c., first on silica [propan-1-ol–ammonia (7 : 3) as developer], the arsenical being located at R_F 0.35, and then on cellulose, R_F 0.28. Finally, the arsenic-containing material was passed through a small Sephadex LH-20–methanol column to yield, after work-up, a clear syrup (70 mg; 9 mg As). On dissolution in dry methanol and being kept overnight, crystals separated out, which had m.p. 163—175 °C (decomp.); δ_H (400 MHz; D_2O) 4.98 (1 H, s), 4.23 (2 H, m), 4.09 (1 H, d, *J* 3.3 Hz), 4.02 (3 H, m), 3.78 (1 H, dd, J_{gem} 11.0, J_{vic} 4.6 Hz), 3.56 (1 H, dd, J_{gem} 11.0 J_{vic} 3.3 Hz), 2.65 (1 H, dd, J_{gem} 14.8, J_{vic} 3.0 Hz), 2.50 (1 H, dd, J_{gem} 14.8, J_{vic} 10.0 Hz), 1.88 (3 H, s), and 1.83 (3 H, s). From these data we assigned structure (1a) to this compound.

Isolation of non-acidic arsenic compound. The non-acidic fraction was made up to 360 ml in 0.05M ammonium oxalate buffer (pH 2.4), and was chromatographed on a Dowex 50 × 4, 100—200 mesh column (250 × 16 mm) equilibrated with 0.05M ammonium oxalate buffer at pH 2.4. The buffer was removed by passage through Dowex 50 (H^+), the column being washed with water, and then the bases were eluted with 0.03M NH_4OH. In this way the non-acidic arsenic material was concentrated in the basic fraction (420 mg; 14.3 mg As). Further chromatography on Sephadex LH-20 (1 000 × 26 mm column; methanol as eluant) concentrated

