2374

# The Crystal and Molecular Structure of Ergocalciferol (Vitamin D₂)

By S. E. Hull,* I. Leban,† P. Main, P. S. White‡ and M. M. Woolfson

*Department of Physics, University of York, Heslington, York YO1 5DD, England*

The orthorhombic form of ergocalciferol ($C_{28}H_{44}O$) crystallizes in space group $P2_12_12_1$ with $Z=8$, $a=$ 21·695 (10), $b=6·857$ (4), $c=35·320$ (15) Å. Data were collected on a four-circle diffractometer and the structure was solved by the successive application of two novel direct-method techniques. Refinement resulted in an $R$ of 9·5 %. The two molecules in the asymmetric unit differ in the conformations of the rings $A$ and $D$, whilst the conformations of the rings $C$ and the side chains are similar. The steroid molecules are joined by hydrogen bonds (O–H$\cdots$O) which form infinite helices in the direction of **b**.

## Introduction

The fat-soluble D vitamins are important in bone formation and the prevention of rickets. Ergosterol (provitamin D) is prepared from yeast and the other isomeric members of the vitamin $D_2$ series are obtained from it by irradiation with ultraviolet light.

Because of its biological importance, considerable attention has been devoted to clarifying the exact role of vitamin D in chemical and biological processes. The results of investigations in this field have been reviewed (Fieser & Fieser, 1959; Dyke, 1965; Sanders, Pot & Havinga, 1969; Norman & Henry, 1974). Bernal & Crowfoot (1935) summarized the early crystallographic work on compounds in the vitamin D series, and since then the crystal structures of several derivatives have been reported. Amongst these are 3,20-bis(ethylene-dioxy)ergosterol (Braun, Hornstra, Knobler, Rutten & Romers, 1973), calciferyl 4-iodo-3-nitrobenzoate (hereafter INC) (Hodgkin, Rimmer, Dunitz & Trueblood, 1963) and 3,20-bis(ethylenedioxy)ergocalciferol (hereafter ECF) (Knobler, Romers, Braun & Hornstra, 1972). These derivatives were chosen for investigation, either because a heavy atom was needed in the structure so that it could be solved by the crystallographic techniques then available, or because no suitable crystals of the biologically active compound could be obtained. Although the study of these derivatives provided important information, knowledge of the crystal structures of the native compounds is likely to be more useful, since both the molecular conformations and the crystal packing are probably different for the derivatives. This can be illustrated by reference to the two calciferol derivatives (INC and ECF), in which the $A$ rings have markedly different conformations. The crystal structures of lumisterol₃ (de Kok & Romers,

1973) and ergosterol (Hull & Woolfson, 1976) have been reported. This paper deals with ergocalciferol [vitamin $D_2$ or 9,10-secoergosta-5,7,10(19),22-tetraen-3β-ol].

## Experimental

Two crystal forms were isolated when commercial ergocalciferol (Sigma London Chemical Co.) was recrystallized from acetone. The crystal data for the two forms are given in Table 1. The orthorhombic form was used in this work, since the monoclinic form decayed when exposed to X-rays, and the diffraction data extended only up to $\theta=35°$ for Cu $K\alpha$ radiation. About 70 crystals were examined by photographic methods before a well shaped single crystal with the dimensions $0·4 \times 0·4 \times 0·6$ mm was obtained for data collection. The cell parameters were refined by a least-squares fit of the angles $2\theta$, $\chi$ and $\varphi$ for 12 independent reflexions obtained on a four-circle diffractometer. The densities were measured by flotation.

Table 1. *Crystal data for the two forms of ergocalciferol, $C_{28}H_{44}O$, at 20°C*

|  | Orthorhombic | Monoclinic |
|---|---|---|
| Space group | $P2_12_12_1$ (No. 19) | $P2_1$ (No. 4) |
| $a$ | 21·695 (10) Å | 21·079 (15) Å |
| $b$ | 6·857 (4) | 7·267 (6) |
| $c$ | 35·320 (15) | 35·253 (18) |
| $\beta$ |  | 94·1 (2)° |
| Volume | 5254·3 Å³ | 5386·1 Å³ |
| $D_m$ | 1·00 (6) g cm⁻³ | – |
| $Z$ | 8 | 8 |
| $D_x$ | 1·003 g cm⁻³ | 0·978 g cm⁻³ |
| $F(000)$ | 1760 | 1760 |
| Radiation used | Cu $K\alpha$, Ni filter |  |
| $\mu$(Cu $K\alpha$) | 4·4 cm⁻¹ |  |
| Unique reflexions | 3128 |  |
| Observed reflexions | 2060 |  |
| Unobserved reflexions | 1068 |  |

Intensities for two octants of reciprocal space were collected on a Hilger & Watts four-circle computer-controlled diffractometer (Y290/FA128) with Ni-filtered Cu $K\alpha$ radiation to a maximum $\theta$ value of 55° and an $\omega$–$2\theta$ scan. During data collection the reflexions

* Present address: University Chemical Laboratory, Lensfield Road, Cambridge, CB2 1EW, England.

† Present address: Department of Chemistry, University of Ljubljana, Murnikova 6, P.O.B. 537, 61001 Ljubljana, Jugoslavia.

‡ Present address: Department of Chemistry, University of New Brunswick, Fredericton, New Brunswick, Canada.

400 and 0,0,10 were monitored every 15 reflexions and the data were rescaled accordingly. The scan rate was $1 \cdot 2°$ min$^{-1}$ for $2\theta$ and the background was counted for 45 s at each end of the scan. The symmetry-related reflexions were averaged (the mean discrepancy on $F_o^2$ was 5·5% for the 3832 contributing reflexions) to give 3128 unique reflexions, of which 1068 had intensities ($I$) less than $1 \cdot 5\sigma(I)$ and were classified as unobserved. The intensities were reduced to structure-factor moduli in the usual way.* No absorption corrections were applied.

