

738. The X-Ray Crystallographic Determination of the Structure of Bromomiroestrol.

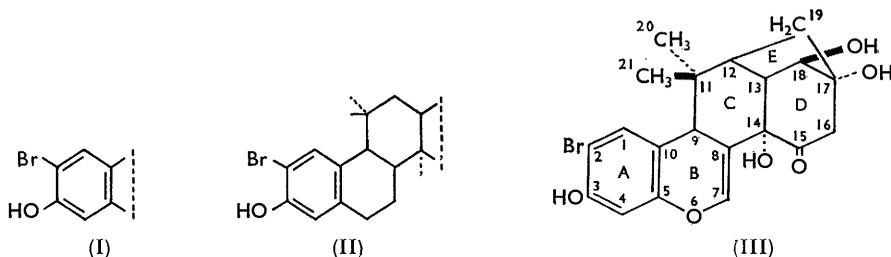
By NOEL E. TAYLOR, DOROTHY CROWFOOT HODGKIN, and J. S. ROLLETT.

The structure of the plant oestrogen, miroestrol, in the form of its monobromo-derivative, has been investigated by two- and three-dimensional methods of X-ray analysis. These lead to the structure (VI), which corresponds to the chemical formula $C_{20}H_{22}O_6$ and incorporates a partly reduced isoflaven nucleus. The bromine atom was first placed by the calculation of a Patterson projection along [010] in the monoclinic crystal structure, and the positions of the carbon and oxygen atoms were then found by three successive calculations of the three-dimensional electron-density distribution. The distinction between carbon and oxygen atoms in the molecule was made from a projection of the difference between the electron density derived with the observed F values and with a set calculated by assuming that all the atoms present (other than bromine) were carbon.

The structure found for miroestrol appears to have interesting biogenetic relations and to account both for the chemical properties of the compound and, in general terms, for its biological activity.

THE study of the structure of the plant oestrogen, miroestrol, described in this paper, has been carried out in close collaboration with Drs. D. G. Bounds and G. S. Pope, who prepared the active compound from the plant *Pueraria mirifica* found in Thailand. The use of X-ray methods to investigate its structure was initially suggested on account of the very small amount of material they could obtain. The structure now found by the X-ray analysis of monobromomiroestrol appears to account very well for the chemical and spectroscopic properties of miroestrol which they describe in the following paper.¹

Miroestrol itself was first examined crystallographically by Dr. M. S. Webster in two differently hydrated crystalline modifications. The molecular weight determined from the measured density and unit-cell dimensions agreed with the formula $C_{20}H_{22}O_6$ but the accuracy was not sufficient to rule out the alternative analytically permissible formula $C_{20}H_{24}O_6$.¹ The complexity of the molecule was clearly too great to permit a detailed X-ray analysis of its structure without the introduction of a heavy atom into the molecule. A bromine derivative was sought, since the average scattering factor for X-rays of the miroestrol molecule as a whole, measured as $\sqrt{\sum f^2}$, is approximately equivalent to that of a bromine atom;² the monobromo-derivative, when prepared, appeared to be a suitable simple substitution product.



The molecular structure for monobromomiroestrol (III) has been deduced by two- and three-dimensional methods of X-ray analysis. Initially the positions of the bromine atoms in the crystal were determined from Patterson projections. Two following electron-density projections served to place the nine additional atoms of the structural unit (I), and the remaining carbon and oxygen atoms were located in stages (II) and (III), through

¹ Bounds and Pope, following paper.

² Cf. Mathieson, *Rev. Pure Appl. Chem.*, 1955, 5, 113.

successive calculations of the electron density in three dimensions. The distinction between carbon and oxygen atoms in the molecule was made by the calculation of difference maps. The structure (III) corresponds with the formula $C_{20}H_{22}O_6$ for miroestrol itself.

EXPERIMENTAL

Bromomiroestrol was recrystallised from aqueous alcohol in thin, transparent, lath-like crystals, elongated along the b axis, with $\{001\}$ dominating. All crystals from the first sample studied (type I) proved to have a disordered crystal structure. They gave sharp X -ray reflections of widely varying intensity, interleaved by weak diffuse reflections. The sharp reflections could be referred to a monoclinic unit cell of dimensions $a = 7.23 \text{ \AA}$, $b = 5.98 \text{ \AA}$, $c = 20.33 \text{ \AA}$, $\beta = 92.4^\circ$ with $0k0$ reflections absent for k odd and two molecules in the unit cell (calculated density 1.65_6 g./c.c.). The weak diffuse reflections corresponded with a unit cell having a doubled and four molecules per cell.