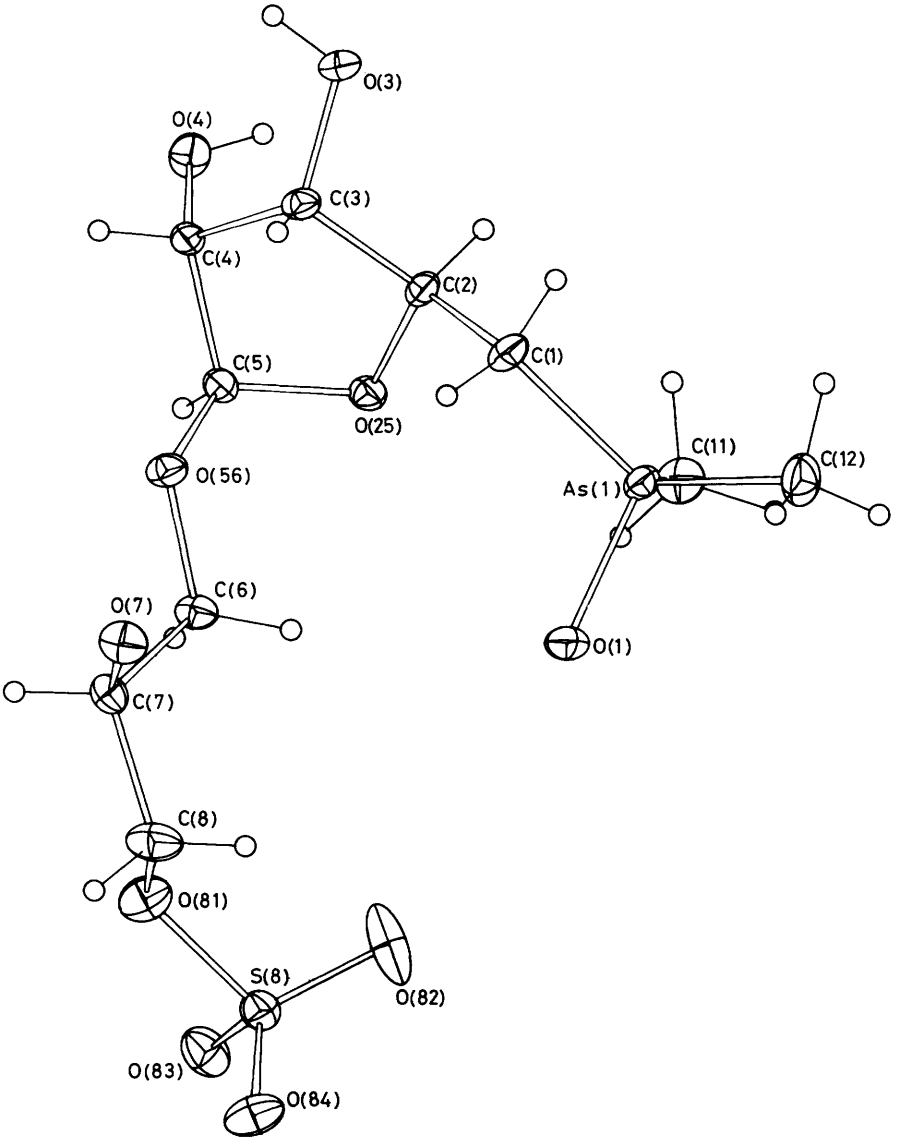


Figure 2. A projection of the molecule (1a) showing the non-hydrogen atom crystallographic labelling scheme. Hydrogen atoms have been given an arbitrary radius of 0.1 Å

the arsenic-containing material to a pale, orange syrup (135 mg; 9.7 mg As). This material was chromatographed on cellulose plates (200 × 200 × 1 mm, Whatman CC41). The arsenical fraction was located in a single band at R_F 0.51. Further p.l.c. on silica (200 × 200 × 1 mm, Merck silica G) with propan-2-ol–ethyl acetate–water (7 : 1 : 2) as developer, (R_F 0.14), followed by passage through a small Sephadex LH-20–methanol column, yielded a syrup (43.1 mg; 7.42 mg As). This material was shown, by co-t.l.c. on cellulose and silica and comparison of ^{1}H N.m.r. spectra, to be identical with the basic arsenic-containing sugar (1b) previously isolated from *Ecklonia radiata*. In addition, the ^{13}C n.m.r. spectrum of compound (1b) was in accord with the assigned structure: δ_C (CD_3OD) 109.4 [C(1′)], 78.4 [C(4′)] 77.6 [C(3′)], 76.2 [C(2′)], 72.1 [C(2)], 70.4 [C(3)], 64.1 [C(1)], 38.0 [C(5′)], and 15.7, 15.0 [As($Me)_2$].

Examination of Arsenic Compounds in Adductor Muscle.—Adductor muscles (540 g) were steeped in methanol for 14 d, then were homogenized with methanol (2 l) and the mixture was filtered. The residue was further extracted with methanol (1 l) and the combined filtrates were evaporated to dryness to give a brown syrup (34.6 g; <2 mg As) t.l.c. of which (cellulose) showed the presence of two arsenic-containing compounds (R_F 0.26 and 0.50) in the ratio *ca.* 1 : 2. These two compounds were also separable by g.p.c. (Sephadex LH-20–water) and their elution positions were identical with compounds (1a) and (1b) isolated from the kidneys (see above). Although there was insufficient material for isolation of the arsenic compounds from the adductor muscle, their chromatographic properties suggested identity with compounds (1a) and (1b) from the kidneys.

Crystallography

Crystal Data for Compound (1a).—$C_{10}H_{21}AsO_{10}S$, M = 407.9. Orthorhombic, space group $P2_12_12_1$ (D_2^4, No. 19), a = 29.85(2), b = 8.128(6), c = 6.365(4) Å, U = 1 544(2) Å^3. D_m = 1.79(1), D_c (Z = 4) = 1.79 g cm^{-3}. μ = 24.3 cm^{-1}, Mo-K_α radiation. Specimen: 0.18 × 0.09 × 0.26 mm.

Table 1. Some atomic co-ordinates for compound (1a) (standard deviations in parentheses)

Atom	*x*	*y*	*z*
As(1)	0.187 48(3)	0.220 7(1)	0.214 2(1)
O(1)	0.191 9(2)	0.046 1(7)	0.059 4(12)
C(1)	0.128 7(3)	0.218 5(12)	0.333 0(12)
C(11)	0.198 1(3)	0.387 4(12)	0.008 7(16)
C(12)	0.232 2(2)	0.209 4(14)	0.421 6(17)
C(2)	0.094 2(2)	0.318 4(9)	0.213 8(15)
O(25)	0.095 5(2)	0.276 5(8)	−0.002 9(8)
C(3)	0.046 3(2)	0.293 4(10)	0.288 9(12)
O(3)	0.035 1(2)	0.415 6(8)	0.436 5(11)
H(O3)	0.008 0(—)	0.423 6(—)	0.449 8(—)
C(4)	0.019 7(2)	0.292 1(11)	0.086 2(12)
O(4)	0.006 8(2)	0.451 5(7)	0.021 2(12)
H(O4)	0.024 0(—)	0.538 6(—)	0.018 0(—)
C(5)	0.052 5(2)	0.215 0(11)	−0.064 0(13)
O(56)	0.051 0(2)	0.042 7(7)	−0.040 7(10)
C(6)	0.082 1(3)	−0.043 5(10)	−0.170 6(13)
C(7)	0.070 6(3)	−0.224 4(11)	−0.166 9(12)
O(7)	0.066 2(2)	−0.282 4(8)	0.044 9(9)
C(8)	0.107 1(4)	−0.326 5(11)	−0.266 4(17)
O(81)	0.113 5(2)	−0.270 9(9)	−0.480 6(11)
S(8)	0.164 02(8)	−0.279 0(3)	−0.566 4(4)
O(82)	0.189 9(3)	−0.157 7(11)	−0.464 5(14)
O(83)	0.179 1(2)	−0.444 9(8)	−0.539 3(13)
O(84)	0.155 1(2)	−0.234 6(8)	−0.783 1(11)

