Acta Cryst. (1969). B**25**, 811

Phase Determination for the Estriol Structure*

By Herbert Hauptman, Janet Fisher and Harrison Hancock
U.S. Naval Research Laboratory, Washington, D.C. 20390, *U.S.A.*

and Dorita A. Norton†
Roswell Park Memorial Institute, Buffalo, N.Y. 14203, *U.S.A.*

(*Received* 1 *May* 1968)

Some recently secured results make possible the implementation of a program for direct phase determination first proposed over ten years ago. Application is made to the evaluation of the phases for the estriol structure. The values of 1805 structure invariants, $\cos(\varphi_1+\varphi_2+\varphi_3)$, lead, by means of a novel least-squares technique, to the determination of 103 phases. Employing the tangent formula, these 103 phases are used to obtain the values of the 1023 phases whose corresponding normalized structure factors are not less than unity.

1. Introduction

A procedure for the direct determination of the phases of the structure factors for noncentrosymmetric structures, *via* the values of the structure invariants $\varphi_1+\varphi_2+\varphi_3$, where $\mathbf{h}_1+\mathbf{h}_2+\mathbf{h}_3=0$, was first described over ten years ago (Karle & Hauptman, 1957). This procedure appears never to have been implemented, the probable reason being that the formula there suggested for computing the values of the $\cos(\varphi_1+\varphi_2+\varphi_3)$ is subject to considerable error if significant overlap occurs in the Patterson function. More recent work (Hauptman, 1964), concerned with a study of the various interactions which may occur among the interatomic vectors, led to improved formulas for $\cos(\varphi_1+\varphi_2+\varphi_3)$ which presumably could serve as a reliable base for the phase determination. A first application of these methods has now been made and the present paper contains a description of the actual procedure used for finding the values of the phases required to determine the estriol structure, $C_{18}H_{24}O_3$, which crystallizes in the space group $P2_1$ and has two molecules in the asymmetric unit. Not only is the phase determination carried out through the medium of the structure invariants, $\cos(\varphi_1+\varphi_2+\varphi_3)$, but a novel least-squares technique is here employed which exploits in a systematic way the redundancy inherent in such a program. An essential feature of the method is its dependence on the conditional probability distribution of $\cos(\varphi_1+\varphi_2+\varphi_3)$ which is described in the Appendix.

2. The strategy

Of the 3193 observed, independent reflections, 1023 had normalized structure factor magnitudes not less than unity. It was decided to attempt to determine the values of the 1023 phases, hereafter referred to as permissible phases, corresponding to these largest normalized structure factors.

Two phases, $\varphi_{200}=0$ and $\varphi_{400}=0$, both permissible, were determined by special methods. The phase φ_{400} was found by means of Σ_1 (Hauptman & Karle, 1953), and the phase φ_{200} by an argument which the reader is challenged to supply, and which depends on the facts that $|E_{200}|=5\cdot40$, $|E_{100}|=0\cdot00$, and that there are two molecules in the asymmetric unit.*

Next, three permissible phases, $\varphi_{30\bar{9}}$, $\varphi_{70\bar{4}}$, and $\varphi_{21\bar{1}}$ were arbitrarily specified in order to fix the origin. Of course $\varphi_{30\bar{9}}$ and $\varphi_{70\bar{4}}$ had to be chosen, on account of the space group symmetry, to be 0 or π, while the value of $\varphi_{21\bar{1}}$ could be truly arbitrary (Hauptman & Karle, 1956).

Now we had a basic set of five phases φ_{200}, φ_{400}, $\varphi_{30\bar{9}}$, $\varphi_{70\bar{4}}$, and $\varphi_{21\bar{1}}$ on which the whole phase determination rested. Owing to the space group symmetry, knowledge of a phase φ_{hkl} with $k\neq0$ implied knowledge of the three additional phases $\varphi_{h\bar{k}l}$, $\varphi_{\bar{h}k\bar{l}}$, and $\varphi_{\bar{h}\bar{k}\bar{l}}$, while if $k=0$ only one additional phase $\varphi_{\bar{h}0\bar{l}}$, was known. Thus we started with a knowledge of twelve phases.