## Structure solution

The data were put on an absolute scale and $|E|$ values calculated by the computer program *NORMAL* [part of the *MULTAN* package of Main, Woolfson, Lessinger, Germain & Declercq (1974)]. Spherically averaged molecular scattering factors were calculated for two fragments of the molecule (one of 9 atoms, the

---

* A list of structure factors has been deposited with the British Library Lending Division as Supplementary Publication No. SUP 31666 (10 pp.). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 13 White Friars, Chester CH1 1NZ, England.
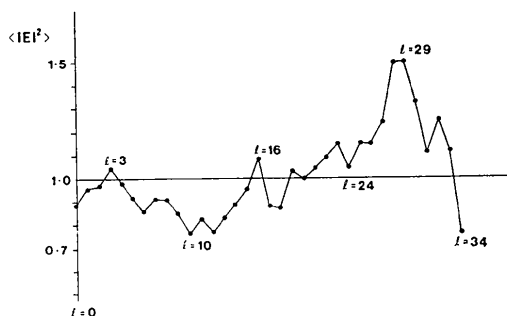


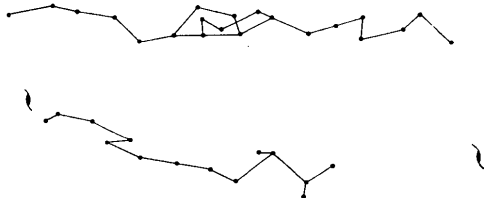Fig. 1. A plot of the average value of $|E|^2$ against the index $l$.



Fig. 2. A projection onto the *ac* plane of the 35 peaks used to recalculate phases and weights for the structure invariants.
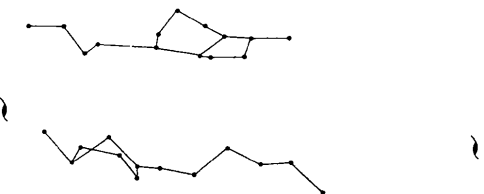


Fig. 3. A projection onto the *ac* plane of the 27 peaks obtained from the fifth $E$ map after the invariant-estimation procedure.

other of 11) with the coordinates for ECF given by Knobler *et al.* (1972): the overall temperature factor was calculated as 6·2 Å$^2$. The normalized data had a peculiar distribution, shown graphically in Fig. 1, of $|E|^2$ with respect to the index $l$. This distribution was probably partly responsible for the rather strange *MULTAN* convergence map which indicated that phases would initially be developed for two distinct sets of reflexions: one with low $l$ index ($l \le 5$) and the other with high $l$ index ($l \ge 23$). All our attempts to solve the structure by the usual *MULTAN* procedure failed.

We then adopted the 'magic integer' approach (White & Woolfson, 1975). Several two-dimensional maps were calculated with the set of six integers 3, 4, 5, 30, 40, 50 and used to provide phases for seven special and nine general reflexions in addition to the three origin-defining reflexions. The phases corresponding to the six highest peaks in these maps were input as known phases into the *FASTAN* section of *MULTAN* and $E$ maps were calculated for all six phase sets. The best of these maps showed only a dubious fragment of one molecule; however when 22 peaks of this fragment were used to calculate phases for re-input to the tangent formula, the resultant $E$ map revealed the 35 peaks which are shown projected onto the *ac* plane in Fig. 2. At this stage no further atoms could be located either by tangent formula recycling or by weighted Fourier synthesis.

The structure was eventually solved by using the molecular fragments, obtained as described above, to recalculate phases and weights for the three-phase structure invariants ($\Sigma_2$'s) according to a method developed by Main (1975). Fig. 2 shows the fragments, one of 21 atoms and one of 14, which we assumed to be correctly orientated but not necessarily correctly positioned, either with respect to the symmetry elements of the space group or with respect to each other. Probable phases and new weights ($\kappa$) were calculated for 8000 invariants; in these calculations the $|E|$ values were not recalculated to take the known structure into account, otherwise the procedure followed that given by Main (1975). After these calculations *MULTAN* was re-entered at the *CONVERGE* stage and phase determination was allowed to proceed in the usual way, except that, since the enantiomorph was now defined by the invariants, it was necessary to develop twice as many phase sets (64 against 32) as would normally have been necessary. $E$ maps were calculated for the phase sets with the largest *MULTAN* combined figures of merit and the fifth of these maps could be interpreted as shown in Fig. 3. At this stage only 27 of the 58 atoms in the asymmetric unit had been found and it was necessary to find more atoms by repeated application of the tangent formula before the structure could be completed by ordinary Fourier methods.

Examination of the invariant estimation process after the structure had been solved showed that the greatest improvement in the phase estimation was for those

relationships used early in the *MULTAN* tangent formula procedure. The r.m.s. weighted (with $\kappa$) error of the structure invariants used in the determination of the phases of the first 100 reflexions was reduced from 70° when no information about the molecular orientation was used, to 64° when the information was used. This is not a large improvement, and is much less than would be expected if the orientation of such a large part of the structure were known exactly. However comparison of Fig. 2 with Fig. 7 shows that the 35 atoms used in the estimation process are positioned so that they give a good idea of the general orientations of the molecules, but that the detailed structure is not very accurately represented. It seems that, in this case,

## Table 2. *Fractional coordinates and thermal parameters*

Fractional coordinates are multiplied by $10^4$. Estimated standard deviations in the least significant digits are given in parentheses. Thermal parameters and their e.s.d.'s are $\times 10$ Å$^2$. The temperature factor is defined as $\exp[-2\pi^2\sum_{ij}h_ih_ja_i^*a_j^*U(ij)]$, $i,j=1,2,3$.