Structure Determination.—Unique sets of *hkl* and $\bar{h}k\bar{l}$ data were measured at 295 K to the limit $2\theta_{max.} = 60°$ using a Syntex $P2_1$ four-circle diffractometer, using monochromatic Mo-K_α radiation ($\lambda = 0.710\,69$ Å), and operating in the conventional $2\theta/\theta$ scan mode. 2 502 Independent *hkl* reflections were obtained; 1 624 with $I > 2\sigma(I)$ were considered as 'observed' and were used in the full-matrix least-squares refinement after analytical absorption correction; for the $\bar{h}k\bar{l}$ set, 1 605 reflections were 'observed'. Because of profile overlap, reflections (14 ⟶ 38, 0, 1) and (22 ⟶ 28, 0, 2) and their $\bar{h}k\bar{l}$ equivalents were deleted from the refinement. Anisotropic thermal parameters were refined for the non-hydrogen atoms; hydrogen atoms were located clearly in difference maps and included at 'improved' estimates of (*x*, *y*, *z*) with U_H set at 1.25 U_{ii} (parent C, O) excepting the sulphate hydrogen *and* the hydrogen of the central glycerol hydroxy-group, neither of which were clearly defined and which are probably associated in hydrogen-bonding interactions associated with a number of short O · · · O contacts observed (see below). At convergence, *R* and *R'* (*hkl*) were 0.054 and 0.055, respectively, for the preferred chirality (corresponding to that of D-ribose), *R*,*R'* ($\bar{h}k\bar{l}$) being 0.059 and 0.061, respectively; for the opposite hand *R*,*R'* (*hkl*) were 0.068, 0.071 and *R*,*R'* ($\bar{h}k\bar{l}$) 0.070, 0.074, respectively. Reflection weights were $[\sigma^2(F_o) + 0.0005\ (F_o)^2]^{-1}$. Neutral-atom scattering factors were employed, all except those for H being corrected for anomalous dispersion (*f'*,*f''*); [10] computation was carried out using the *X*-RAY 76 program system [11] implemented on a Perkin-Elmer 3240 computer by S. R. Hall. The crystal packing in the unit cell (as viewed down the *c* axis) is shown in Figure 1.

Material deposited comprises tables of structure-factor amplitudes, thermal parameters, and hydrogen-atom parameters.* The non-hydrogen atom numbering scheme used for the crystallographic work is given in Figure 2.

Table 2. Bond lengths (Å) and bond angles (°) (standard deviations in parentheses)

Bond lengths	
As(1)–C(1)	1.910(8)
As(1)–O(1)	1.733(7)
As(1)–C(11)	1.88(1)
As(1)–C(12)	1.91(1)
C(1)–C(2)	1.51(1)
C(2)–C(3)	1.52(1)
C(2)–O(25)	1.42(1)
C(3)–C(4)	1.52(1)
C(3)–O(3)	1.41(1)
C(4)–C(5)	1.51(1)
C(4)–O(4)	1.41(1)
C(5)–O(25)	1.43(1)
C(5)–O(56)	1.41(1)
O(56)–C(6)	1.43(1)
C(6)–C(7)	1.51(1)
C(7)–O(7)	1.43(1)
C(7)–C(8)	1.51(1)
C(8)–O(81)	1.45(1)
O(81)–S(8)	1.604(7)
S(8)–O(82)	1.410(9)
S(8)–O(83)	1.431(7)
S(8)–O(84)	1.451(7)
Bond angles	
C(1)–As(1)–O(1)	106.7(4)
C(1)–As(1)–C(11)	115.5(4)
C(1)–As(1)–C(12)	112.0(4)
O(1)–As(1)–C(11)	100.3(4)
O(1)–As(1)–C(12)	107.7(4)
C(11)–As(1)–C(12)	113.4(4)
As(1)–C(1)–C(2)	114.9(6)
C(1)–C(2)–C(3)	114.2(7)
C(1)–C(2)–O(25)	109.7(6)
C(3)–C(2)–O(25)	107.4(6)
C(2)–C(3)–C(4)	103.1(7)
C(2)–C(3)–O(3)	109.9(6)
O(3)–C(3)–C(4)	116.6(7)
C(3)–C(4)–C(5)	101.6(6)
C(3)–C(4)–O(4)	112.7(7)
C(5)–C(4)–O(4)	111.9(7)
C(4)–C(5)–O(25)	105.4(6)
C(4)–C(5)–O(56)	109.0(7)
O(25)–C(5)–O(56)	110.4(6)
C(2)–O(25)–C(5)	108.8(5)
C(5)–O(56)–C(6)	113.9(6)
O(56)–C(6)–C(7)	108.7(6)
C(6)–C(7)–C(8)	111.4(7)
C(6)–C(7)–O(7)	110.8(7)
C(8)–C(7)–O(7)	106.2(7)
C(7)–C(8)–O(81)	108.6(7)
C(8)–O(81)–S(8)	115.6(6)
O(81)–S(8)–O(82)	109.2(5)
O(81)–S(8)–O(83)	107.0(4)
O(81)–S(8)–O(84)	98.1(4)
O(82)–S(8)–O(83)	115.6(5)
O(82)–S(8)–O(84)	111.4(5)
O(83)–S(8)–O(84)	114.0(4)