Corresponding to each distinct pair $\varphi_{\mathbf{k}},\varphi_{-\mathbf{h}-\mathbf{k}}$ of the basic set such that the phase $\varphi_{\mathbf{h}}$ was permissible, we constructed the associated structure invariant $\varphi_{\mathbf{h}}+\varphi_{\mathbf{k}}+\varphi_{-\mathbf{h}-\mathbf{k}}$ and the product $|E_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{h}+\mathbf{k}}|$. For each vector $\mathbf{h}$ so defined, the sum, $\sum_{\mathbf{k}}|E_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{h}+\mathbf{k}}|$, (which may have consisted of only a single term) taken over all allowed vectors $\mathbf{k}$ was also computed. These sums (one for each $\mathbf{h}$) were arranged in decreasing order and the largest one was selected, thus defining a unique vector $\mathbf{h}$ and several (perhaps only one) structure invariants

* Presented at the meeting of the American Crystallographic Association in Tucson, Arizona, February, 1968.

† Present address: Medical Foundation of Buffalo, Buffalo, N.Y. 14203.

* It should be noted that even if the value of φ_{200} had not been determined at this time the whole phase determination could have been carried out by use of the two possible values (0 or π) for φ_{200} with only a negligible increase in total computer time required. Two Fourier series could then have been computed and the incorrect one rejected.

$\cos(\varphi_{\mathbf{h}}+\varphi_{\mathbf{k}}+\varphi_{-\mathbf{h}-\mathbf{k}})$ the values of which were determined in a manner to be described (§ 3). Since the values $\varphi_{\mathbf{k}}$, $\varphi_{-\mathbf{h}-\mathbf{k}}$, and the several $\cos(\varphi_{\mathbf{h}}+\varphi_{\mathbf{k}}+\varphi_{-\mathbf{h}-\mathbf{k}})$ were presumed known, the value of $\varphi_{\mathbf{h}}$ was actually overdetermined (in general). The new phase, $\varphi_{\mathbf{h}}$, and its three (or possibly only one) symmetry related phases were added to the basic set of known phases and the process repeated. During the second cycle, however, the largest three sums $\sum_{\mathbf{k}} |E_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{h}+\mathbf{k}}|$ were selected so that three new phases were determined, rather than only one as in the first cycle. During the third, fourth, fifth, ... cycles, the numbers of new phases determined were five, seven, nine, ... Furthermore, during each cycle, all phases determined in the earlier cycles were automatically redetermined so that the procedure possessed a refinement feature whereby the values of new phases were used in order to improve the values of phases initially determined at an earlier stage. The process could have been continued until all 1023 permissible phases were obtained although, as explained later, not all 1023 phases were evaluated in this way. Thus, the order in which the values of the phases were to be obtained was determined, and the identity of the structure invariants, $\cos(\varphi_{\mathbf{h}}+\varphi_{\mathbf{k}}+\varphi_{-\mathbf{h}-\mathbf{k}})$, whose values were required was also found.

A 'dummy' run showed that all 1023 permissible phases could have been reached in the manner described. However, the values of only 103 of these phases were actually determined in this way (§ 4). The more efficient tangent formula was then employed to determine the remaining phases (§ 5).

3. Evaluation of the structure invariants $\cos(\varphi_1+\varphi_2+\varphi_3)$

In order to carry out twelve cycles of the process described in § 2, it was necessary to find the values of 1805 structure invariants $\cos(\varphi_{\mathbf{h}}+\varphi_{\mathbf{k}}+\varphi_{-\mathbf{h}-\mathbf{k}})$. To this end it was decided to use a variant of equation 5·12 (Hauptman, 1964):

$$|E_1E_2E_3|\cos(\varphi_1+\varphi_2+\varphi_3) \simeq \frac{\sigma_1^3-3\sigma_1\sigma_2+2\sigma_3}{\sigma_2^{3/2}\langle(|E_{\mathbf{k}}|^2-1)^3\rangle_{\mathbf{k}}} \times \langle(|E_{\mathbf{k}}|^2-1)(|E_{\mathbf{h}_1+\mathbf{k}}|^2-1)(|E_{-\mathbf{h}_3+\mathbf{k}}|^2-1)\rangle_{\mathbf{k}} + \frac{\sigma_3}{\sigma_2^{3/2}}(|E_1|^2+|E_2|^2+|E_3|^2-2)\,, \qquad (3{\cdot}1)$$

