Table 12. *Alternate hydrogen positions*

| H attached to | *x* | *y* | *z* | X–H distance | Peak height (arbitrary units) |
|---|---|---|---|---|---|
| $SO_3$ | 0·0500 | 0·3000 | 0·0005 | 0·815 Å | 113 |
| N(2) | 0·3680 | 0·1267 | 0·4200 | 0·698 | 114 |

Hamilton, W. C. & Ibers, J. A. (1968). *Hydrogen Bonding in Solids*, p. 208. New York: Benjamin.

Hauptman, H. (1972). *Crystal Structure Determination – The Role of Cosine Invariants*. New York: Plenum Press.

Hauptman, H. & Karle, J. (1953). *Solution of the Phase Problem*. I. *The Centrosymmetric Crystal*, A.C.A. Monograph No. 3.

Karle, J. & Hauptman, H. (1956). *Acta Cryst*. **9**, 635–651.

Kitaigorodskii, A. I. (1961). *Organic Chemical Crystallography*, p. 439. New York: Consultants Bureau.

Leuchs, H. & Rauch, H. (1914). *Ber. Dtsch. Chem. Ges.* **47**, 370–380.

Leuchs, H. & Schwaebel, G. (1914). *Ber. Dtsch. Chem. Ges.* **47**, 1552–1560.

Leuchs, H. & Schwaebel, G. (1915). *Ber. Dtsch. Chem. Ges.* **48**, 1009–1015.

Leuchs, H. & Weber, L. E. (1909). *Ber. Dtsch. Chem. Ges.* **42**, 770–777.

Manske, R. H. F. (1965). *The Alkaloids*. Vol. 8, Chapter 17.

Robertson, J. H. & Beevers, C. A. (1950). *Nature, Lond.* **165**, 690–691.

Robinson, R. & Stephen, A. M. (1948). *Nature, Lond.* **162**, 177–178.

Stout, G. H. & Jensen, L. H. (1968). *X-ray Structure Determination*, p. 370. London: Macmillan.

Sutor, D. J. (1962). *Nature, Lond.* **195**, 68–69.

Woodward, R. B. & Brehm, B. J. (1948). *J. Amer. Chem. Soc.* **70**, 2107–2115.

---

*Acta Cryst.* (1973). B**29**, 2770

# The Crystal and Molecular Structure of an Aldotriouronic Acid Trihydrate: *O*-(4-*O*-Methyl-α-D-Glucopyranosyluronic Acid)-(1-2)-*O*-β-D-xylopyranosyl-(1-4)-D-xylopyranose Trihydrate

By Robert A. Moran and Gerald F. Richards

*The Institute of Paper Chemistry, Appleton, Wisconsin* 54911, *U.S.A.*

(*Received* 6 *November* 1972; *accepted* 1 *August* 1973)

The crystal structure of an aldotriouronic acid trihydrate [*O*-(4-*O*-methyl-α-D-glucopyranosyluronic acid)-(1-2)-*O*-β-D-xylopyranosyl-(1-4)-D-xylopyranose trihydrate] has been determined from three-dimensional X-ray data obtained from crystals maintained at a temperature of approximately −193°C. The structure was solved by a combination of direct acentric phasing methods and Patterson search methods, and the anisotropic refinement was terminated at a final *R* index of 0·066 for 2232 observations. One trisaccharide molecule and three water molecules of crystallization ($C_{17}H_{28}O_{15}.3H_2O$) comprise the asymmetric unit in a monoclinic unit cell with the low-temperature cell parameters of $a = 10·136$ (3), $b = 10·925$ (8), $c = 10·120$ (3) Å, $\beta = 91·66$ (3)°, and space group $P2_1$. All three rings in the trisaccharide molecule were found to have the *C*1 chair conformation with all the substituents, except C(1″)–O(2), in an equatorial position. The trisaccharide molecule has no intramolecular hydrogen bonds. The xylose rings in the xylobiose residue are arranged in a left-handed helical conformation with a glycosidic oxygen angle of 113·8° between the two rings. The glycosidic oxygen angle between the two rings linked α-1,2 in the aldobiouronic acid residue is 116·0°. Orientation of the trisaccharide in the crystal is such that the xylobiose residue extends along the *b* axis and the aldobiouronic acid residue extends along the *c* axis. The three water molecules are clustered between the trisaccharide molecules and participate in 10 of the 15 unique hydrogen bonds in the crystal structure.

## Introduction

The aldotriouronic acid trisaccharide is a segment of the naturally occurring 4-*O*-methylglucuronoxylans found in hardwood hemicelluloses. The acid trisaccharide contains two different disaccharides: xylobiose (two xylose units connected by a β-1,4-linkage) and an aldobiouronic acid (a 4-*O*-methyl-D-glucuronic acid ring attached to the nonreducing xylose ring by an α-1,2-linkage), neither of which have been studied crystallographically. The conformation of the xylobiose residue of the trisaccharide was of interest because it

---

* A portion of a thesis submitted by Robert A. Moran in partial fulfillment of the requirements of The Institute of Paper Chemistry for the degree of Doctor of Philosophy from Lawrence University, Appleton, Wisconsin, June 1972.

could be compared to the minimum energy conformation position defined by potential energy functions for a xylobiose disaccharide (Sundararajan & Rao, 1969). In addition, the orientation of the two rings could also be compared to the conformation of the other β-1,4-linked disaccharide structures, cellobiose (Chu & Jeffrey, 1968), methyl β-cellobioside (Ham & Williams, 1970), and α-lactose (Fries, Rao & Sundaralingam, 1971), all of which contain an intramolecular hydrogen bond between O(3′) and O(5). An examination of the orientation of the 4-*O*-methyl-D-glucuronic acid ring with respect to the xylobiose backbone was also of interest in that it could provide some information as to how this ring might be oriented in the polymer as a single branched unit. Although the α-1,2 glycosidic link has been studied in sucrose (Brown & Levy, 1963) and in the sucrose residue in raffinose (Berman, 1970), planteose (Rohrer, 1972), and l-ketose (Jeffrey & Park) it has not been studied previously between two six-membered rings.

## Crystal data

The aldotriouronic acid samples used in this work were obtained from Dickey (1969). Large, clear, twinned crystals were grown in a bacteria-free environment at room temperature conditions by slow evaporation from a methanol–water solvent mixture. Crystals suitable for data collection were obtained by fracturing the large twinned crystals [the twin plane is the (001) plane] and then shaping the fragments into near-rectangular specimens. Table 1 gives the crystal data.

Table 1. *Crystal data**

| | 20°C | −193°C |
|---|---|---|
| $a$ | 10·172 (2) Å | 10·136 (3) Å |
| $b$ | 10·947 (8) | 10·925 (8) |
| $c$ | 10·288 (3) | 10·120 (3) |
| $\beta$ | 91·04 (14)° | 91·66 (3)° |
| $V$ | 1145·4 (9) Å$^3$ | 1120·1 (9) Å$^3$ |
| $D_f$ | 1·528 (1) g cm$^{-3}$ | |
| $D_x$ | 1·525 (1) g cm$^{-3}$ | 1·560 (1) g cm$^{-3}$ |
| $\mu$(Cu $K\alpha$) | 12·35 cm$^{-1}$ | 12·62 cm$^{-1}$; |
| Melting point | 179 (3)°C | |
| Molecular formula | $C_{17}H_{28}O_{15}.3H_2O$ | |
| M.W. | 526·34 | |
| Crystal habit | Lancet-like | |
| Crystal system | Monoclinic | |
| Space group | $P2_1$ (from systematic absences $0k0$ with $k=2n+1$) | |
| $Z$ | 2 | |

* The estimated standard deviations are in parentheses and refer to the last decimal position of the respective value.

The unit-cell dimensions were determined from the high-angle Cu $K\alpha$ and Cu $K\beta$ reflections on back-reflection photographs by least-squares procedures. The experimental density was determined by flotation in benzene and carbon tetrachloride.

## Experimental

Fourteen layers of intensity data consisting of 7 layers (0–6) from the $a$ axis and 7 layers (0–6) from the $c$ axis were recorded at a temperature near −193°C by the multiple-film equi-inclination Weissenberg technique and nickel-filtered Cu $K\alpha$ ($\lambda=1·54178$ Å) radiation. The three crystals used for collecting the intensity data, two of which were needed to complete the 7-layer $a$-axis set, were approximately cubic in shape and had edge lengths ranging from 0·17 to 0·31 mm. The crystals were mounted in capillaries and maintained at a temperature near the boiling point of liquid nitrogen with a gas flow cryostat system similar to the one described by Richards (1964). The intensities were estimated visually by comparison with a calibrated film strip. The data from both axes were corrected for spot-shape, Lorentz, and polarization factors with a program written by Norment (1962). Absorption corrections were not made. The method of Rollett & Sparks (1960) was used to calculate interlayer scale factors and to place the structure-factor amplitudes on the same relative basis. 2456 unique reflections were measured of which 195 had intensities below that of the background. Normalized structure-factor magnitudes were calculated from the relative observed structure-factor magnitudes using the coefficients obtained from the $K$-curve method (Karle & Hauptman, 1956). The computed values of the statistical averages $\langle|E|\rangle$ and $\langle|E^2-1|\rangle$ were 0·88 and 0·74, respectively, for the 2456 observed and unobserved reflections.

## Solution of the crystal structure

A phasing model large enough to initiate the solution of the crystal structure was established by using the structure invariants method in a manner similar to that described by Duax, Weeks & Hauptman (1970). Invariants were calculated for 763 vector triples with the *MDKS* formula [equation (1)] at $E$-threshold levels ($t$) of 1·3 and 1·5. The vector triples were obtained from 145 centrosymmetric reflections with $A=2\sigma_3|E_{h_1}E_{h_2}E_{h_3}|/\sigma_2^{3/2}$ and $|E|$ lower limits of 0·5.

$$\cos(\varphi_{h_1}+\varphi_{h_2}+\varphi_{h_3})=M(D-KS) \qquad (1)$$

where

$$D=\langle(|E_{-h3+k}|^2-1)||E_k|>t,|E_{h1+k}|>t\rangle_k$$

and

$$S=\langle(|E_{-h3+k}|^2-1)||E_k|>t\rangle_k + \langle(|E_{-h3+k}|^2-1)||E_{h1+k}|>t\rangle_k\,.$$

The 1·3 and 1·5 threshold levels allowed 449 and 279 reflections to enter into the calculations. Vector triples were considered risks and were not used in the subsequent phase determination procedure if their calculated invariants changed greatly as a function of the number of contributors in the *MDKS* formula (Karle, 1970). $K$ and $M$ scale factors were calculated within groups of data of approximately constant $A$ values. Construction of the vector triples and normalization of the calculated invariants was performed with a set of computer programs written by Weeks (1969).