| | x | y | z | U(11) | U(22) | U(33) | U(12) | U(13) | U(23) |
|---|---|---|---|---|---|---|---|---|---|
| O(3) | 5681 (4) | 2465 (12) | 2474 (2) | 11 (1) | 8 (1) | 8 (1) | 0 (1) | 0 (1) | 2 (1) |
| C(1) | 6609 (7) | 4905 (19) | 2877 (3) | 15 (2) | 7 (1) | 8 (1) | 1 (1) | 0 (1) | 1 (1) |
| C(2) | 6649 (7) | 4209 (22) | 2467 (3) | 15 (2) | 11 (1) | 6 (1) | −2 (1) | 0 (1) | 1 (1) |
| C(3) | 6315 (6) | 2283 (18) | 2399 (3) | 10 (1) | 8 (1) | 7 (1) | 4 (1) | −1 (1) | 0 (1) |
| C(4) | 6591 (7) | 738 (19) | 2667 (3) | 15 (1) | 9 (1) | 5 (1) | 1 (1) | 0 (1) | 0 (1) |
| C(5) | 6588 (6) | 1382 (17) | 3084 (3) | 12 (1) | 7 (1) | 5 (1) | 3 (1) | −2 (1) | 0 (1) |
| C(6) | 6403 (6) | 215 (19) | 3353 (3) | 10 (1) | 9 (1) | 4 (1) | 2 (1) | 1 (1) | 2 (1) |
| C(7) | 6349 (5) | 616 (18) | 3756 (3) | 9 (1) | 8 (1) | 6 (1) | 1 (1) | 0 (1) | 1 (1) |
| C(8) | 6294 (6) | −709 (18) | 4035 (5) | 9 (1) | 7 (1) | 8 (1) | 1 (1) | −1 (1) | 1 (1) |
| C(9) | 6361 (6) | −2858 (18) | 3981 (3) | 11 (1) | 7 (1) | 8 (1) | 0 (1) | 0 (1) | −1 (1) |
| C(10) | 6831 (6) | 3320 (19) | 3140 (3) | 11 (1) | 8 (1) | 6 (1) | 3 (1) | 1 (1) | 0 (1) |
| C(11) | 6942 (6) | −3571 (17) | 4187 (3) | 12 (1) | 6 (1) | 6 (1) | 2 (1) | 0 (1) | 0 (1) |
| C(12) | 6906 (6) | −3043 (15) | 4617 (3) | 13 (1) | 4 (1) | 6 (1) | 0 (1) | 1 (1) | 0 (1) |
| C(13) | 6816 (5) | −843 (16) | 4661 (3) | 6 (1) | 7 (1) | 5 (1) | 0 (1) | 0 (1) | 0 (1) |
| C(14) | 6214 (5) | −200 (18) | 4451 (3) | 8 (1) | 9 (1) | 4 (1) | 0 (1) | −1 (1) | 0 (1) |
| C(15) | 6097 (6) | 1836 (18) | 4582 (3) | 12 (1) | 7 (1) | 5 (1) | 3 (1) | 0 (1) | 0 (1) |
| C(16) | 6343 (6) | 1888 (18) | 4998 (3) | 10 (1) | 8 (1) | 7 (1) | −1 (1) | 0 (1) | 0 (1) |
| C(17) | 6630 (5) | −168 (17) | 5076 (3) | 10 (1) | 7 (1) | 5 (1) | 0 (1) | 0 (1) | 0 (1) |
| C(18) | 7381 (5) | 241 (19) | 4530 (3) | 8 (1) | 9 (1) | 7 (1) | 0 (1) | 1 (1) | 1 (1) |
| C(19) | 7270 (7) | 3773 (26) | 3412 (3) | 11 (1) | 14 (2) | 7 (1) | 0 (1) | 1 (1) | 0 (1) |
| C(20) | 7111 (6) | −27 (18) | 5397 (3) | 11 (1) | 7 (1) | 5 (1) | 0 (1) | 0 (1) | 0 (1) |
| C(21) | 7466 (7) | −2011 (22) | 5444 (3) | 11 (1) | 11 (1) | 8 (1) | 2 (1) | −1 (1) | 0 (1) |
| C(22) | 6776 (6) | 475 (22) | 5746 (3) | 9 (1) | 12 (1) | 7 (1) | 2 (1) | −2 (1) | 0 (1) |
| C(23) | 6855 (7) | 2252 (21) | 5920 (3) | 12 (1) | 10 (1) | 6 (1) | 3 (1) | −1 (1) | 0 (1) |
| C(24) | 6512 (7) | 2699 (23) | 6285 (4) | 14 (2) | 10 (2) | 9 (1) | 2 (1) | 0 (1) | −2 (1) |
| C(25) | 6942 (10) | 3278 (35) | 6617 (5) | 18 (2) | 17 (2) | 10 (1) | −2 (2) | −1 (2) | 0 (2) |
| C(26) | 7352 (11) | 1681 (42) | 6716 (5) | 20 (2) | 24 (3) | 10 (1) | 9 (2) | −2 (2) | 2 (2) |
| C(27) | 7289 (12) | 5095 (41) | 6523 (5) | 23 (3) | 20 (3) | 11 (1) | 3 (2) | −3 (2) | −5 (2) |
| C(28) | 6036 (8) | 4306 (30) | 6204 (5) | 14 (2) | 14 (2) | 14 (2) | 1 (1) | 1 (1) | −4 (1) |
| O(103) | 4876 (4) | 437 (14) | 2905 (2) | 11 (1) | 10 (1) | 7 (1) | −1 (1) | 1 (1) | 0 (1) |
| C(101) | 4411 (11) | 3776 (29) | 3681 (5) | 21 (2) | 12 (2) | 11 (1) | 0 (2) | 0 (1) | 0 (1) |
| C(102) | 4505 (9) | 3055 (25) | 3274 (4) | 19 (2) | 10 (1) | 8 (1) | 1 (1) | 0 (1) | 1 (1) |
| C(103) | 4810 (6) | 1178 (19) | 3286 (3) | 11 (1) | 8 (1) | 7 (1) | 2 (1) | 1 (1) | 0 (1) |
| C(104) | 4442 (7) | −284 (19) | 3509 (3) | 16 (2) | 7 (1) | 6 (1) | −1 (1) | 1 (1) | 2 (1) |
| C(105) | 4267 (6) | 433 (19) | 3895 (3) | 9 (1) | 8 (1) | 7 (1) | 0 (1) | 0 (1) | 1 (1) |
| C(106) | 4322 (7) | −783 (19) | 4193 (3) | 13 (1) | 8 (1) | 7 (1) | 0 (1) | 1 (1) | 0 (1) |
| C(107) | 4138 (6) | −301 (18) | 4588 (3) | 11 (1) | 7 (1) | 7 (1) | 0 (1) | 2 (1) | 0 (1) |
| C(108) | 4150 (6) | −1532 (17) | 4876 (4) | 12 (1) | 5 (1) | 9 (1) | −1 (1) | 2 (1) | 1 (1) |
| C(109) | 4370 (9) | −3606 (23) | 4877 (5) | 17 (2) | 9 (1) | 12 (1) | 1 (1) | 3 (1) | 1 (1) |
| C(110) | 4051 (6) | 2461 (21) | 3906 (3) | 9 (1) | 10 (1) | 7 (1) | 1 (1) | 0 (1) | 2 (1) |
| C(111) | 4837 (8) | −3980 (22) | 5173 (4) | 14 (1) | 9 (1) | 12 (1) | 4 (1) | 0 (1) | 2 (1) |
| C(112) | 4634 (7) | −3271 (22) | 5583 (4) | 14 (1) | 9 (1) | 9 (1) | 0 (1) | 2 (1) | 2 (1) |
| C(113) | 4431 (5) | −1159 (16) | 5563 (0) | 6 (1) | 6 (1) | 9 (1) | 0 (1) | 1 (1) | 1 (1) |
| C(114) | 3918 (5) | −948 (18) | 5269 (3) | 9 (1) | 7 (1) | 6 (1) | 0 (1) | 1 (1) | 0 (1) |
| C(115) | 3642 (6) | 1066 (18) | 5332 (3) | 11 (1) | 7 (1) | 7 (1) | 2 (1) | 0 (1) | 0 (1) |
| C(116) | 3649 (6) | 1209 (21) | 5769 (0) | 10 (1) | 11 (1) | 7 (1) | 1 (1) | 2 (1) | −1 (1) |
| C(117) | 4089 (6) | −344 (19) | 5919 (3) | 8 (1) | 8 (1) | 8 (1) | 2 (1) | 1 (1) | 1 (1) |
| C(118) | 4967 (6) | 206 (22) | 5477 (4) | 9 (1) | 11 (1) | 9 (1) | 0 (1) | 1 (1) | 1 (1) |
| C(119) | 3529 (8) | 2986 (37) | 4087 (4) | 13 (2) | 23 (2) | 9 (1) | 6 (2) | 0 (1) | −1 (1) |
| C(120) | 4463 (7) | 356 (27) | 6264 (4) | 12 (1) | 13 (2) | 10 (1) | 0 (1) | 1 (1) | 0 (1) |
| C(121) | 4905 (9) | −1112 (29) | 6435 (4) | 16 (2) | 15 (2) | 10 (1) | 3 (2) | −1 (1) | 2 (1) |
| C(122) | 4020 (7) | 1134 (23) | 6565 (3) | 13 (1) | 10 (1) | 8 (1) | 0 (1) | 0 (1) | 0 (1) |
| C(123) | 4030 (7) | 2970 (25) | 6703 (4) | 12 (1) | 14 (2) | 11 (1) | −1 (2) | 0 (1) | −1 (1) |
| C(124) | 3575 (10) | 3895 (32) | 6986 (5) | 17 (2) | 16 (2) | 12 (1) | 1 (2) | 0 (2) | 0 (2) |
| C(125) | 3871 (20) | 4080 (71) | 7370 (13) | 26 (4) | 26 (4) | 30 (5) | −7 (4) | −5 (4) | 3 (4) |
| C(126) | 3986 (32) | 2131 (98) | 7541 (12) | 45 (7) | 40 (7) | 23 (4) | −3 (7) | −2 (4) | −1 (5) |
| C(127) | 4407 (17) | 5229 (67) | 7391 (9) | 25 (4) | 30 (5) | 20 (3) | 0 (4) | 0 (3) | 1 (4) |
| C(128) | 3262 (12) | 5636 (30) | 6840 (8) | 19 (2) | 10 (2) | 24 (3) | 2 (2) | 6 (2) | 1 (2) |