where $\mathbf{k}$ ranges over all the (nearly 4×3193) vectors corresponding to observed intensities. As shown in the latter reference, earlier versions of this formula, *e.g.* equation 2·2 (Karle & Hauptman, 1957) or equation 2·1·3 (Karle & Hauptman, 1958), would not be valid if the structure contained a significant number of induced or chance interactions. It is known that the presence of such interactions causes the average values of $(|E_{\mathbf{k}}|^2-1)^2$ and $(|E_{\mathbf{k}}|^2-1)^3$ to increase. These averages were computed and compared with the theoretical values which obtain if no induced or chance interactions are present. The results are exhibited in Tables 1 and 2. These Tables clearly show the presence of substantial numbers of induced and chance interactions so that the earlier formulas would not be expected to hold. They were, in fact, grossly in error. Even the more recent (3·1), however, while a decided improvement over the earlier versions, occasionally led to values of 4 or 5 for $\cos(\varphi_1+\varphi_2+\varphi_3)$! The source of the difficulty is not hard to find. While the structure is heavily infested with induced and chance interactions, these interactions are only approximate. In order to show this, the average values of $(|E_{\mathbf{k}}|^2-1)^2$ and $(|E_{\mathbf{k}}|^2-1)^3$ were computed as functions of $s=(\sin\theta)/\lambda$. It was observed that the local averages of $(|E_{\mathbf{k}}|^2-1)^2$ ranged from about 3 for the smallest values of s to about 1·2 for the larger s values. Similarly the average values of $(|E_{\mathbf{k}}|^2-1)^3$ ranged from about 27 to about 2·6 over the same range of s values. In short, the induced and chance interactions, even if only approximate, lead to extremely large local averages of $(|E_{\mathbf{k}}|^2-1)^q$, $q=2,3$, for small s since approximate interactions 'appear' to be exact at low resolution. However, the almost normal local averages of $(|E_{\mathbf{k}}|^2-1)^q$, $q=2,3$, when s is large, show conclusively that most of the induced and chance interactions are only approximate and that almost all of the exact interactions must be the valid ones; otherwise the local averages of $(|E_{\mathbf{k}}|^2-1)^q$ for large s would be larger than actually found (Hauptman, 1964).

Table 1. *Values of* $\langle(|E_{hkl}|^2-1)^2\rangle_{hkl}$

	Actual	Theoretical	No. of contributors
$k\neq0$	1·70	0·99	3030
$k=0$	6·42	1·96	163
All k	1·94	1·04	3193

Table 2. *Values of* $\langle(|E_{hkl}|^2-1)^3\rangle_{hkl}$

	Actual	Theoretical	No. of contributors
$k\neq0$	8·23	1·92	3030
$k=0$	141·80	7·55	163
All k	15·06	2·21	3193

These facts clearly call for replacing the coefficient of the average on the right side of (3·1) by a sliding scale factor. This was done by substituting for this constant coefficient a function of $|E_1E_2E_3|$, in such a way that the empirical conditional distribution, for fixed $|E_1E_2E_3|$, of the values of $\cos(\varphi_1+\varphi_2+\varphi_3)$ which resulted, coincided with the known theoretical conditional probability distribution (Appendix) of $\cos(\varphi_1+\varphi_2+\varphi_3)$ (*cf.* equation 7·13 of Hauptman (1964) where, however, a different method of obtaining the scale factor is described). One final modification was introduced. Instead of using equation (3·1), which is the analogue of the case $p=q=r=2$ of the earlier equation 2·1·3 (Karle & Hauptman, 1958), we employed the analogue corresponding to the case $p=q=r=\frac{1}{2}$ of the latter equation in order to reduce the variance caused by the finite sampling (since all computed averages are of necessity only estimates of the true averages based on the finite number of data available from experiment):

$$|E_1E_2E_3|\cos(\varphi_1+\varphi_2+\varphi_3) \simeq K\langle(|E_{\mathbf{k}}|^{1/2}-\overline{|E|^{1/2}})(|E_{\mathbf{h}_1+\mathbf{k}}|^{1/2}-\overline{|E|^{1/2}})(|E_{-\mathbf{h}_3+\mathbf{k}}|^{1/2}-\overline{|E|^{1/2}})\rangle_{\mathbf{k}}+R_3\,, \qquad (3\cdot2)$$

where

$$R_3=\frac{\sigma_3}{4\sigma_2^{3/2}}[\tfrac{3}{2}(|E_1E_2|^2+|E_2E_3|^2+|E_3E_1|^2)+|E_1|^2+|E_2|^2+|E_3|^2-\tfrac{7}{2}]\,, \qquad (3\cdot3)$$

and

$$\overline{|E|^{1/2}}=\langle|E_{\mathbf{k}}|^{1/2}\rangle_{\mathbf{k}}\,. \qquad (3\cdot4)$$