Two centrosymmetric origin reflections ($\bar{1}02$ and 703) and three centrosymmetric structure invariant reflections (606, $\bar{2}02$, and 406) were used to determine phases from the list of vector triples and their corresponding invariants. Systematic searches through the list to an $A$ level of 1·0, introducing each of the three structure invariant reflections when necessary and requiring that the calculated invariants be greater than unity and that there be no conflicting phase indications, allowed phases to be determined for a total of 45 centrosymmetric reflections. Eleven of these were structure invariant reflections and had probability predictions regarding their phase assignments (Killean, 1966). Only 25 of the 45 centrosymmetric reflections had $E$ values large enough to assist in the early stages of tangent-formula phase expansion and refinement.

When phase values of $\pi$ were assigned to both centrosymmetric origin reflections and phase values of 0, 0, and $\pi$ were assigned to the structure invariant reflections (606, $\bar{2}02$, and 406, respectively), a set of internally consistent phases were derived for the 25 centrosymmetric reflections with $E$ values greater than 1·4. These centrosymmetric data and the data indicated in Table 2 were used as a starting set for tangent-formula expansion and refinement. The symbol reflection, 395, was tested in quadrants and fixed at $-135°$. Six cycles of tangent-formula refinement with 5 iterations per cycle at threshold levels from 1·6 to 1·37 gave 375 phased $|E|$'s. The corresponding $E$ map showed a rough image of a pyranose ring with the plane of the ring nearly perpendicular to the $x=z$ plane, and with most of the peaks located on or very near the $x=z$ diagonal. Two large peaks located at either end of the ring, later shown to be O(1) and O(1′), were related by a pseudosymmetry.

The set of interatomic vectors belonging to the pyranose ring was compared to the set of vectors around the origin of a sharpened three-dimensional Patterson map in order to verify the validity of the ring's orientation. The placement of the ring in the unit cell was checked by comparing the coordinates of the ring peaks with those corresponding to several possible ring positions identified by using a 10 atom β-D-xylose model in a vector verification search program written by Braun, Hornstra & Leenhouts (1969). After confirmation of these two points, the structure was completed by the tangent-formula recycling technique described by Karle (1968). Three cycles based on phase models from 6-, 8-, and 18-atom connected fragments gave the 32-atom L-enantiomorphic form of the trisaccharide molecule and the three oxygen atoms present in the crystal structure as water molecules of crystallization. The structure was converted from the L-enantiomorph to the D-enantiomorph before refinement. Reexamination of the initial $E$ map showed that peaks were present corresponding to all 35 atoms in the asymmetric unit.

A special feature of the aldotriouronic acid trihydrate crystal, which may have been responsible for some of the difficulty experienced in applying the symbolic addition method and the multiple-solution approach of direct methods to the phase problem, was noticed after the structure was solved. Three pairs of atoms in the trisaccharide structure were found to have coordinates such that when one of the pair was subjected to a $P2_1$ symmetry operation, a $P2$ pseudosymmetry resulted. Two of the three pairs are in the xylobiose residue [C(2′) and O(5), and O(1) and O(1′)], whereas the third pair is in the glucuronic acid ring [C(2″) and O(7″)]. The $P2$ pseudosymmetry was a perplexing problem because when it was present in combination with the normal $P2_1$ symmetry, the resulting $E$-map often contained a mirror plane normal to the $b$ axis and located halfway between the atoms of some of the pairs. There was also evidence for the existence of vector overlap in the crystal structure. The average values of $(|E|^2-1)^2$ and $(|E|^2-1)^3$ for all reflections were 1·32 and 4·34, as compared to the theoretical values of 1·04 and 2·21 (Hauptman, Fisher, Hancock & Norton, 1969).

## Structure refinement

Least-squares refinement of the atomic positional and thermal parameters was carried out by using the block-diagonal approximation in a program written by Ahmed (Ahmed, Hall, Pippy & Huber, 1966). This program is part of the NCR programs used throughout this investigation. Five cycles of refinement with isotropic temperature parameters followed by three cycles with anisotropic parameters, with unit weights applied to the observed reflections and the acceleration factors of Hodgson & Rollett (1963) applied to the parameter shifts, reduced $R\,[=\sum|(|F_o|-|F_c|)|/\sum|F_o|]$ to 0·092 and 0·084, respectively. The hydrogen atoms were located

Table 2. *Starting set reflections used in E-map preparation*

| *hkl* | $\lvert E\rvert$ | Assignment | Phase assignment | Number of interactions $\lvert E\rvert>2·0$ | $\lvert E\rvert>1·5$ |
|---|---|---|---|---|---|
| $\bar{1}02$ | 2·02 | origin | $\pi$ | 12 | 118 |
| 703 | 2·08 | origin | $\pi$ | 18 | 64 |
| 018 | 2·52 | origin | $-\pi/2$ | 2 | 51 |
| 482 | 3·36 | enantiomorph | $-\pi/2$ | 13 | 73 |
| 395 | 1·85 | symbol | variable | – | 39 |

Table 3. *Fractional atomic coordinates and temperature factors for the carbon and oxygen atoms*

The anisotropic temperature factors and the estimated standard deviations in parentheses for the anisotropic temperature factors and the fractional coordinates are multiplied by $10^5$. The expression for the temperature factor exponent which is consistent with the $B$ values is $\exp[-(h^2B_{11}+k^2B_{22}+l^2B_{33}+klB_{23}+hlB_{13}+hkB_{12})]$.

| | $x$ | $y$ | $z$ | $B_{11}$ | $B_{22}$ | $B_{33}$ | $B_{23}$ | $B_{13}$ | $B_{12}$ | $B_t$ |
|---|---|---|---|---|---|---|---|---|---|---|
| C(1) | −0·09863 (51) | 0·14253 (48) | −0·27317 (51) | 164 (42) | 115 (38) | 142 (43) | −56 (71) | 30 (76) | 175 (66) | 0·429 |
| C(2) | −0·12145 (50) | 0·02462 (46) | −0·19642 (47) | 127 (41) | 131 (38) | 17 (39) | 20 (62) | 25 (73) | −13 (64) | 0·381 |
| C(3) | 0·00095 (46) | −0·05547 (45) | −0·20343 (48) | 75 (39) | 50 (35) | 125 (40) | 13 (66) | 62 (71) | −45 (64) | 0·342 |
| C(4) | 0·04344 (49) | −0·07364 (51) | −0·34564 (53) | 113 (42) | 186 (40) | 204 (46) | −50 (74) | 35 (79) | 101 (70) | 0·595 |
| C(5) | 0·05019 (52) | 0·04878 (54) | −0·41797 (52) | 185 (46) | 229 (41) | 189 (45) | 127 (79) | 363 (81) | 139 (78) | 0·898 |
| O(1) | −0·21308 (34) | 0·21300 (33) | −0·27078 (35) | 117 (30) | 79 (26) | 189 (32) | 63 (49) | 90 (55) | 0 (55) | 0·417 |
| O(2) | −0·14288 (32) | 0·05223 (34) | −0·06052 (32) | 81 (25) | 132 (27) | 55 (29) | 6 (49) | −41 (51) | 49 (51) | 0·391 |
| O(3) | −0·02564 (33) | −0·17031 (32) | −0·14084 (35) | 145 (31) | 57 (25) | 171 (31) | 155 (48) | −65 (55) | 33 (48) | 0·496 |
| O(4) | 0·17271 (36) | −0·12398 (35) | −0·35308 (36) | 253 (33) | 128 (28) | 181 (32) | 158 (52) | 208 (58) | 185 (52) | 0·793 |
| O(5) | −0·07342 (36) | 0·11273 (36) | −0·40893 (35) | 264 (33) | 149 (28) | 100 (31) | 40 (48) | 141 (58) | 283 (52) | 0·714 |
| C(1′) | −0·25386 (52) | 0·58772 (47) | −0·22280 (51) | 172 (45) | 89 (38) | 173 (45) | −71 (67) | −180 (82) | 81 (64) | 0·486 |
| C(2′) | −0·33020 (51) | 0·53271 (51) | −0·33918 (53) | 149 (44) | 197 (42) | 203 (46) | 197 (74) | 150 (81) | 116 (70) | 0·653 |
| C(3′) | −0·32735 (48) | 0·39151 (47) | −0·33666 (49) | 98 (41) | 136 (39) | 106 (42) | −40 (67) | 19 (74) | −92 (66) | 0·502 |
| C(4′) | −0·19180 (48) | 0·34145 (48) | −0·29846 (50) | 115 (42) | 124 (39) | 154 (44) | −33 (69) | 126 (78) | 38 (67) | 0·518 |
| C(5′) | −0·13666 (55) | 0·40881 (48) | −0·17776 (56) | 307 (49) | 61 (38) | 249 (47) | 182 (68) | −315 (87) | −100 (70) | 0·771 |
| O(1′) | −0·23909 (40) | 0·71273 (36) | −0·24712 (42) | 326 (37) | 88 (28) | 416 (40) | 68 (58) | −537 (67) | 11 (55) | 0·977 |
| O(2′) | −0·46161 (37) | 0·57658 (37) | −0·33561 (39) | 157 (32) | 223 (32) | 301 (36) | −89 (55) | −199 (61) | 54 (52) | 0·842 |
| O(3′) | −0·37201 (36) | 0·34352 (36) | −0·46225 (37) | 228 (33) | 165 (30) | 209 (33) | −132 (53) | −71 (59) | −60 (53) | 0·752 |
| O(5′) | −0·12684 (36) | 0·53584 (34) | −0·21064 (37) | 199 (32) | 105 (27) | 259 (34) | 184 (52) | −362 (59) | −29 (50) | 0·757 |
| C(1″) | −0·27582 (47) | 0·06173 (45) | −0·02436 (50) | 69 (40) | 53 (36) | 180 (43) | 60 (69) | −58 (75) | 29 (64) | 0·322 |
| C(2″) | −0·27945 (47) | 0·13301 (47) | 0·10409 (49) | 94 (40) | 90 (37) | 154 (44) | −37 (67) | −183 (77) | −64 (66) | 0·435 |
| C(3″) | −0·20351 (50) | 0·06248 (46) | 0·21215 (48) | 189 (44) | 49 (37) | 126 (43) | −92 (68) | −1 (77) | −100 (68) | 0·434 |
| C(4″) | −0·26447 (48) | −0·06435 (48) | 0·22448 (48) | 138 (42) | 117 (38) | 86 (41) | 150 (65) | −9 (75) | 5 (68) | 0·493 |
| C(5″) | −0·26785 (48) | −0·13026 (46) | 0·09019 (47) | 107 (40) | 51 (35) | 148 (44) | −17 (68) | −302 (74) | 44 (63) | 0·424 |
| C(6″) | −0·34966 (50) | −0·24763 (47) | 0·09466 (47) | 184 (43) | 116 (38) | 21 (40) | −19 (65) | −222 (73) | −21 (67) | 0·426 |
| C(7″) | −0·25889 (60) | −0·17667 (56) | 0·42658 (63) | 390 (54) | 189 (44) | 389 (55) | 151 (84) | 219 (98) | 10 (83) | 1·308 |
| O(2″) | −0·22142 (35) | 0·25062 (33) | 0·08710 (36) | 176 (31) | 55 (26) | 224 (34) | −15 (50) | −80 (57) | −58 (47) | 0·547 |
| O(3″) | −0·21576 (37) | 0·13063 (34) | 0·33286 (35) | 359 (36) | 108 (28) | 92 (31) | −38 (49) | −203 (59) | 71 (54) | 0·565 |
| O(4″) | −0·18616 (35) | −0·13193 (35) | 0·31871 (34) | 208 (32) | 155 (28) | 131 (31) | 56 (53) | −182 (56) | −23 (51) | 0·596 |
| O(5″) | −0·33294 (33) | −0·05601 (33) | −0·00813 (35) | 95 (28) | 71 (26) | 172 (30) | 21 (48) | −166 (52) | −74 (46) | 0·416 |
| O(6″) | −0·46717 (36) | −0·24887 (36) | 0·07465 (41) | 167 (32) | 168 (29) | 398 (39) | 99 (57) | −328 (62) | −114 (51) | 0·874 |
| O(7″) | −0·27630 (36) | −0·34479 (34) | 0·12673 (39) | 222 (34) | 96 (28) | 323 (36) | 38 (56) | −81 (61) | −6 (52) | 0·732 |
| O(*A*) | −0·47014 (42) | 0·10593 (40) | −0·40650 (42) | 434 (42) | 278 (34) | 334 (39) | −22 (63) | −363 (70) | −198 (63) | 1·384 |
| O(*B*) | 0·39988 (35) | −0·06313 (36) | −0·11039 (36) | 178 (32) | 171 (30) | 187 (32) | 5 (53) | −101 (56) | 93 (52) | 0·700 |
| O(*C*) | −0·45353 (39) | −0·13414 (39) | −0·31730 (38) | 351 (38) | 228 (32) | 253 (35) | −6 (59) | −171 (64) | 149 (57) | 1·049 |