the small improvement in invariant estimation achieved by using the available knowledge of the approximate orientations of the molecules was sufficient to enable us to solve the structure.

## Refinement

Refinement, with one scale factor and isotropic vibrational parameters for the C atoms, excluding four terminal atoms in the side chain of molecule II and the H atoms, gave an $R$ of 16·0%. Since the positions of the atoms in the side chains were somewhat doubtful because of the large thermal motion, they were omitted from the next structure-factor calculation. At this stage a difference map revealed all the atoms of the side chains without any spurious peaks and thus the possibility of disorder could be ruled out.

Three cycles of isotropic full-matrix least-squares refinement resulted in an $R$ of 12·5%. At this stage, positions were calculated for the H atoms, assuming a C–H length of 1·08 Å and that the methyl group H atoms would adopt the staggered conformation. During the final stage of the refinement all 58 heavy atoms were refined with anisotropic temperature factors, whereas the 88 H atoms were kept fixed at their calculated positions with isotropic $B$ values 2 Å$^2$ larger than those of the atoms to which they are bonded. The final $R$ and $R_w$ for the 2060 observed reflexions were 9·5 and 12·5% respectively.

Scattering factors for C and O were taken from Cromer & Mann (1968) and for H from Stewart, Davidson & Simpson (1965). The function minimized was $w(|F_o| - |F_c|)^2$ and the following weighting scheme was determined empirically: $w = w_F \times w_S$, where $w_F(|F_o| < 10·0) = |F_o|/10·0$, $w_F(|F_o| > 24·0) = (24·0/|F_o|)^{3/2}$, $w_F(10·0 < |F_o| < 24·0) = 1·0$; and $w_S(\sin \theta < 0·45) = (\sin \theta/0·45)^3$, $w_S(\sin \theta > 0·65) = (0·65/\sin \theta)^3$, $w_S(0·45 < \sin \theta < 0·65) = 1·0$.

The final positional and vibrational parameters of the heavy atoms are listed in Table 2. The atoms are numbered according to chemical convention. The same numbering scheme is used for each of the two molecules in the asymmetric unit, except that the numbers used for (II) are obtained by adding 100 to those for (I) (Fig. 4). Because of the limited resolution of the data along **b** the standard deviations of the $y$ coordinates are several times larger than the corresponding deviations in $x$ and $z$. The average estimated standard deviations in the bond lengths are 0·017 for bonds in the steroid nuclei and 0·032 Å for bonds in the side chains. The average estimated standard deviations for the valency angles are 1·0 and 1·8° in the steroid nuclei and the side chains respectively. The estimated deviation for the torsion angles is 2·0°. The estimated standard deviations are rather high mainly because of the unfavourable ratio between the number of parameters and of observed reflexions.

Table 2 shows that several atoms, especially those in the side chains, have rather high thermal parameters. This is not unexpected since the largest motions are usually found in tail or side-group atoms and not in the steroid nucleus. Despite the large thermal vibrations, the distribution of bond lengths and angles does not suggest that the structure is disordered.