The averages on the right hand side of (3·2) were computed for the 1805 structure invariants $\cos(\varphi_1+\varphi_2+\varphi_3)$ where the abbreviations

$$\varphi_1=\varphi_{\mathbf{h}_1}=\varphi_{\mathbf{h}}\,,\quad \varphi_2=\varphi_{\mathbf{h}_2}=\varphi_{\mathbf{k}}\,,\quad \varphi_3=\varphi_{\mathbf{h}_3}=\varphi_{-\mathbf{h}-\mathbf{k}}\,, \qquad (3\cdot5)$$

have been used, so that

$$\mathbf{h}_1+\mathbf{h}_2+\mathbf{h}_3=0\,. \qquad (3\cdot6)$$

These averages were arranged in decreasing order of $|E_1E_2E_3|$ and grouped into sets having 200 to 300 members in each group. Thus, in each group the values of $|E_1E_2E_3|$ were essentially constant. The parameter K appearing on the right hand side of (3·2) was determined for each group in such a way that the resulting empirical distribution coincided with the theoretical distribution (equation (1) of the Appendix).

4. Least squares determination of initial phases

The values of the 1805 structure invariants $\cos(\varphi_1+\varphi_2+\varphi_3)$ called for by twelve cycles of the procedure described in § 2 were computed by the method of § 3. The situation then was as follows: for each fixed **h**, whose corresponding phase $\varphi_{\mathbf{h}}$ was to be determined, the values of several structure invariants

$$\cos(\varphi_{\mathbf{h}}+\varphi_{\mathbf{k}}+\varphi_{-\mathbf{h}-\mathbf{k}})=c_{\mathbf{k}} \qquad (4\cdot1)$$

were known, with weights

$$w_{\mathbf{k}}=|E_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{h}+\mathbf{k}}|\sqrt{n}\,, \qquad (4\cdot2)$$

where n is the number of contributors to the average (3·2) from which $c_{\mathbf{k}}$ was found. The values of the phases $\varphi_{\mathbf{k}},\varphi_{-\mathbf{h}-\mathbf{k}}$ were also known. The phase $\varphi_{\mathbf{h}}$ was then determined by minimizing

$$\Phi=\frac{\sum\limits_{\mathbf{k}} w_{\mathbf{k}}[\cos(\varphi_{\mathbf{h}}+\varphi_{\mathbf{k}}+\varphi_{-\mathbf{h}-\mathbf{k}})-c_{\mathbf{k}}]^2}{\sum\limits_{\mathbf{k}} w_{\mathbf{k}}} \qquad (4\cdot3)$$

which, after some simplification, finally reduces to

$$\Phi=\tfrac{1}{2}C_2\cos 2\varphi_{\mathbf{h}}-\tfrac{1}{2}S_2\sin 2\varphi_{\mathbf{h}}-2C_1\cos\varphi_{\mathbf{h}}+2S_1\sin\varphi_{\mathbf{h}}+c\,, \qquad (4\cdot4)$$

where

$$C_2=\frac{\sum\limits_{\mathbf{k}} w_{\mathbf{k}}\cos 2(\varphi_{\mathbf{k}}+\varphi_{-\mathbf{h}-\mathbf{k}})}{\sum\limits_{\mathbf{k}} w_{\mathbf{k}}}=\langle\cos 2(\varphi_{\mathbf{k}}+\varphi_{-\mathbf{h}-\mathbf{k}})\rangle_{\mathbf{k}}\,, \qquad (4\cdot5)$$

$$S_2=\frac{\sum\limits_{\mathbf{k}} w_{\mathbf{k}}\sin 2(\varphi_{\mathbf{k}}+\varphi_{-\mathbf{h}-\mathbf{k}})}{\sum\limits_{\mathbf{k}} w_{\mathbf{k}}}=\langle\sin 2(\varphi_{\mathbf{k}}+\varphi_{-\mathbf{h}-\mathbf{k}})\rangle_{\mathbf{k}}\,, \qquad (4\cdot6)$$