by interpretation of successive Fourier difference syntheses and were assigned the isotropic temperature factor of the carbon or oxygen atom to which they were bonded. Two additional cycles of anisotropic refinement with the temperature factors and coordinates of the hydrogen atoms fixed reduced $R$ to 0·067. The scattering factors for the carbon and oxygen atoms used in the refinement were calculated by Hanson, Herman, Lea & Skillman (1964). The scattering factor curve calculated by Stewart, Davidson & Simpson (1965) was used for the hydrogen atoms.

Least-squares curve-fitting methods were used to determine coefficients of a polynomial expression on $|F_o|$, and then on $\sin\theta$, in order to correctly assign relative weights to the measured set of structure-factor magnitudes. The appropriateness of the weighting scheme was judged by its ability to maintain the weighted residual, $\omega(\Delta F)^2$, constant, as a function of $|F_o|$ and $\sin\theta$. The most intense reflection, $\bar{2}02$, believed to be affected by secondary extinction, was given the weight of zero. Twenty other reflections were also found to have unacceptable discrepancies and were given weights of zero (these reflections are noted with a ** in Table 5). Six cycles of anisotropic refinement using the observed reflections and the weighting function described by equation (2), with the constants $C_1$ to $C_6$ at values of

$$w=1/[(C_1|F_o|^2+C_2|F_o|+C_3) \times (C_4\sin^2\theta+C_5\sin\theta+C_6)] \quad (2)$$

0·0021, 0·041, 1·01, 5·24, −6·71, and 2·95, reduced $R$ to 0·066. Least-squares refinement was terminated when the average shift in the parameters was less than $\frac{1}{30}$ of the estimated standard deviations of the atomic coordinates. The final positional coordinates and thermal parameters for the carbon and oxygen atoms and their estimated standard deviations are listed in Table 3. The hydrogen atom coordinates and bond lengths are listed in Table 4. The calculated and observed structure factor magnitudes and the phases for each reflection are listed in Table 5.

## Description of the structure

### *Nomenclature of the aldotriouronic acid trihydrate complex*

The formal nomenclature for the aldotriouronic acid molecule is *O*-(4-*O*-methyl-α-D-glucopyranosyluronic acid)-(1-2)-*O*-β-D-xylopyranosyl-(1-4)-D-xylopyranose. The atoms of the β-1,4-linked xylobiose residue in the trisaccharide are labeled according to the convention used in other β-1,4-linked disaccharides such as β-cellobiose (Chu & Jeffrey, 1968), methyl β-cellobioside (Ham & Williams, 1970), and α-lactose (Fries, Rao & Sundaralingam, 1971). The disaccharide numbering convention requires that the nonreducing ring be labeled as the unprimed ring (referred to as ring I) and the reducing ring be labelled as the primed ring (referred to as ring II). The glucuronic acid ring in the trisaccharide is treated as an aglycone and labeled

Table 4. *Hydrogen atom positions and bond lengths*

| Ring | Atom | Fractional coordinates |  |  | Bond length (Å) |  |
|---|---|---|---|---|---|---|
|  |  | $x$ | $y$ | $z$ | C–H | O–H |
| I | H(1) | −0·008 | −0·180 | −0·248 | 1·033 |  |
|  | H(2) | −0·194 | −0·020 | −0·239 | 0·971 |  |
|  | H(3) | 0·073 | −0·022 | −0·155 | 0·936 |  |
|  | H(4) | −0·028 | −0·125 | −0·402 | 1·073 |  |
|  | H(5–1) | 0·084 | 0·046 | −0·510 | 1·000 |  |
|  | H(5–2) | 0·120 | 0·103 | −0·397 | 0·943 |  |
|  | H(03) | 0·042 | −0·192 | −0·088 |  | 0·891 |
|  | H(04) | 0·181 | −0·203 | −0·328 |  | 0·905 |
| II | H(1′) | −0·314 | 0·578 | −0·144 | 1·025 |  |
|  | H(2′) | −0·283 | 0·551 | −0·421 | 0·988 |  |
|  | H(3′) | −0·392 | 0·365 | −0·275 | 0·961 |  |
|  | H(4′) | −0·117 | 0·354 | −0·369 | 1·063 |  |
|  | H(5′–1) | −0·201 | 0·404 | −0·099 | 1·040 |  |
|  | H(5′–2) | −0·035 | 0·386 | −0·161 | 1·070 |  |
|  | H(01′) | −0·171 | 0·728 | −0·207 |  | 0·812 |
|  | H(02′) | −0·512 | 0·569 | −0·399 |  | 0·807 |
|  | H(03′) | −0·319 | 0·359 | −0·521 |  | 0·833 |
| III | H(1″) | −0·342 | 0·090 | −0·088 | 0·969 |  |
|  | H(2″) | −0·363 | 0·132 | 0·133 | 0·904 |  |
|  | H(3″) | −0·097 | 0·051 | 0·193 | 1·105 |  |
|  | H(4″) | −0·342 | −0·061 | 0·278 | 0·961 |  |
|  | H(5″) | −0·170 | −0·157 | 0·082 | 1·045 |  |
|  | H(02″) | −0·283 | 0·306 | 0·091 |  | 0·869 |
|  | H(03″) | −0·152 | 0·114 | 0·407 |  | 0·997 |
|  | H(7″–1) | −0·306 | −0·231 | 0·394 | 0·824 |  |
|  | H(7″–2) | −0·200 | −0·238 | 0·480 | 1·036 |  |
|  | H(7″–3) | −0·324 | −0·119 | 0·470 | 1·023 |  |
|  | H(07″) | −0·331 | −0·408 | 0·123 |  | 0·888 |
| Water |  |  |  |  |  |  |
| *A* | H(O*A*–1) | −0·437 | 0·042 | −0·375 |  | 0·839 |
|  | H(O*A*–2) | −0·399 | 0·153 | −0·431 |  | 0·922 |
| *B* | H(O*B*–1) | 0·459 | −0·080 | −0·056 |  | 0·821 |
|  | H(O*B*–2) | 0·450 | −0·071 | −0·182 |  | 0·905 |
| *C* | H(O*C*–1) | −0·526 | −0·146 | −0·378 |  | 0·947 |
|  | H(O*C*–2) | −0·391 | −0·194 | −0·315 |  | 0·913 |

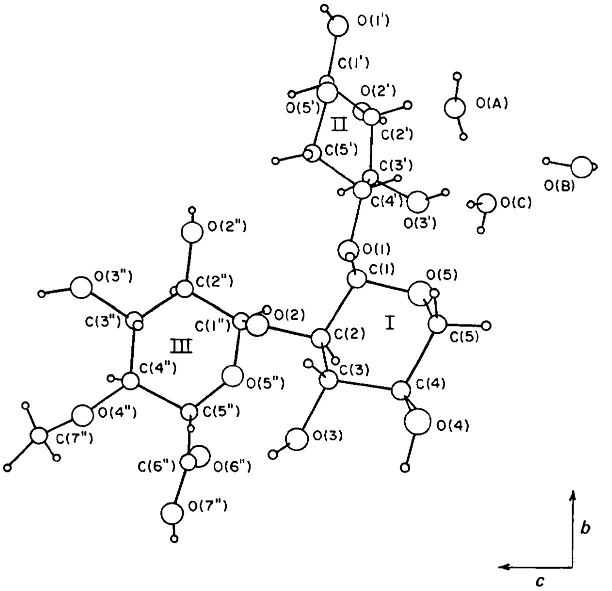


Fig. 1. The configuration and numbering of the atoms in one asymmetric unit of the aldotriouronic acid trihydrate crystal.