Neither the positional nor the vibrational parameters of the H atoms could be refined and these atoms will not be discussed further.

### Computational details

The structure solution was carried out on the Leeds University ICL 1906A computer accessed through the University of York Computing Service remote terminal. The refinement and final interpretation of the results were performed on a CDC Cyber 72 computer at the RRC Ljubljana with the X-RAY 72 system (Stewart, Kruger, Ammon, Dickinson & Hall, 1972) adapted for this computer by Professor L. Golič.

## Discussion of the structure

### Molecular geometry

The bond lengths, valency and torsion angles for the two independent molecules (I) and (II) are summarized in Fig. 4. Some additional geometrical data not shown in Fig. 4 are presented in Table 3. Table 4 gives the weighted average values for the different types of bond length with average standard deviations $\sigma(\bar{x})$ calculated as $\bar{x} = \sum_i (x_i/\sigma_i^2)/(\sum_i \sigma_i^{-2})$ and $\sigma(\bar{x}) = (1/\sum_i \sigma_i^{-2})^{1/2}$. The average bond lengths are in good agreement with the commonly observed values and also with those found in INC and ECF.

Table 3. *Additional geometrical data not included in Fig. 4*

| Bond angles (°) | (I) | (II) |
|---|---|---|
| 12–13–17 | 114 | 116 |
| 14–13–18 | 113 | 111 |

| Intramolecular distances (Å) | | |
|---|---|---|
| 6···9 | 3·06 (2) | 3·10 (2) |
| 7···15 | 3·09 (1) | 2·99 (2) |
| 7···19 | 3·19 (2) | 3·15 (2) |
| 15···19 | 5·03 (2) | 4·60 (2) |

| Hydrogen bonds (Å) | | |
|---|---|---|
| O(3)–O(103) | | 2·70 (1) |
| O(3)–O(103) | | 2·72 (1) |
| at $(1-x,\frac{1}{2}+y,\frac{1}{2}-y)$ | | |

Because there are two molecules in the asymmetric unit, a check on the accuracy of the structure determination is possible by comparing the chemically equivalent bond lengths and angles by a one-parameter significance test (Cruickshank & Robertson, 1953). For bond lengths the mean value of $|l_I - l_{II}|/(\sigma_I^2 + \sigma_{II}^2)^{1/2}$ is 1·3; for valency angles the corresponding quantity is 1·6. This shows that the internal consistency is fairly good in spite of the moderate accuracy of the structure determination.

Although the geometry of the two molecules does not vary a great deal, significant differences are found in the conformations of the rings A and D, and to a smaller extent in the side chain. Assuming the same absolute configuration for ERC as found in INC and ECF, Fig. 5 and comparison of the torsion angles in the two molecules reveal that although the rings A in both molecules are somewhat distorted, flattened chairs with mean torsion angles of 51 and 52°, the angles have opposite signs in the two molecules. Moreover, in (I) the hydroxyl group is in an axial position ($\alpha$ chair), whereas in (II) it is equatorial ($\beta$ chair). A similar situation has also been observed recently, by NMR methods, for ERC in non-polar solution (La Mar & Budd, 1974; Wing, Okamura, Rego, Pirio & Norman, 1975). It was found that ERC exists as a mixture containing approximately equal amounts of the vitamin with the ring A in the $\alpha$ and $\beta$ chair conformations in dynamical equilibrium. From these results, we can deduce that the two derivatives INC and ECF, on crystallization, assume the most stable conformation, which is the $\alpha$ chair for INC and the $\beta$ chair for ECF.

Because the two molecules of ERC have different conformations in the crystalline state, there are differences in the planarity of the conjugated regions of the molecules. A comparison between selected least-squares planes in the two molecules is given in Table 5. In particular it can be seen that whilst in (II) the diene system (atoms 4, 5, 6, 7, 8, 9, 10 and 14) is planar, in (I) atoms 9 and 10 are 0·3 Å from the mean plane. The separate ene systems (atoms 6,7,8,9,14 and 4,5,6,7,10) are approximately planar in both molecules; however the dihedral angles between the planes are 25 and 4°, in (I) and (II) respectively, compared to the values 9 and 12·5° observed for INC and ECF respectively. The torsion angle about the bond 6–7 is − 164 in (I) and 175° in (II). The NMR results of Wing et al. (1975) gave a value for this angle not detectably different from that observed in INC (168°). The distortion of the rings A (Fig. 6) can be attributed mainly to the effects of molecular packing.

In both molecules ring C is a regular chair: the mean torsion angles are 60 and 56° in (I) and (II) respectively. Rings C and D are trans fused in the usual way.

Another conformational difference between the two molecules is that between the rings D which can be best expressed in terms of the 'phase angle' of pseudorotation $\Delta$ and the maximum puckering angle $\varphi_m$ (Al-

tona, Geise & Romers, 1968). The values are $\Delta = 27$ and $\varphi_m = 49°$ for (I) and $\Delta = -12$ and $\varphi_m = 45°$ for (II). This indicates that the conformation of ring D in (I) is
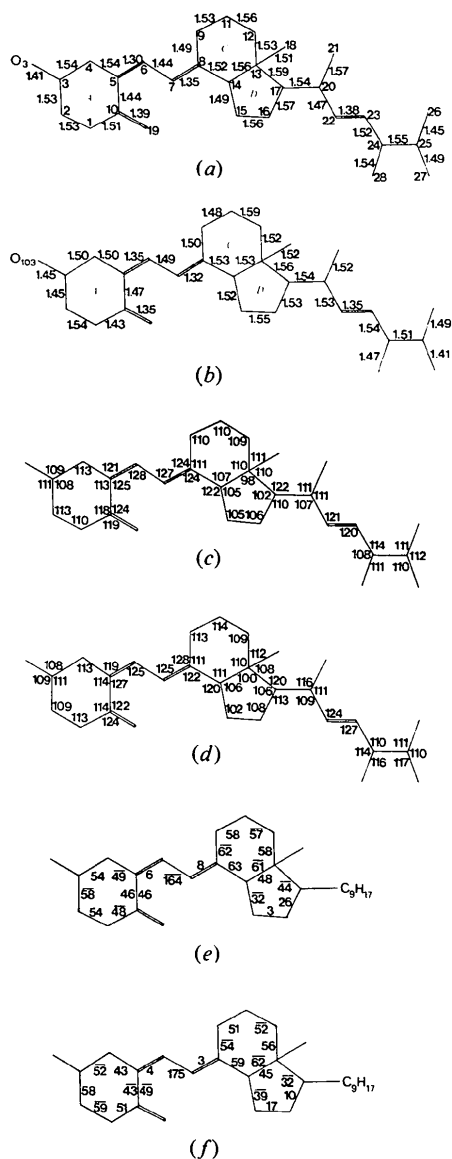


Fig. 4. (a), (b) Bond lengths and numbering scheme, (c) and (d) valency angles, and (e) and (f) torsion angles of ergocalciferol.