$$C_1=\frac{\sum\limits_{\mathbf{k}} w_{\mathbf{k}}c_{\mathbf{k}}\cos(\varphi_{\mathbf{k}}+\varphi_{-\mathbf{h}-\mathbf{k}})}{\sum\limits_{\mathbf{k}} w_{\mathbf{k}}}=\langle c_{\mathbf{k}}\cos(\varphi_{\mathbf{k}}+\varphi_{-\mathbf{h}-\mathbf{k}})\rangle_{\mathbf{k}}\,, \qquad (4\cdot7)$$

$$S_1=\frac{\sum\limits_{\mathbf{k}} w_{\mathbf{k}}c_{\mathbf{k}}\sin(\varphi_{\mathbf{k}}+\varphi_{-\mathbf{h}-\mathbf{k}})}{\sum\limits_{\mathbf{k}} w_{\mathbf{k}}}=\langle c_{\mathbf{k}}\sin(\varphi_{\mathbf{k}}+\varphi_{-\mathbf{h}-\mathbf{k}})\rangle_{\mathbf{k}}\,, \qquad (4\cdot8)$$

$$c=\frac{\sum\limits_{\mathbf{k}} w_{\mathbf{k}}(\frac{1}{2}+c_{\mathbf{k}}^2)}{\sum\limits_{\mathbf{k}} w_{\mathbf{k}}}=\langle(\tfrac{1}{2}+c_{\mathbf{k}}^2)\rangle_{\mathbf{k}}\,. \qquad (4\cdot9)$$

The first structure invariant $\cos(\varphi_{\mathbf{h}}+\varphi_{\mathbf{k}}+\varphi_{-\mathbf{h}-\mathbf{k}})$ whose value was different from unity yielded two values for $\varphi_{\mathbf{h}}+\varphi_{\mathbf{k}}+\varphi_{-\mathbf{h}-\mathbf{k}}$ differing only in sign. One of these values was arbitrarily chosen, thus fixing the enantiomorph and leading to a unique value for the corresponding phase $\varphi_{\mathbf{h}}$. The remaining phases were then uniquely determined. Twelve cycles of this least-squares technique yielded stable values for 103 phases which were assumed to be reliably determined and which served as the base for the determination of the remaining phases by means of the tangent formula.

5. Application of the tangent formula

Using the values of the 103 phases determined in § 4, the tangent formula [Karle & Hauptman, 1956, equation (5·62)],

$$\tan\varphi_{\mathbf{h}}=\frac{\sum\limits_{\mathbf{k}}|E_{\mathbf{k}}E_{\mathbf{h}-\mathbf{k}}|\sin(\varphi_{\mathbf{k}}+\varphi_{\mathbf{h}-\mathbf{k}})}{\sum\limits_{\mathbf{k}}|E_{\mathbf{k}}E_{\mathbf{h}-\mathbf{k}}|\cos(\varphi_{\mathbf{k}}+\varphi_{\mathbf{h}-\mathbf{k}})}\,, \qquad (5\cdot1)$$

was employed to find the values of all 1023 permissible phases. The order and the rate (21, 23, 25, ...) in which new phases were found had already been determined by the procedure described in § 2. In addition, in each cycle, phases obtained earlier were redetermined.

Once the 1023 permissible phases were found, an E map yielded the crystal structure which was refined by standard least-squares techniques and is described in the accompanying paper (Cooper, Norton & Hauptman, 1969).

APPENDIX

The conditional probability distribution of cos $(\varphi_{\mathbf{h}}+\varphi_{\mathbf{k}}+\varphi_{-\mathbf{h}-\mathbf{k}})$