Table 5. *Observed and calculated structure-factor magnitudes and their phase angles*

The column headings are $k$, 10 $F_o$, 10 $F_c$, and $\varphi$ in degrees. The unobserved reflections are marked with one * and the reflections excluded from least-squares refinement are marked with a double **.

| K | FO | FC | ALPHA |
|---|---|---|---|
| [illegible] | [illegible] | [illegible] | [illegible] |

with double primes (referred to as ring III). The configuration and identification of the atoms in one asymmetric unit of the crystal is shown in Fig. 1. The three water molecules in the asymmetric unit are labeled *A*, *B*, and *C*.

### *Bond distances*

The bond distances and bond angles involving the nonhydrogen atoms in the trisaccharide molecule are listed in Tables 6 and 7. The mean standard deviations corresponding to the carbon–carbon and carbon–oxygen bond distances, and the carbon and oxygen valence bond angles are 0·007 Å, 0·006 Å, and 0·4°.

In general, all the bond lengths involving the carbon and oxygen atoms in the trisaccharide are in good agreement with the values reported by Arnott & Scott (1972), and Sundaralingam (1968), for bonds present in mono- and disaccharide carbohydrate crystal structures. The average of the twelve ring carbon–carbon bonds in the molecule is 1·525 Å, with a range of 1·515 to 1·543 Å. Both the longest and the shortest carbon–carbon bonds in the molecule are in ring II, the reducing xylose ring, and are associated with atom C(2′). On an individual basis, rings I, II, and III have means of 1·525, 1·523, and 1·526 Å, with ranges of 0·006, 0·028, and 0·020 Å, respectively, in their carbon–carbon bonds. The one exterior carbon–carbon bond in the molecule, C(5″)–C(6″), linking the carboxyl group to the glucose ring, has a length of 1·528 Å and is slightly longer than what one would expect for a carbon–carbon bond with $sp^3$–$sp^2$ bond character (Dewar & Schmeising, 1960).

The equatorial anomeric carbon–oxygen bonds, C(1)–O(1) and C(1′)–O(1′), in the xylose rings, have

Table 6. *Bond distances involving nonhydrogen atoms in the trisaccharide molecule**

| Xylose (I) | | Xylose (II) | | 4-*O*-Methyl-D-glucuronic acid (III) | |
|---|---|---|---|---|---|
| C(1)–C(2) | 1·525 (7) Å | C(1′)–C(2′) | 1·515 (7) Å | C(1″)–C(2″) | 1·517 (7) Å |
| C(2)–C(3) | 1·521 (7) | C(2′)–C(3′) | 1·543 (8) | C(2″)–C(3″) | 1·528 (7) |
| C(3)–C(4) | 1·527 (7) | C(3′)–C(4′) | 1·518 (7) | C(3″)–C(4″) | 1·524 (7) |
| C(4)–C(5) | 1·527 (8) | C(4′)–C(5′) | 1·519 (8) | C(4″)–C(5″) | 1·537 (7) |
| | | | | C(5″)–C(6″) | 1·528 (7) |
| C(1)–O(5) | 1·442 (6) | C(1′)–O(5′) | 1·409 (7) | C(1″)–O(5″) | 1·422 (6) |
| C(5)–O(5) | 1·440 (7) | C(5′)–O(5′) | 1·431 (7) | C(5″)–O(5″) | 1·430 (6) |
| C(1)–O(1) | 1·393 (6) | C(1′)–O(1′) | 1·397 (7) | C(1″)–O(2) | 1·410 (6) |
| C(2)–O(2) | 1·431 (6) | C(2′)–O(2′) | 1·417 (6) | C(2″)–O(2″) | 1·426 (6) |
| C(3)–O(3) | 1·434 (6) | C(3′)–O(3′) | 1·436 (6) | C(3″)–O(3″) | 1·439 (6) |
| C(4)–O(4) | 1·425 (6) | C(4′)–O(1) | 1·448 (6) | C(4″)–O(4″) | 1·429 (6) |
| | | | | O(4″)–C(7″) | 1·421 (7) |
| | | | | C(6″)–O(6″) | 1·202 (6) |
| | | | | C(6″)–O(7″) | 1·331 (6) |

* The estimated standard deviations are in parentheses and refer to the last decimal position of the respective value.

Table 7. *Bond angles involving nonhydrogen atoms in the trisaccharide molecule**

| Xylose (I) | | Xylose (II) | | 4-*O*-Methyl-D-glucuronic acid (III) | |
|---|---|---|---|---|---|
| C(1)–C(2)–C(3) | 109·0° | C(1′)–C(2′)–C(3′) | 112·0° | C(1″)–C(2″)–C(3″) | 109·3° |
| C(2)–C(3)–C(4) | 111·8 | C(2′)–C(3′)–C(4′) | 112·4 | C(2″)–C(3″)–C(4″) | 108·7 |
| C(3)–C(4)–C(5) | 110·8 | C(3′)–C(4′)–C(5′) | 109·9 | C(3″)–C(4″)–C(5″) | 110·6 |
| C(4)–C(5)–O(5) | 110·1 | C(4′)–C(5′)–O(5′) | 108·0 | C(4″)–C(5″)–O(5″) | 110·3 |
| C(5)–O(5)–C(1) | 110·4 | C(5′)–O(5′)–C(1′) | 110·0 | C(5″)–O(5″)–C(1″) | 114·3 |
| O(5)–C(1)–C(2) | 109·1 | O(5′)–C(1′)–C(2′) | 110·7 | O(5″)–C(1″)–C(2″) | 110·2 |
| O(5)–C(1)–O(1) | 108·2 | O(5′)–C(1′)–O(1′) | 107·9 | O(5″)–C(1″)–O(2) | 111·0 |
| C(2)–C(1)–O(1) | 108·6 | C(2′)–C(1′)–O(1′) | 107·8 | C(2″)–C(1″)–O(2) | 107·9 |
| C(1)–C(2)–O(2) | 109·9 | C(1′)–C(2′)–O(2′) | 108·0 | C(1″)–C(2″)–O(2″) | 109·8 |
| C(3)–C(2)–O(2) | 108·2 | C(3′)–C(2′)–O(2′) | 110·8 | C(3″)–C(2″)–O(2″) | 109·9 |
| C(2)–C(3)–O(3) | 108·6 | C(2′)–C(3′)–O(3′) | 110·2 | C(2″)–C(3″)–O(3″) | 107·0 |
| C(4)–C(3)–O(3) | 111·4 | C(4′)–C(3′)–O(3′) | 110·7 | C(4″)–C(3″)–O(3″) | 110·9 |
| C(3)–C(4)–O(4) | 112·6 | C(3′)–C(4′)–O(1) | 105·0 | C(3″)–C(4″)–O(4″) | 107·8 |
| C(5)–C(4)–O(4) | 105·0 | C(5′)–C(4′)–O(1) | 111·6 | C(5″)–C(4″)–O(4″) | 110·2 |
| | | | | C(4″)–C(5″)–C(6″) | 111·4 |
| | | | | O(5″)–C(5″)–C(6″) | 104·9 |
| | | | | C(4″)–O(4″)–C(7″) | 113·6 |
| | | | | C(5″)–C(6″)–O(6″) | 122·7 |
| | | | | C(5″)–C(6″)–O(7″) | 112·1 |
| | | | | O(6″)–C(6″)–O(7″) | 125·2 |

| | |
|---|---|
| β-1,4 Glycosidic oxygen angle [C(1)–O(1)–C(4′)] | 113·8° |
| α-1,2 Glycosidic oxygen angle [C(1″)–O(2)–C(2)] | 116·0° |

* The estimated standard deviation for all the bond angles was 0·4°.

values of 1·393 and 1·397 Å. Both these values are in good agreement with the average value of 1·389 Å reported by Arnott & Scott (1972) for sugars containing the equatorial anomeric carbon–oxygen bond. The intraring carbon–oxygen bonds in rings I and II, C(1)–O(5), C(5)–O(5), C(1′)–O(5′), and C(5′)–O(5′), have values of 1·442, 1·440, 1·409, and 1·431 Å, respectively. The similarity of length for the C(1)–O(5) and C(5)–O(5) bonds compares well with that observed in other sugars (Ham & Williams, 1970; Sundaralingam, 1968) containing the equatorial C(1)–O(1) bond. The C(1′)–O(5′) and C(5′)–O(5′) bond lengths, in contrast, differ by 0·022 Å, when they were expected to be similar. The greater variation of the carbon–oxygen and carbon–carbon bond lengths, and the valence bond and conformational angles in ring II (in contrast to those in rings I and III), suggests that this ring might be in a slightly distorted chair conformation.

The carbon–oxygen bonds of C(1″)–O(2), C(1″)–O(5″), and C(5″)–O(5″), in ring III have lengths of 1·410, 1·422, and 1·430 Å. These bond lengths are in good agreement with the lengths of similar bonds (1·415, 1·414, and 1·439 Å) reported previously (Arnott & Scott, 1972).