Table 1 gives the non-hydrogen atomic co-ordinates and Table 2 gives bond distances and angles; skeletal dimensions are largely as expected but the following features of interest are noted. (i) The ribose ring adopts an evelope conformation with C(4) lying well out of the plane (by 0.55 Å) of the other four atoms (σ 0.02 Å). (ii) Irregularities are observed in the

* Supplementary Publication No. SUP 23417 (9 pp.). For details see Notice to Authors No. 7, *J. Chem. Soc., Perkin Trans. 1*, 1981, Index issue.

geometry about the arsenic and sulphur atoms, notably in regard to 'equivalent' angles. While in part this may be a result of unsymmetrical interactions of the substituents with nearby methylene hydrogen atoms in the adjoining chains, it may also be a consequence of (iii) considerable hydrogen bonding in the lattice involving not only hydroxy-groups but also the arsine oxide and sulphate groups. O ··· O contacts <3.0 Å are: O(1) ··· O(84) (x, y, $1 + z$), 2.724(9); O(3) ··· O(56) ($\bar{x}$, $\frac{1}{2} + y$, $\frac{1}{2} - z$), 2.847(9); O(4) ··· O(7) (x, $1 + y$, z), 2.801(9); and O(7) ··· O(84) (x, y, $1 + z$), 2.895(9) Å.

Acknowledgements

We thank Mrs. S. Slack-Smith for identifying specimens of *Tridacna maxima* and Dr. M. J. Crossley for obtaining the 400 MHz ^{1}H n.m.r. spectra.

References

1 C. M. Yonge, *Sci. Rep. Gt. Barrier Reef Exp. Brit. Mus.* (*Nat. Hist.*), 1937, **1**, 283.
2 A. A. Benson and R. E. Summons, *Science*, 1981, **211**, 482.
3 R. V. Cooney, R. O. Mumma, and A. A. Benson, *Proc. Natl. Acad. Sci. USA*, 1978, **75**, 4262.
4 J. S. Edmonds and K. A. Francesconi, *Nature* (*London*), 1981, **289**, 602.
5 J. R. Cannon, J. S. Edmonds, K. A. Francesconi, C. L. Raston, J. B. Saunders, B. W. Skelton, and A. H. White, *Aust. J. Chem.*, 1981, **34**, 787.
6 J. S. Edmonds and K. A. Francesconi, *Mar. Pollut. Bull.*, 1981, **12**, 92.
7 J. S. Edmonds, K. A. Francesconi, and J. A. Hansen, *Experientia*, 1982, **38**, 643.
8 J. S. Edmonds and K. A. Francesconi, *Chemosphere*, 1981, **10**, 1041.
9 S. Kurosawa, K. Yasuda, M. Taguchi, S. Yamazaki, S. Toda, M. Morita, T. Uehiro, and K. Fuwa, *Agric. Biol. Chem.*, 1980, **44**, 1993.
10 'International Tables for X-ray Crystallography,' ed. J. A. Ibers and W. C. Hamilton, Kynoch Press, Birmingham, 1974, Vol. 4.
11 'The X-ray System—Version of March 1976,' Technical Report, TR-466, Computer Science Center, University of Maryland, USA (ed. J. M. Stewart).

Received 8th April 1982; *Paper* 2/604