Table 4. The average values of bond lengths (Å) of different types

|  | Accepted value | Average value | Number | Range | Average deviation |
|---|---|---|---|---|---|
| C($sp^3$)–C($sp^3$) | 1·533[a] | 1·537 | 36 | 1·406–1·592 | 0·004 |
| C($sp^3$)–C($sp^2$) | 1·505[b] | 1·509 | 12 | 1·433–1·542 | 0·006 |
| C($sp^2$)–C($sp^2$) | – | 1·463 | 4 | 1·444–1·486 | 0·009 |
| C($sp^2$)=C($sp^2$) | 1·337[c] | 1·345 | 8 | 1·303–1·388 | 0·010 |
| C($sp^3$)–OH | 1·426[c] | 1·429 | 2 | 1·406–1·447 | 0·020 |

(a) Bartell (1959), (b) Bartell & Bonham (1960), (c) Sutton (1965).

intermediate between a C(13) envelope and a C(16) half-chair, whilst in (II) it is intermediate between a C(14) envelope and a C(16) half-chair.

The dihedral angles characterizing the $(17\beta)$ side chains are listed in Table 6. It can be seen that the conformations are similar in both molecules. The side chain curls at C(24) so that C(25) is a long way from the plane through the atoms 20, 22, 23 and 24. The same side-chain conformation was found in INC and suprasteryl II 4-iodo-5-nitrobenzoate (Saunderson, 1965). The side-chain conformation in ERC and other vitamin $D_2$ derivatives is considerably different from that in lumisterol₃ (de Kok & Romers, 1974), in which the chain consists of only eight atoms and is saturated.
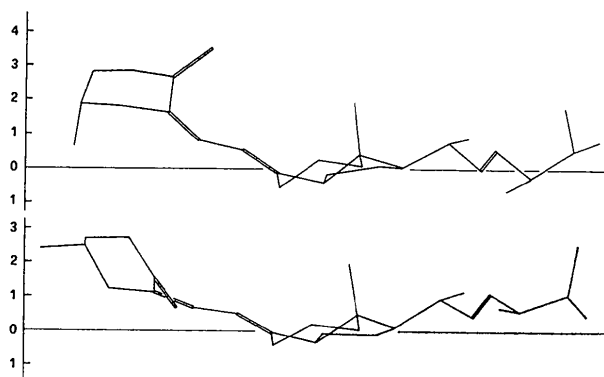


Fig. 5. Projection of the molecules I and II parallel to the least-squares plane through atoms of rings $C$ and $D$. The scale is in Å.
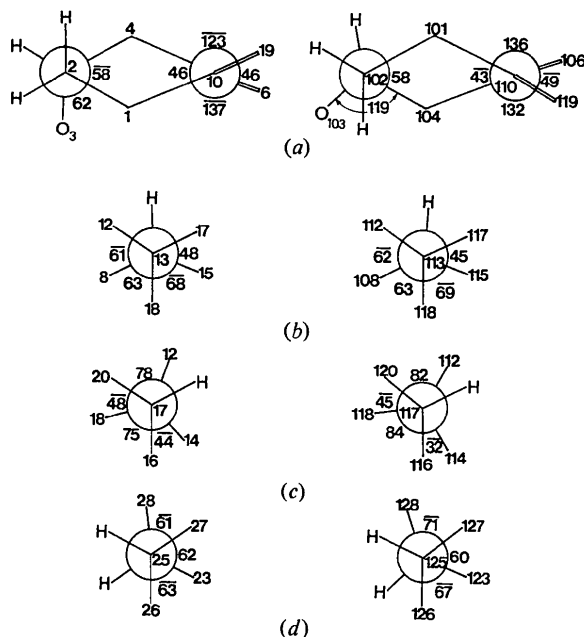
Table 5. *Displacements (Å) of atoms from least-squares planes and dihedral angles (°) in ergocalciferol*

Plane (1): 4,5,6,7,8,9,10,14

|  | (I) | (II) |
| --- | --- | --- |
| C(4) | −0·11 | −0·06 |
| C(5) | −0·02 | 0·00 |
| C(6) | −0·18 | 0·05 |
| C(7) | −0·18 | 0·05 |
| C(8) | −0·04 | 0·01 |
| C(9) | 0·31 | 0·01 |
| C(10) | 0·31 | 0·01 |
| C(14) | −0·10 | −0·06 |
| C(19) | 1·31 | −0·81 |

Plane (2): 6,7,8,9,14

|  | (I) | (II) |
| --- | --- | --- |
| C(6) | −0·02 | −0·02 |
| C(7) | 0·04 | 0·03 |
| C(8) | −0·02 | 0·01 |
| C(9) | 0·01 | 0·00 |
| C(14) | −0·01 | −0·02 |

Plane (3): 4,5,6,7,10

|  | (I) | (II) |
| --- | --- | --- |
| C(4) | −0·01 | −0·01 |
| C(5) | −0·01 | 0·00 |
| C(6) | 0·03 | 0·03 |
| C(7) | −0·02 | −0·02 |
| C(10) | 0·01 | 0·00 |

Plane (4): 1,5,10,19

|  | (I) | (II) |
| --- | --- | --- |
| C(1) | 0·00 | −0·01 |
| C(5) | 0·00 | −0·01 |
| C(10) | −0·01 | 0·03 |
| C(19) | 0·00 | −0·01 |

Dihedral angles

|  | (I) | (II) |
| --- | --- | --- |
| Plane (2)–Plane (3) | 25 | 4 |
| Plane (3)–Plane (4) | 47 | 45 |



Fig. 6. Newman projections for molecules I and II along the bonds (a) 2–3 and 10–5, (b) 13–14, (c) 17–13, and (d) 25–24.