The carbon–oxygen bond lengths in the carboxyl group have values of 1·331 and 1·202 Å, the shorter length corresponding to the double-bonded oxygen atom. These bond lengths compare reasonably well with the values of 1·321 and 1·206 Å for the same bonds in the structure of acetic acid at low temperature (Jönsson, 1971).

The six carbon–hydroxyl oxygen bonds in the molecule, excluding the anomeric C(1′)–O(1′) bond in ring II, have a range of 1·417 to 1·439 Å and a mean of 1·429 Å. Interestingly, the C(4′)–O(1) bond linking the aldobiouronic acid disaccharide to the reducing xylose ring, is the longest carbon–oxygen bond in the molecule. It is noted that the C(4′)–O(1) bond has also been observed to be consistently longer in the *β*-1,4-linked disaccharides of methyl *β*-cellobioside (Ham & Williams, 1970), *β*-cellobiose (Chu & Jeffrey, 1968), and *α*-lactose (Fries, Rao & Sundaralingam, 1971). With the exception of the methyl *β*-cellobioside structure, the C(4′)–O(1) bond is the longest exocyclic carbon–oxygen bond in each of these molecules. In methyl *β*-cellobioside, the C(4′)–O(1) bond is only 0·002 Å shorter than the longest exocyclic bond, C(2′)–O(2′). Since the C(4″)–O(4″) bond in the methyl substituted C(4″)–O(4″)–C(7″) glycosidic link is 0·019 Å shorter than that of the C(4′)–O(1) bond between the two xylose rings, it may be that this bond length is dependent on steric factors which are influenced by the size of the substituent attached at the glycosidic oxygen atom. In contrast, the C(2)–O(2) bond connecting the two rings in the aldobiouronic acid disaccharide through an *α*-1,2 configuration has a value of 1·431 Å and is quite similar to the average value of 1·429 Å for the unsubstituted hydroxyl bonds in the molecule. Thus, it may be that steric differences and the configuration of the atoms in the glycosidic link influence the length of the bonds about a glycosidic oxygen atom.

The carbon–hydrogen and oxygen–hydrogen bond distances in the trisaccharide and the oxygen–hydrogen bond distances in the three water molecules are listed in Table 4. The average value of the twenty carbon–hydrogen bonds is 1·00 Å with a range of 0·82 to 1·10 Å. The average value of the fourteen oxygen–hydrogen bonds in the range of 0·81 to 1·00 Å is 0·88 Å. These values agree with values reported by other workers (Ham & Williams, 1970; Stewart, Davidson & Simpson 1965) for similar bonds in carbohydrate crystal structures.

### *Bond angles*

The interior bond angles in rings I, II, and III, excluding those at the ring oxygen atoms, have mean values of 110·1, 110·6, and 109·8°, with spreads of 2·8, 4·4, and 1·9°, respectively. The interior ring angles involving the ring oxygen atoms in the two xylose rings have values of 110·4° (ring I) and 110·0° (ring II). The largest interior ring angle, 114·3°, is associated with the ring oxygen atom in the glucuronic acid ring. The interior bond angles, in general, agree quite well with the values for similar angles reported by Arnott & Scott (1972).

The exterior bond angles in the trisaccharide molecule show a wide range, 105·0 to 112·6°. The exterior bond angles associated with a ring carbon atom are generally unequal, with the greatest variation occurring about the C(4) atom of the nonreducing xylose ring. In agreement with the observations of Sundaralingam (1968), the exterior angle between the ring oxygen atom and the glycosidic oxygen atom, seems to depend on the configuration of the anomeric carbon–oxygen bond. In the two xylose units containing the equatorial anomeric bond, the O(5)–C(1)–O(1) and O(5′)–C(1′)–O(1′) angles are 108·2 and 107·9°. In the *α*-D-glucuronic acid ring, the O(5″)–C(1″)–O(2) angle is 111·0°.

The bridge oxygen angles corresponding to the glycosidic bonds, C(1)–O(1)–C(4′) and C(1″)–O(2)–C(2), have values of 113·8 and 116·0°. Interestingly, the *β*-1,4-link between the two xylose units has an angle which is very close to the values of 113·2, 113·4, and 113·1° found in the methyl *β*-D-pyranosyl residues of methyl *β*-maltoside (Chu & Jeffrey, 1967), methyl *β*-D-xyloside (Brown, Cox & Llewellyn, 1966), and methyl *β*-D-cellobioside (Ham & Williams, 1970). The difference between this angle and the average of the three angles in the three previously mentioned disaccharides containing the *β*-1,4-link is 2·6°. It is suspected that the difference is probably related to the absence of the intramolecular hydrogen bond in the xylobiose residue. An increase in flexibility between rings due to the absence of the hydrogen bond apparently permits the reducing xylose unit to seek a conformation that diminishes the bridge oxygen angle. The values of 113·0, 113·5, and 113·2° for the glycosidic oxygen angles in methyl *α*-D-glucoside (Berman & Kim, 1968), methyl α-D-galactoside (Gatehouse & Poppleton, 1971*a*), and methyl α-D-altropyranoside (Gatehouse & Poppleton, 1971*b*) are about 3 to 4° smaller than the bridge oxygen angles of 117·6 and 117·2° for methyl β-D-maltoside (Chu & Jeffrey, 1967) and maltose (Quigley, Sarko & Marchessault, 1970). Both of these disaccharide structures contain an intramolecular hydrogen bond. The α-1,2-glycosidic oxygen angle between the glucuronic acid ring and the nonreducing xylose ring in the trisaccharide is 116° (standard deviation, 0·4°). The glycosidic oxygen angle between the C(7″) methyl carbon atom and the C(4″) ring carbon atom in the glucuronic acid ring, 113·6°, agrees quite well with the values reported for the methyl-substituted monosaccharides.

### *Conformational analysis*

All three rings in the trisaccharide molecule have the *C*l chair conformation with all the substituents, except C(1″)–O(2), in an equatorial position. Rings I, II, and III have the mean values of 57·6, 56·0, and 57·1°, and the corresponding ranges of 49·3 to 66·9°, 42·1 to 70·7°, and 54·1 to 60·2°, respectively, for the intracyclic conformational angles (Table 8). The bond geometry statistics for each of the three rings in the molecule are summarized in Table 9.

Ring III, the glucuronic acid ring, has the least spread in the ring bond angles and the ring conformation angles. The 6·1° variation in the conformation angles is quite close to the 5·9° variation in the idealized pyranose ring (Berman & Kim, 1968) and is consistent with variations in other structures that are heavily substituted at the O(1) position (Berman & Kim, 1968). The extracyclic conformation angles, in contrast, have a rather large variation but this is probably due to the orientation of the methoxyl and carboxyl group substituents and the axial anomeric bond. The methoxyl and carboxyl substituents are oriented with respect to the ring so that there is a minimum of steric interaction between themselves and the adjacent ring atoms, C(3″), C(5″), and O(5″), C(4″), respectively (Fig. 2). The carboxyl group is planar and the acceptor oxygen atom, O(*B*), deviates from the plane by only 0·32 Å. The coplanarity of the four atoms in the carboxyl group, the near-linearity of the three atoms O(7″)H···O(*B*) in the hydrogen bond, and the positioning of the acceptor oxygen atom so that the system O(6″)=C(6″)–O(7″)H···O(*B*) is synplanar, are features that have been observed in many carboxylic acid structures (Donohue, 1960). Although the slight twist of the carboxyl group with respect to the O(5″) and C(4″) atoms would seem to be unfavorable from a steric point of view, the energy barrier between conformations is small and it appears that the hydrogen bond is the determining factor.

Ring II, the reducing xylose ring, is the most irregular of the three rings in the trisaccharide molecule. This ring has the greatest variation in the carbon–carbon and carbon–oxygen bond distances, and also the greatest variation in the ring conformational angles. The nonreducing xylose ring, in contrast, has a very

Table 8. *Intracyclic and extracyclic conformation angles in the trisaccharide molecule*

| | Xylose (I) | | Xylose (II) | | 4-*O*-Methyl-D-glucuronic acid (III) | |
|---|---|---|---|---|---|---|
| Intracyclic conformation angles | C(1)–C(2) | 59·9° | C(1′)–C(2′) | 50·4° | C(1″)–C(2″) | 59·1° |
| | C(2)–C(3) | −51·9 | C(2′)–C(3′) | −42·1 | C(2″)–C(3″) | −57·3 |
| | C(3)–C(4) | 49·3 | C(3′)–C(4′) | 46·9 | C(3″)–C(4″) | 55·1 |
| | C(4)–C(5) | −54·1 | C(4′)–C(5′) | −60·2 | C(4″)–C(5″) | −54·1 |
| | C(5)–O(5) | 63·7 | C(5′)–O(5′) | 70·7 | C(5″)–O(5″) | 57·1 |
| | O(5)–C(1) | −66·9 | O(5′)–C(1′) | −65·7 | O(5″)–C(1″) | −60·2 |
| Extracyclic conformation angles | O(1)–C(1)–C(2)–O(2) | −63·9 | O(1′)–C(1′)–C(2′)–O(2′) | −69·7 | O(2)–C(1″)–C(2″)–O(2″) | 58·5 |
| | O(2)–C(2)–C(3)–O(3) | 65·4 | O(2′)–C(2′)–C(3′)–O(3′) | 73·5 | O(2″)–C(2″)–C(3″)–O(3″) | 62·3 |
| | O(3)–C(3)–C(4)–O(4) | −71·7 | O(3′)–C(3′)–C(4′)–O(1) | −69·3 | O(3″)–C(3″)–C(4″)–O(4″) | −67·0 |
| | | | | | C(3″)–C(4″)–O(4″)–C(7″) | 122·9 |
| | | | | | C(5″)–C(4″)–O(4″)–C(7″) | −116·3 |
| | | | | | C(4″)–C(5″)–C(6″)–O(7″) | −91·5 |
| | | | | | O(5″)–C(5″)–C(6″)–O(7″) | 149·2 |
| | | | | | C(4″)–C(5″)–C(6″)–O(6″) | 86·6 |
| | | | | | O(5″)–C(5″)–C(6″)–O(6″) | −32·7 |

Table 9. *Bond geometry statistics for the three rings in the molecule**

| Ring | $\overline{\text{C–C}}$ (Å) | $\Delta$† | $\overline{\text{C–O}}$ (Å)‡ | $\Delta$ | $\overline{\text{C–C–C}}$ (°) | $\Delta$ | $\overline{\text{C–C–O}}$ (°) | $\Delta$ | $\overline{\text{I.C.A.}}$§ | $\Delta$ |
|---|---|---|---|---|---|---|---|---|---|---|
| I | 1·525 | 0·006 | 1·430 | 0·009 | 110·1 | 2·8 | 109·1 | 7·6 | 57·6 | 17·6 |
| II | 1·523 | 0·028 | 1·434 | 0·031 | 110·6 | 4·4 | 109·0 | 6·6 | 56·0 | 28·6 |
| III | 1·526 | 0·020 | 1·428 | 0·013 | 109·8 | 1·9 | 109·1 | 6·5 | 57·1 | 6·1 |

* The mean estimated standard deviations for the carbon–carbon and carbon–oxygen bonds are 0·007 and 0·006 Å, respectively. The mean estimated standard deviation for all angles is 0·4°.