Let the vector **h** be fixed and assume that **k** ranges uniformly throughout reciprocal space. Denote by $P(x|\,|E_{\mathbf{k}}|,|E_{\mathbf{h}+\mathbf{k}}|)$ the conditional probability distribution of the random variable $X=\cos(\varphi_{\mathbf{h}}+\varphi_{\mathbf{k}}+\varphi_{-\mathbf{h}-\mathbf{k}})$, given that $|E_{\mathbf{k}}|$ and $|E_{\mathbf{h}+\mathbf{k}}|$ have specified, fixed values. By means of an analysis to be published at a later date we find

$$\left.\begin{aligned} P(x|\,|E_{\mathbf{k}}|,|E_{\mathbf{h}+\mathbf{k}}|) &\simeq \frac{\exp Ax}{\pi I_0(A)\sqrt{1-x^2}} \quad \text{if } |x|<1\,, \\ &= 0 \qquad\qquad\qquad\quad \text{if } |x|>1\,, \end{aligned}\right\} \tag{1}$$

where

$$A=\frac{2}{N^{1/2}}\,|E_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{h}+\mathbf{k}}|\,, \tag{2}$$

I_0 is the Bessel function of imaginary argument, and N is the number of atoms, assumed identical, in the unit cell.* Since, as explained in § 3, this distribution was needed in order to evaluate the structure invariants $\cos(\varphi_1+\varphi_2+\varphi_3)$, (1) has been tabulated for appropriate values of A. This function is readily computed with the aid of modern computers and the tabulation is, therefore, not reproduced here.

References

Cochran, W. (1965). *Acta Cryst.* **8**, 473.

Cooper, A., Norton, D. & Hauptman, H. (1969). *Acta Cryst.* B**25**, 0 .

Hauptman, H. (1964). *Acta Cryst.* **17**, 1421.

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

Hauptman, H. & Karle, J. (1956). *Acta Cryst.* **9**, 45.

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

Karle, J. & Hauptman, H. (1957). *Acta Cryst.* **10**, 515.

Karle, J. & Hauptman, H. (1958). *Acta Cryst.* **11**, 264.

* A related probability distribution, from which (1) may be derived, is given by Cochran (1955).

Acta Cryst. (1969). B**25**, 814

Estrogenic Steroids. III. The Crystal and Molecular Structure of Estriol

By A. Cooper and D. A. Norton

The Medical Foundation of Buffalo, 73 High Street, Buffalo, N.Y., U.S.A.

and H. Hauptman

U.S. Naval Research Laboratory, Washington, D.C. 20390, U.S.A.

(*Received* 1 *May* 1968)

Crystal data for estriol ($C_{18}H_{24}O_3$) are: $a=9{\cdot}270$, $b=23{\cdot}001$, $c=7{\cdot}560$ Å, $\beta=110{\cdot}90$. Space group $P2_1$ with two molecules in the asymmetric unit. The structure was solved from a set of phases derived initially by application of a phase determining formula of the type

$$K|E_1E_2E_3|\cos(\varphi_1+\varphi_2+\varphi_3)=\langle(|E_k|^p-\langle|E|^p\rangle)(|E_{h_1+k}|^p-\langle|E|^p\rangle)(|E_{-h_3+k}|^p-\langle|E|^p\rangle)\rangle_k+R$$

and extended by application of the tangent formula. The structure was refined by block diagonal least-squares with anisotropic thermal parameters for the non-hydrogen atoms, to $R=5{\cdot}6\%$. All hydrogen atoms were located. Standard deviations of non-hydrogen distances and angles are 0·007 Å and 0·4° respectively.

Differences in hydrogen bonding, packing environment and intramolecular steric effects cause the two molecules in the asymmetric unit to be non-identical. The two molecules are hydrogen bonded head-to-tail *via* the 3-hydroxyl oxygen of one molecule and the 16α-hydroxyl hydrogen of another, in such a way that the 18-methyl group of the first lies under the *A* ring of the second, producing distortion of this aromatic ring. Steric hindrance between the *C* ring equatorial hydrogen atom at C(11), and the hydrogen atom at C(1) produces twisting about the C(9)–C(10) bond, in opposite directions in the two molecules, causing the *B* ring to take up a half chair conformation in the first and a twist boat conformation in the second.

The molecules pack with molecules of the first kind and molecules of the second kind, hydrogen bonded in separate chains parallel to **b**. These chains are cross linked *via* the asymmetric unit and *via* weaker hydrogen bonds involving all three hydroxyl groups. Three distinct types of hydrogen bond exist in which the O···H–O angles are 159–177°, 136°, and 128°.

Introduction

The estrogens, of which the principal members are estradiol, estrone and estriol, are essential for the development and maintenance of the secondary female sex characteristics. They are produced mainly by the ovary and to a small extent by the adrenal cortex but during pregnancy, the placenta produces relatively