Although the differences may be partly due to molecular packing, it seems quite likely that these changes in the $(17\beta)$ side chain may be responsible for the fact that vitamin $D_3$ (cholecalciferol) is much more active in humans than vitamin $D_2$ (ergocalciferol) (Dyke, 1965). The melting point of vitamin $D_2$ is 115–118°, that of vitamin $D_3$ 84–85°, and that of vitamin $D_4$ (dihydroergocalciferol), which has intermediate activity, 96–98°.

Table 6. *Dihedral angles (°) defining the side chains*

Torsional and dihedral angles are taken positive if the bond in front has to be turned clockwise in order to eclipse the rear bond.

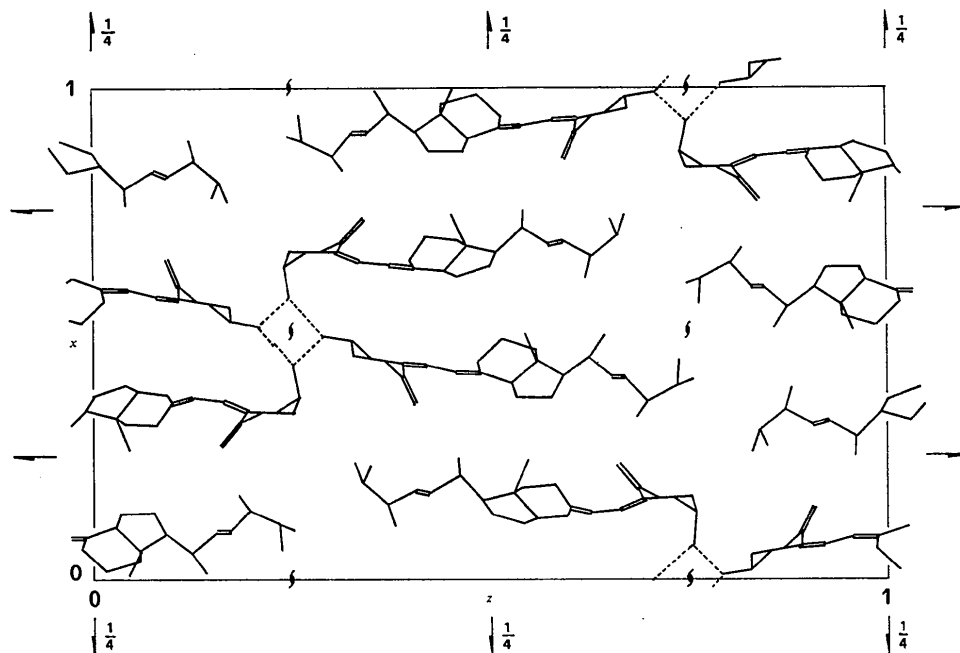| Atoms | (I) | (II) |
| --- | --- | --- |
| 13–17–20–21 | −54 | −56 |
| 13–17–20–22 | −175 | 178 |
| 16–17–20–22 | 66 | 53 |
| 17–20–22–23 | −112 | −121 |
| 21–20–22–23 | 126 | 110 |
| 20–22–23–24 | −179 | 176 |
| 22–23–24–25 | 124 | 106 |
| 23–24–25–26 | −63 | −67 |
| 23–24–25–27 | 62 | 60 |
| 28–24–25–26 | 174 | 162 |
| 28–24–25–27 | −61 | −71 |

Fig. 7. A projection of a unit cell along [010]. The hydrogen bonds are indicated by broken lines.

Several Newman projections, showing in more detail the dihedral angles about certain bonds in the two molecules of ERC, are given in Fig. 6.

*Packing*

The molecular packing, projected along [010], is illustrated in Fig. 7. The main feature of the structure is the hydrogen-bonded (O-H···O) spiral running in the direction of **b**. The bonding scheme, with O···O lengths of 2·72 and 2·70 Å, can be described as ···O(3$^i$)–H···O(103)–H···O(3)–H···O(103$^{ii}$)–H···, in which each O donates and accepts one H atom and the superscripted atoms are related to those in the reference cell by screw-axis symmetry operations, thus forming infinite hydrophilic helices parallel to *b*. Each molecule is surrounded at a maximum distance of 4·0 Å by 12 neighbouring steroid molecules. There are no unusually close intermolecular contacts that would suggest molecular instability leading to significant conformational changes in solution. This is in accord with the previously mentioned NMR results of La Mar & Budd (1974) and Wing *et al.* (1975).

Further work is being carried out in order to obtain data at low temperature on both ergocalciferol and ergosterol. Attempts to solve the structure of the monoclinic form of ergocalciferol are also being made.

**References**

ALTONA, C., GEISE, H. J. & ROMERS, C. (1968). *Tetrahedron*, **24**, 13–32.
BARTELL, L. S. (1959). *J. Amer. Chem. Soc.* **81**, 3497–3498.
BARTELL, L. S. & BONHAM, R. A. J. (1960). *J. Chem. Phys.* **32**, 824–826.
BERNAL, J. D. & CROWFOOT, D. (1935). *J. Soc. Chem. Ind.* **54**, 701–702.
BRAUN, P. B., HORNSTRA, J., KNOBLER, C., RUTTEN, E. W. M. & ROMERS, C. (1973). *Acta Cryst.* B**29**, 463–469.
CROMER, D. T. & MANN, J. B. (1968). *Acta Cryst.* A**24**, 321–324.
CRUICKSHANK, D. W. J. & ROBERTSON, A. P. (1953). *Acta Cryst.* **6**, 698–705.
DYKE, S. F. (1965). *The Chemistry of Natural Products,* Vol. VI, 1st ed., pp. 271–317. London: Interscience.
FIESER, L. F. & FIESER, M. (1959). *Steroids,* pp. 90–168. New York: Reinhold.
HODGKIN, D., RIMMER, B. M., DUNITZ, J. D. & TRUEBLOOD, K. N. (1963). *J. Chem. Soc.* pp. 4945–4956.
HULL, S. E. & WOOLFSON, M. M. (1976). *Acta Cryst.* B**32**, 2370–2373.
KNOBLER, C., ROMERS, C., BRAUN, P. B. & HORNSTRA, J. (1972). *Acta Cryst.* B**28**, 2097–2103.
KOK, A. J. DE & ROMERS, C. (1974). *Acta Cryst.* B**30**, 1695–1701.
LA MAR, G. N. & BUDD, D. L. (1974). *J. Amer. Chem. Soc.* **96**, 7317–7324.
MAIN, P. (1975). Int. Summer School on Crystallographic Computing, Prague, Paper No. 48.
MAIN, P., WOOLFSON, M. M., LESSINGER, L., GERMAIN, G. & DECLERCQ, J. P. (1974). *MULTAN 74, A System of Computer Programs for the Automatic Solution of Crystal Structures from X-ray Diffraction Data.*
NORMAN, A. W. & HENRY, H. (1974). *Recent Progress in Hormone Research,* Vol. 30, pp. 431–480. New York: Academic Press.