† $\Delta$ refers to the spread between the highest and lowest values in the average.

‡ These averages exclude the ring C–O bonds and the anomeric C–O bonds.

§ Mean intracyclic conformation angle (°).

narrow spread in the carbon–carbon and carbon–oxygen bond distances and a reasonable variation in the conformational and exterior bond angles.

### *Disaccharide residues*

The two xylose rings that form the xylobiose residue in the aldotriouronic acid molecule are rotated in such a fashion that the distance between O(3′) and O(5), 3·964 Å, is greater than that allowable for formation of an intramolecular hydrogen bond. The conformation angles between the two xylose rings are compared to those of other structures in Table 10.

Table 10. *Conformation angles about the bridge bonds*

| | 1*e*,4*e*-Glycosides* | | | |
|---|---|---|---|---|
| | $\varphi_1$(°) | $\varphi_1'$(°) | $\varphi_2$(°) | $\varphi_2'$(°) |
| Theoretical values for the fully extended conformation | −120·0 | 120·0 | 120·0 | −120·0 |
| Methyl β-cellobioside | −88·9 | 152·0 | 80·3 | −160·7 |
| β-Cellobiose | −77·8 | 167·3 | 106·0 | −127·3 |
| α-Lactose | −92·6 | 146·2 | 94·6 | −143·0 |
| Xylobiose portion of the aldotriouronic acid | −98·1 | 159·8 | 161·5 | −100·4 |
| | 1*a*,2*e*-Glycosides† | | | |
| | $\varphi_3$(°) | $\varphi_3'$(°) | $\varphi_4$(°) | $\varphi_4'$(°) |
| Aldobiouronic acid portion of the aldotriouronic acid | 78·8 | −160·3 | 95·5 | −145·5 |

* $\varphi_1$, $\varphi_1'$, $\varphi_2$, and $\varphi_2'$ are defined according to Sundaralingam (1968). $\varphi_1$[O(5)–C(1)–O(1)–C(4′)]; $\varphi_1'$[C(2)–C(1)–O(1)–C(4′)]; $\varphi_2$[C(1)–C(1)–C(4′)–C(3′)]; $\varphi_2'$[C(1)–O(1)–C(4′)–C(5′)].

† $\varphi_3$ [O(5″)–C(1″)–O(2)–C(2)]; $\varphi_3'$ [C(2″)–C(1″)–O(2)–C(2)]; $\varphi_4$ [C(1″)–O(2)–C(2)–C(1)]: $\varphi_4'$ [C(1″)–O(2)–C(2)–C(3)].

A comparison of the conformational angles of α-lactose (Fries, Rao & Sundaralingam, 1971) and methyl β-cellobioside (Ham & Williams, 1970) with the theoretical angles listed in Table 10 indicates that both these disaccharides have near-symmetrical deviations about their bridge bonds, C(1)–O(1) and O(1)–C(4′), while the deviations from the theoretical angles in cellobiose (Chu & Jeffrey, 1968), in contrast, are asymmetric. The spread in each of the angles, $\varphi_1$, $\varphi_1'$, $\varphi_2$, and $\varphi_2'$, within these three disaccharides ranges from 15 to 34°. Since each of these disaccharides contains an intramolecular hydrogen bond, it appears that there is some flexibility in the conformational requirements for hydrogen bond formation. The major difference in the conformation angles of these three disaccharides and the conformation angles in the xylobiose residue occurs in $\varphi_2$ and $\varphi_2'$. The rotation about the O(1)–C(4′) bond causes a 67·9° increase in the $\varphi_2$ angle, and a 46·8° decrease in the $\varphi_2'$ angle, when the xylobiose angles are compared to the averages of the conformation angles in these three structures. This rotation orients the reducing xylose ring so that an intramolecular hydrogen bond cannot be formed.

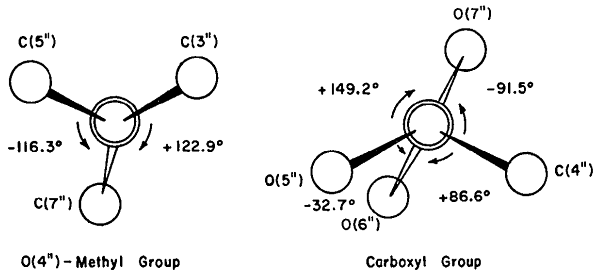


Fig. 2. Conformation angles and orientation of the substituents on the D-glucose ring.

The conformation of the xylobiose residue in the trisaccharide can be compared to the minimum energy conformation of a xylobiose disaccharide by expressing the conformational angles about the C(1)–O(1) and O(1)–C(4′) bonds in the convention followed by Sundararajan & Rao (1969). The conformation of the xylobiose residue, $(\varphi, \psi) = (30°, 35°)$, is such that the difference in potential energy from the minimum energy conformation position, (50°,30°), is less than 0·5 kcal mole$^{-1}$ per residue. The conformation proposed by Settineri & Marchessault (1965) for the threefold helical arrangement of xylose rings containing the intramolecular hydrogen bond, has a value of (63°,25°), and is practically on the same energy contour level as the xylobiose residue, indicating that both conformations experience about the same amount of van der Waals interaction between rings. The absence of the intramolecular hydrogen bond and the orientation of the rings such that the glycosidic oxygen angle is 113·8° in the xylobiose residue, are features that are contrary to the assumptions for the polymer conformation.

The α-1,2-linkage between the glucuronic acid ring and the nonreducing xylose ring has a glycosidic oxygen angle of 116·0°. This value is between the values of 122·2° and 114·6° determined for the α-1,2-linkage in sucrose (Brown & Levy, 1963), and the sucrose residue in raffinose (Berman, 1970). The large difference between the two sucrose bridge angles is in part related to the intramolecular hydrogen bond in the sucrose molecule and to the intermolecular hydrogen bonding of the raffinose molecule with five molecules of water. The two rings in the aldobiouronic acid disaccharide, however, are oriented in such a fashion that an intramolecular hydrogen bond does not form between them. The angle between the two least-squares planes formed by the two six-membered rings in this disaccharide is 122°. The departure of the two rings from a planar conformation which would presumably allow the hydrogen bond to form, causes the O(5″) ring oxygen atom to be pulled away from the O(3) hydroxyl oxygen atom on the nonreducing ring. The nonbonded distance between O(5″) and O(3) is 3·648 Å.

The values of the conformation angles about the α-1,2-link in the aldobiouronic acid disaccharide are listed in Table 10. The $\varphi_3$ and $\varphi_3'$ values, +78·8° and −160·3°, respectively, deviate considerably from the values of +120° and −120° corresponding to the fully extended conformation. In general, since there is little variation in the bond geometry of both rings in this disaccharide, it might be reasonable to assume that this disaccharide lies in a strainless conformation that may not be much different if it were part of the 4-*O*-methylglucuronoxylan polymer.

### *Molecular packing and hydrogen bonding*

The trisaccharide molecule is oriented in the crystal so that there appears to be a planar aldobiouronic acid 'polymer' along the *c* axis and a helical xylan 'polymer' along the *b* axis, with the three water molecules of crystallization clustered between the perpendicular chains and filling in the otherwise open structure (Fig. 3). The arrangement of hydrogen bonds between adjacent molecules in the crystal is illustrated in Fig. 4. [In Fig. 4, (+) and (−) indicate whether the rings are above or below a reference plane at $x=0{\cdot}0$, the arrow indicates the donor–acceptor direction of a hydrogen bond, and the phantom line indicates the nonbonded distances less than 3·3 Å. The $P2_1$ symmetry operations applied to the coordinates of the atoms listed in Table 3 are also listed in this figure.] The oxygen–oxygen and oxygen–hydrogen approach distances with the angles about the hydrogen atoms are given in Table 11.

Table 11. *Hydrogen bonding*

| Donor | | Acceptor | | Distances (Å) | | Angle (°) |
|---|---|---|---|---|---|---|
| O | –H······ | X | O–H | H···X | O···X* | O–H···X |
| O(1′) | H(O1′) | O(3) (*a*) | 0·812 | 1·944 | 2·709 (6) | 156·5 |
| O(2′) | H(O2′) | O(*A*) (*b*) | 0·807 | 2·016 | 2·700 (6) | 142·3 |
| O(3′) | H(O3′) | O(4) (*c*) | 0·833 | 1·988 | 2·814 (5) | 171·1 |
| O(3) | H(O3) | O(2″) (*d*) | 0·891 | 1·919 | 2·690 (5) | 143·8 |
| O(4) | H(O4) | O(3″) (*d*) | 0·905 | 1·850 | 2·723 (6) | 161·4 |
| O(2″) | H(O2″) | O(*B*) (*e*) | 0·869 | 1·872 | 2·737 (5) | 172·9 |
| O(3″) | H(O3″) | O(5) (*f*) | 0·997 | 2·001 | 2·955 (5) | 159·2 |
| O(7″) | H(O7″) | O(*B*) (*d*) | 0·888 | 1·834 | 2·697 (6) | 163·4 |
| O(*A*) | H(O*A*-1) | O(*C*) | 0·839 | 2·015 | 2·777 (6) | 150·7 |
| | H(O*A*-2) | O(3′) | 0·922 | 2·129 | 2·842 (6) | 133·4 |
| O(*B*) | H(O*B*-1) | O(6″) (*g*) | 0·821 | 2·382 | 3·049 (6) | 139·0 |
| | | O(5″) (*g*) | 0·821 | 2·116 | 2·872 (5) | 144·2 |
| | H(O*B*-2) | O(*C*) (*g*) | 0·905 | 1·837 | 2·715 (6) | 162·9 |
| O(*C*) | H(O*C*-1) | O(3′) (*h*) | 0·947 | 1·903 | 2·817 (6) | 161·5 |
| | H(O*C*-2) | O(1′) (*i*) | 0·913 | 1·954 | 2·818 (6) | 157·0 |

Symmetry code

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| (*a*) | $x$ | $y+1{\cdot}0$ | $z$ | (*f*) | $x$ | $y$ | $z+1{\cdot}0$ |
| (*b*) | $-x-1{\cdot}0$ | $y+0{\cdot}5$ | $-z-1{\cdot}0$ | (*g*) | $x+1{\cdot}0$ | $y$ | $z$ |
| (*c*) | $-x$ | $y+0{\cdot}5$ | $-z-1{\cdot}0$ | (*h*) | $-x-1{\cdot}0$ | $y-0{\cdot}5$ | $-z-1{\cdot}0$ |
| (*d*) | $-x$ | $y-0{\cdot}5$ | $-x$ | (*i*) | $x$ | $y-1{\cdot}0$ | $z$ |
| (*e*) | $-x$ | $y+0{\cdot}5$ | $-z$ | | | | |

* The estimated standard deviation in parentheses refers to the last decimal position of the associated value.

The packing arrangement of the molecules in the crystal allows fifteen hydrogen bonds to form, only five of which are between adjacent trisaccharide molecules. The remaining ten involve the three water molecules which play a major part in stabilizing the crystal lattice. The average oxygen–oxygen hydrogen bonding distance corresponding to the fifteen bonds with a range of 2·690 to 3·049 Å, is 2·794 Å. The average angle between the two oxygens and the hydrogen atom is 155·6°, over a range of 133·4 to 172·9°.

The small number of hydrogen bonds between adjacent trisaccharide molecules is primarily a result of the L shape of the molecule and its packing in the crystal. The two hydroxyl groups, O(3″) and O(1′), at the top and toe of the L-shaped molecule, donate their hydrogens to the middle ring of the adjacent molecules and, hence, stabilize the network of 'polymeric' chains (see Fig. 3). Two of the three remaining hydrogen bonds between the trisaccharide molecules originate at O(3) and O(4) of the nonreducing xylose unit. These oxygens donate their hydrogens to O(2″) and O(3″), respectively, of a glucuronic acid ring belonging to a twofold symmetry-related molecule. The packing of the molecules in the crystal is so arranged that these bonds form continuous chains along the twofold screw axis. The last hydrogen bond between trisaccharides is between O(3′) and O(4′) and parallels the *c* axis of the crystal.

A number of general features concerning hydrogen-bonding schemes in carbohydrate crystal structures have been pointed out by several workers (Sundaralingam, 1968; Jeffrey & Rosenstein, 1964), and the features in the aldotriouronic acid trihydrate crystal are consistent with their observations. For example, the glycosidic oxygen atoms, O(1), O(2), and O(4″), do not participate in hydrogen bonding whereas the ring oxygen atoms, O(5) and O(5″) [with the exception of O(5′)] do participate in hydrogen bonding where the hydrogen atom approaches from the equatorial position. In addition, all the hydroxyl oxygens, with the exception of O(2′) and O(3′), act as both donors and acceptors of hydrogen bonds. O(2′) and O(3′), in contrast, participate in one (as donor) and three (one as donor and two as acceptor) hydrogen bonds, respectively. The rather unusual behavior of O(3′) in that it accepts two hydrogen bonds has also been observed in the structures of α-lactose monohydrate (Fries, Rao & Sundaralingam, 1971) and methyl α-galactoside monohydrate (Gatehouse & Poppleton, 1971*a*).

The water molecules in the crystal are positioned so near the hydrophilic oxygen atoms of the adjacent trisaccharide molecules that they participate in two-thirds of the unique hydrogen bonds. O(*B*) and O(*C*) donate both their hydrogen atoms and accept two, whereas O(*A*) donates both its hydrogen atoms and accepts

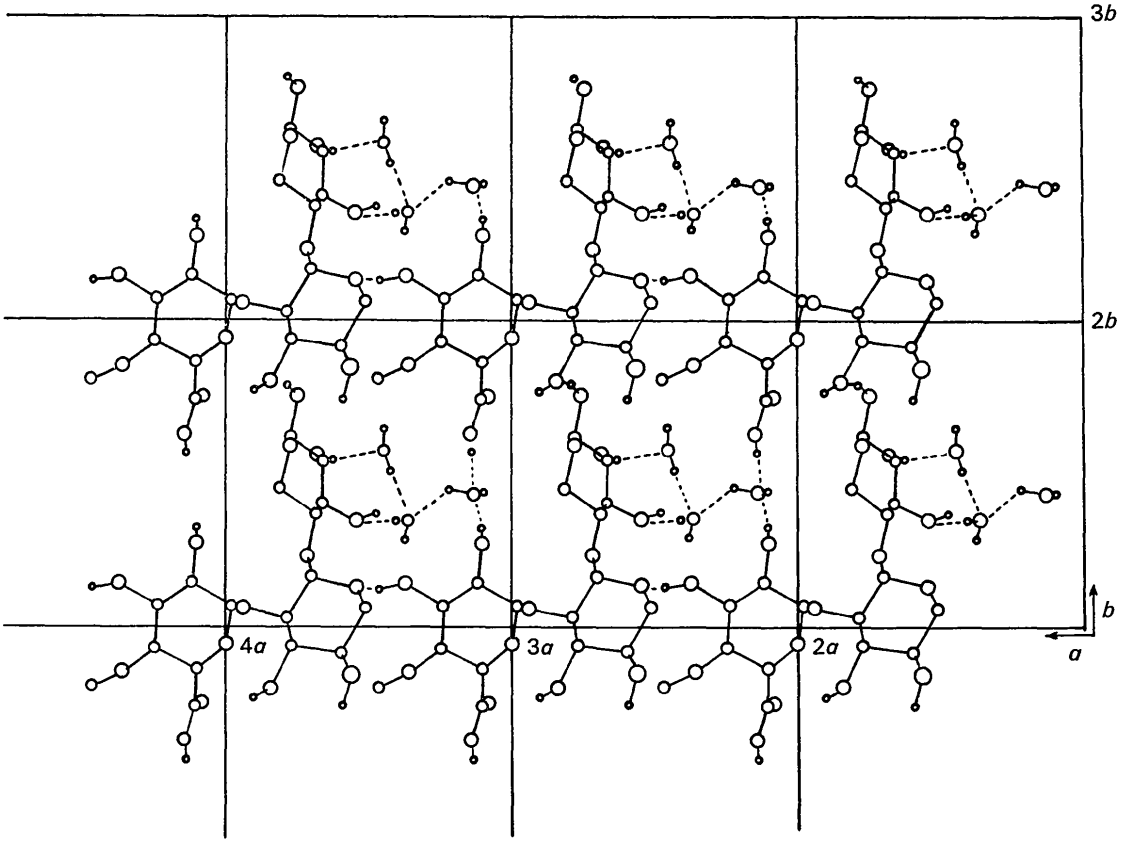


Fig. 3. Multiple unit-cells of the crystal structure in a plane perpendicular to the *a* axis.

only one. One of the hydrogens from O(*B*) appears to be involved in a bifurcated hydrogen bond with O(5″) and O(6″) (see Fig. 4). The hydrogen to oxygen distances of 2·116 and 2·382 Å are both less than the 2·6 Å distance which corresponds to a van der Waals contact (Pauling, 1960), and the interior angles, 144 and 139°, are similar, which suggests that the bonds are nearly equal in strength. Similar bifurcated bond arrangements have been indicated in the structures of methyl $\beta$-cellobioside (Ham & Williams, 1970), methyl $\alpha$-galactoside monohydrate (Gatehouse & Poppleton, 1971*a*), and methyl $\alpha$-altropyranoside (Gatehouse & Poppleton, 1971*b*).

The water molecules and several of the hydroxyl oxygens are involved in two closed spiral systems of hydrogen bonds. The right-handed system resides in a plane parallel to the *bc* plane of the crystal and involves O(2′)-O(*A*)-O(*C*)-O(3′) in the donor-acceptor sequence. The left-handed system of hydrogen bonds joins the right-handed one and involves O(2″)-O(*B*)-O(*C*)-O(3′) in the sequence. O(*B*) also accepts a hydrogen bond from O(7″) so that it links the top and bottom portions of the 4-*O*-methyl-D-glucuronic acid ring of adjacent molecules translated along **b**.

Suggestions given by Dr Dale G. Williams and Dr Norman S. Thompson were appreciated. The authors also thank Professor E. E. Dickey for preparing the samples of the aldotriouronic acid and Dr John A. Heitmann for his many helpful discussions on the crystallographic methods used in this study.

Thanks are also due to Dr Farid R. Ahmed of the National Research Council, Ottawa, Canada, for the NRC series of crystallographic computer programs used throughout this work. Acknowledgement is also extended to Dr Peter Main, University of York, York, England, Dr Charles Weeks of the Medical Foundation of Buffalo, Buffalo, New York, and Dr P. Braun, Philips Reseach Laboratories, Eindhoven, The Netherlands, for their computer programs and explanations on their operation.