SANDERS, G. M.; POT, J. & HAVINGA, E. (1969). *Progress in the Chemistry of Organic Natural Products*, XXVII, pp. 131–157. Wien: Springer-Verlag.

SAUNDERSON, C. P. (1965). *Acta Cryst.* 19, 187–192.

STEWART, J. M., KRUGER, G. J., AMMON, H. L., DICKINSON, C. & HALL, S. R. (1972). The X-RAY system–version of June 1972. Tech. Rep. TR-192. Computer Science Center, Univ. of Maryland, College Park, Maryland.

STEWART, R. F., DAVIDSON, E. R. & SIMPSON, W. T. (1965). *J. Chem. Phys.* 42, 3175–3187.

SUTTON, L. E. (1965). *Tables of Interatomic Distances and Configuration in Molecules and Ions.* Spec. Publ. No. 18. London: Chemical Society.

WHITE, P. S. & WOOLFSON, M. M. (1975). *Acta Cryst.* A31, 53–56.

WING, R. M., OKAMURA, W. H., REGO, A., PIRIO, M. R. & NORMAN, A. W. (1975). *J. Amer. Chem. Soc.* 97, 4980–4985.

---

# The Crystal and Molecular Structures of Lithium Hydrogen Oxydiacetate

BY H. HERBERTSSON

*Department of Minerals and Prospecting Technology,
Section of Inorganic Chemistry, University of Luleå, S-951 87 Luleå, Sweden*

The structure of $LiC_4H_5O_5$ has been determined from X-ray data collected with an automatic diffractometer, and refined to $R=0.034$. The crystals are monoclinic, space group $P2_1/n$, with $a=7.259$ (2), $b=5.468$ (1), $c=14.256$ (2) Å, $\beta=91.84$ (2)°, $Z=4$. The structure consists of layers of infinite chains of hydrogen-bonded $O(CH_2COO^-)_2$ ions held together by $Li^+$ ions. Between the layers there are only van der Waals forces. The oxydiacetate ion is twisted and the hydrogen bond is somewhat longer, 2.561 (2) Å, than in the other alkali hydrogen oxydiacetates. Li is coordinated by five O atoms which form a distorted trigonal bipyramid with Li–O distances from 2.031 (3) to 2.122 (3) Å. The bipyramids form pairs by sharing one edge. The shortest Li–Li distance is 3.156 (6) Å.

## Introduction

With the determination of the structure of lithium hydrogen oxydiacetate, LiHOXY, the series of alkali hydrogen oxydiacetates is complete. Earlier publications describe the monoclinic structures of NaHOXY and KHOXY (Albertsson, Grenthe & Herbertsson, 1973a), and the tetragonal structures of RbHOXY, CsHOXY and TlHOXY (Albertsson, Grenthe & Herbertsson, 1973b). The monoclinic oxydiacetic acid has also been described (Herbertsson & Boman, 1973). In the alkali hydrogen oxydiacetates so far investigated the anion has been almost planar and the oxydiacetate ions are linked by short hydrogen bonds into infinite chains. The planar oxydiacetate ion is also found in the lanthanoid compounds (Albertsson, 1972), in Ca oxydiacetate (Uchtman & Oertel, 1973) and in two phases of Cd oxydiacetate (Boman, 1974). In a third phase of CdOXY, however, as well as in the monoclinic oxydiacetic acid, the oxydiacetate group is strongly twisted.

The purpose of this investigation is to find out how coordination to the small $Li^+$ ion affects the conformation of the hydrogen oxydiacetate ion and the topology of its hydrogen bonds.

### Crystal data

Lithium hydrogen oxydiacetate, $LiC_4H_5O_5$; F.W. 140·0; monoclinic, space group $P2_1/n$; $a=7.259$ (2),
$b=5.468$ (1), $c=14.256$ (2) Å, $\beta=91.84$ (2)°, $V=565.6$ Å$^3$; $Z=4$; $\mu(Cu\ K\alpha)=13.5$ cm$^{-1}$; $D_m=1.65$, $D_x=1.644$ g cm$^{-3}$. Numbers within parentheses represent estimated standard deviations.

## Experimental

Colourless, short prismatic crystals are formed when an aqueous solution containing equimolar amounts of LiOH and oxydiacetic acid is evaporated slowly at room temperature. Percentages of C and H in good agreement with the values calculated for $LiC_4H_5O_5$ were obtained from elemental analyses. The density was determined from the loss of weight in benzene. Weissenberg photographs showed monoclinic symmetry with systematic extinctions $0k0$ for $k \neq 2n$ and $h0l$ for $h+l \neq 2n$ characteristic of space group $P2_1/n$. The dimensions of the unit cell were determined from data collected with a Guinier–Hägg camera (Cu $K\alpha_1$ radiation, $\lambda=1.54051$ Å, 22 °C) with Al (cubic, $a=4.04934$ Å) as internal standard. A single crystal, $0.15 \times 0.15 \times 0.25$ mm, mounted along **b** was used for data collection on an automatic diffractometer of type CAD-4. Intensities were recorded with Cu $K\alpha$ radiation, monochromatized by reflexion from a graphite crystal, at a take-off angle of 5°. The $\omega$–$2\theta$ technique was used with scan interval $\Delta\omega=0.80+0.50\tan\theta$. The background was measured before and after each reflexion. All 1154