## References

AHMED, F. R., HALL, S. R., PIPPY, M. E. & HUBER, C. P. (1966). NRC Crystallographic Programs for the IBM 360 System *World List of Crystallographic Computer Programs*. 2nd ed., Appendix, p. 52.

ARNOTT, S. & SCOTT, W. E. (1972). *J. Chem. Soc. Perkin II*, pp. 324–335.

BERMAN, H. M. (1970). *Acta Cryst.* B**26**, 290–299.

BERMAN, H. M. & KIM, S. H. (1968). *Acta Cryst.* B**24**, 897–904.

BRAUN, P. B., HORNSTRA, J. & LEENHOUTS, J. I. (1969). *Philips Res. Rep.* **24**, 85–118.

BROWN, C. J., COX, G. & LLEWELLYN, F. J. (1966). *J. Chem. Soc.* (*A*), pp. 922–927.

BROWN, G. H. & LEVY, H. A. (1963). *Science*, **141**, 921–923.

CHU, S. C. & JEFFREY, G. A. (1967). *Acta Cryst.* **23**, 1038–1049.

CHU, S. C. & JEFFREY, G. A. (1968). *Acta Cryst.* **24**, 830–838.

DEWAR, M. J. S. & SCHMEISING, H. (1960). *Tetrahedron*, **11**, 96–120.

DICKEY, E. E. (1969). Unpublished work. Appleton, Wisconsin, The Institute of Paper Chemistry.

DONOHUE, J. (1960). *Acta Cryst.* B**24**, 1558–1560.

DUAX, W. L., WEEKS, C. M. & HAUPTMAN, H. (1970). Paper J-3. American Crystallographic Association Symposium. Summer Meeting, 16–22 August, Carleton University.

FRIES, D. C., RAO, S. T. & SUNDARALINGAM, M. (1971). *Acta Cryst.* B**27**, 994–1005.

GATEHOUSE, B. M. & POPPLETON, C. J. (1971*a*). *Acta Cryst.* B**27**, 654–660.

GATEHOUSE, B. M. & POPPLETON, C. J. (1971*b*). *Acta Cryst.* B**27**, 871–876.

HAM, J. T. & WILLIAMS, D. G. (1970). *Acta Cryst.* B**26**, 1373–1383.

HANSON, H. P., HERMAN, F., LEA, J. D. & SKILLMAN, S. (1964). *Acta Cryst.* **17**, 1040–1044.

HAUPTMAN, H., FISHER, J., HANCOCK, H. & NORTON, D. A. (1969). *Acta Cryst.* B**25**, 811–814.

HODGSON, L. I. & ROLLETT, J. S. (1963). *Acta Cryst.* **16**, 329–335.

JEFFREY, G. A. & PARK, Y. J. (1972). *Acta Cryst.* B**28**, 257–267.

JEFFREY, G. A. & ROSENSTEIN, R. D. (1964). *Crystal Structure Analysis in Carbohydrate Chemistry*. In Wolfrom's *Advances in Carbohydrate Research*. Vol. 19, p. 7. New York: Academic Press.

JÖNSSON, P. G. (1971). *Acta Cryst.* B**27**, 893.

KARLE, J. (1968). *Acta Cryst.* B**24**, 182–186.

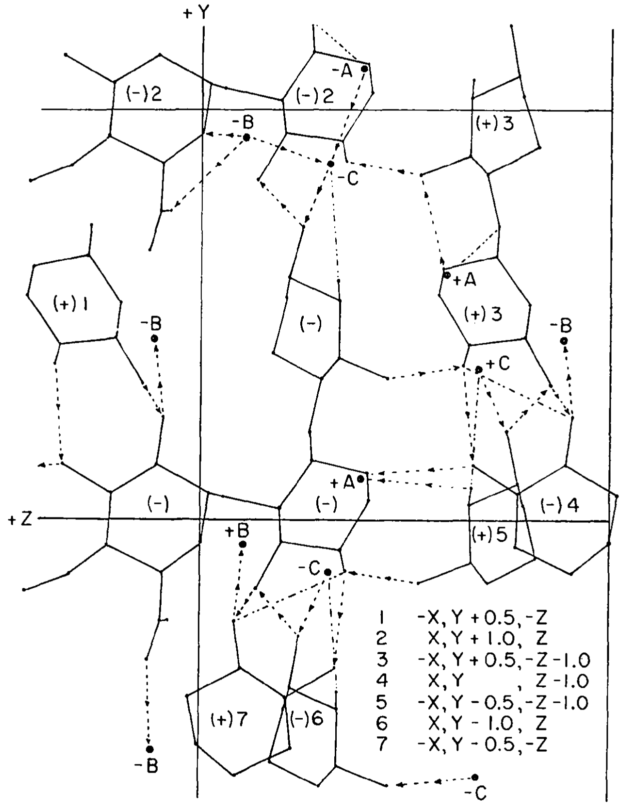


Fig. 4. The molecular packing and hydrogen bonding in the vicinity of a trisaccharide molecule.

KARLE, J. (1970). *Acta Cryst.* **B26**, 1614.
KARLE, J. & HAUPTMAN, H. (1956). *Acta Cryst.* **9**, 635–651.
KILLEAN, R. C. G. (1966). *Acta Cryst.* **20**, 547–549.
NORMENT, H. G. (1962). *An X-ray Diffraction Data Reduction Program for the IBM* 704 *and* 7090. Report 5739. U. S. Naval Research Laboratory, Washington, D. C.
PAULING, L. (1960). *The Nature of the Chemical Bond.* 3rd ed., p. 314. Ithaca: Cornell Univ. Press.
QUIGLEY, G. J., SARKO, A. & MARCHESSAULT, R. H. (1970). *J. Amer. Chem. Soc.* **92**, 5834–5839.
RICHARDS, G. F. (1964). *Some Structures of Metal-Olefin Complexes.* Doctoral Dissertation, Univ. of Iowa. Iowa City, Iowa.
ROHRER, (1972). *Acta Cryst.* **B38**, 425–433.
ROLLETT, J. S. & SPARKS, R. A. (1960). *Acta Cryst.* **13**, 273–274.
SETTINERI, W. J. & MARCHESSAULT, R. H. (1965). *J. Polymer Sci.* (*C*), **11**, 253–264.
STEWART, R. F., DAVIDSON, E. R. & SIMPSON, W. T. (1965). *J. Chem. Phys.* **42**, 3175–3183.
SUNDARALINGAM, M. (1968). *Biopolymers*, **6**, 189–213.
SUNDARARAJAN, P. R. & RAO, V. S. R. (1969). *Biopolymers*, **8**, 305–312.
WEEKS, C. M. (1969). *CDC* 6400 *Computer Programs for the Least-Squares Analysis of Structure Invariants.* Buffalo, New York, The Medical Foundation of Buffalo.

*Acta Cryst.* (1973). B**29**, 2783

# The Crystal and Molecular Structure of Axivalin Hydrate

BY GARY D. ANDERSON, R. S. MCEWEN AND WERNER HERZ

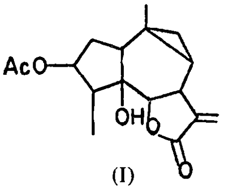
*Department of Chemistry, The Florida State University, Tallahassee, Florida* 32306, *U.S.A.*

(*Received* 8 *December* 1972; *accepted* 9 *July* 1973)

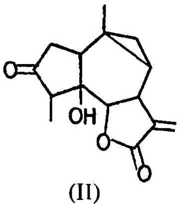
The crystal structure at 25 °C of the hydrate of the cyclopropanoid guaianolide axivalin, $C_{17}O_5H_{22}.H_2O$, has been determined from three-dimensional X-ray diffractometer data with the tangent refinement formula. The crystal is orthorhombic, space group $P2_12_12_1$ with four molecules per unit-cell and the cell dimensions are $a = 8·601$ (3); $b = 10·555$ (4); $c = 18·734$ (9) Å. All hydrogen atoms have been located and the structure refined by least-squares calculations to an $R$ of 0·077. The absolute stereochemistry has been determined from the CD curves of axivalin and its pyrazoline.

## Introduction

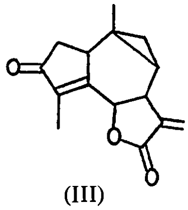
Axivalin (I) is one of a group of novel cyclopropanoid guaianolides isolated from *Iva axillaris* Pursh. ssp *robustior* (Hook.) Bassett (Herz, Sundarsanam & Schmid, 1966). It has been chemically correlated with the other two compounds in this group, ivaxillarin (II) and anhydroivaxillarin (III). Since it was desirable to establish whether the structures of these novel compounds fitted the proposed biogenetic scheme for sesquiterpene lactones and whether the stereochemistry of these compounds was in accord with that of other lactones isolated from the tribe Heliantheae, an X-ray structure analysis was undertaken.



(I)



(II)



(III)

Several efforts to prepare heavy-atom derivatives of axivalin were unsuccessful. These included attempted hydrolysis of the acetate so as to allow preparation of a bromoacetate or bromobenzoate, attempted esterification of the tertiary alcohol group with bromoacetyl bromide, and attempted 1,3 dipolar addition of *o*-bromobenzonitrile oxide to the *exo*-methylene group to form an *o*-bromophenylisoxazole. Hence we decided to attempt the use of direct methods, in particular the tangent refinement formula, for the solution of the crystal structure of axivalin.

## Experimental

Crystallization of axivalin from methanol–water at room temperature yielded crystals of the monohydrate which were colorless, transparent, thin, hexagonal plates elongated in the **b** direction. One such crystal was cleaved with a razor blade to give a rectangular fragment measuring 0·8 by 0·36 by 0·04 mm. The cleavage planes were parallel to the (100) and the (010) planes.

Weissenberg photographs for the $h0l$, $h1l$, and $h2l$ layers and precession photographs for the $0kl$ and $hk0$ layers showed the crystal to be orthorhombic. Systematic absences in the $h00$ ($h = 2n+1$), $0k0$ ($k = 2n+1$), and $00l$ ($l = 2n+1$) reflections uniquely determined the space group as $P2_12_12_1$.