A later preparation of bromomiroestrol, which crystallised in much larger crystals (type II) gave X -ray photographs on which all reflections were sharp, corresponding with an ordered unit cell containing four molecules with $a = 14.46 \text{ \AA}$, $b = 5.98 \text{ \AA}$, $c = 20.33 \text{ \AA}$, $\beta = 92.4^\circ$, space group $P2_1$ and two molecules of bromomiroestrol in the asymmetric unit.

The relation between the two types of crystal is extremely close. All the X -ray reflections common to the two sets of photographs are indistinguishable by eye in relative intensity. Further ($h0l$) reflections when h is odd, given by type II crystals, are represented only by three very weak reflections while type I crystals give no interleaving diffuse ($h0l$) reflections. The effects can be explained by small relative translations parallel to the b axis of alternate molecules following one another along a ; these translations must occur in a regular repetitive manner in type II crystals, in a random way in those of type I.

The structure analysis was almost complete before type II crystals were found and is based solely on the sharp reflections of type I crystals. These were treated as if they referred to a crystal of space group $P2_1$ with two molecules in the unit cell.

(hkl) reflections were recorded for $k = 0$ to $k = 4$ by the multiple-film Weissenberg technique with a crystal of dimensions $0.11 \times 0.75 \times 0.009 \text{ mm.}$, set for rotation about its needle axis. Exposure times of the order of 160 hr. were necessary. The intensities were estimated visually for the 1150 reflections recorded in this way and the data from the individual layers put on to the same scale by comparison with an independent set of ($0kl$) data. Corrections for the Lorentz and polarisation factors were made as usual. The intensity of reflection falls off very rapidly with increasing k and, although reflections with k up to 7 fall within the reflection sphere of $Cu-K\alpha$ radiation, these higher-order terms must be relatively very small. The rapid fall-off in intensity in this direction is obviously associated with the disorder in the crystals, but it could be treated in the same way as the fall-off resulting from an anisotropic thermal movement of the atoms. The data were put on to an approximately absolute scale by Wilson's method,³ Wilson plots of the individual layers being used to give values of the temperature factors normal and parallel to the b -axis, of $B_1 = 2.53 \times 10^{-16} \text{ cm.}^2$ and $B_2 = 8.64 \times 10^{-16} \text{ cm.}^2$. At a later stage, plots of $\log (\sum F_c/F_o)$ against $\sin^2 \theta$, over small ranges of $\sin^2 \theta$, showed the difference to be even greater than this; the final temperature constants were found to be $B_1 = 2.55 \times 10^{-16} \text{ cm.}^2$ and $B_2 = 11.25 \times 10^{-16} \text{ cm.}^2$.

STRUCTURE ANALYSIS

The relative scattering weight of bromine compared with that of miroestrol is very favourable for a direct application of heavy-atom methods of structure analysis. But, in the space group $P2_1$, the relative positions of the two bromine atoms in the crystal are centro-symmetrical when the structure as a whole is not. Except in the single centro-symmetrical projection of the structure down the b axis, electron-density distributions calculated with phase constants derived from the bromine-atom contributions have false symmetry; they show peaks due to the remaining atoms in the molecule superimposed on those of a centro-symmetrically related mirror-image molecule, as in the analysis of cholesteryl iodide.⁴ The main problem in the structure determination of miroestrol there-

³ Wilson, *Nature*, 1942, **150**, 151.

⁴ Carlisle and Crowfoot, *Proc. Roy. Soc.*, 1945, *A*, **184**, 64.

fore turned on distinguishing in three dimensions the atoms of one molecule from its mirror image. The course taken was in outline as follows:

1. The sharpened Patterson function was calculated for the (010) projection by using the derived F^2 values and a modifying function $\exp(8 \sin^2 \theta / \lambda^2)$. The x - and z -parameters of the bromine atom were clearly defined by the heaviest peak present; at the end of the analysis it was found that all the main features of the projection were accounted for in a calculated map of bromine-light atom vectors. The bromine atoms were assigned the initial parameters 0.185, $\frac{1}{4}$, 0.026; 0.815, $\frac{3}{4}$, 0.974.

2. With signs based on the bromine atoms alone electron-density projections were calculated along the b and the a axis.

The b -axis projection is itself centro-symmetrical. In its outline it was possible to recognise immediately that the bromine atom was substituted on to a (probably aromatic)

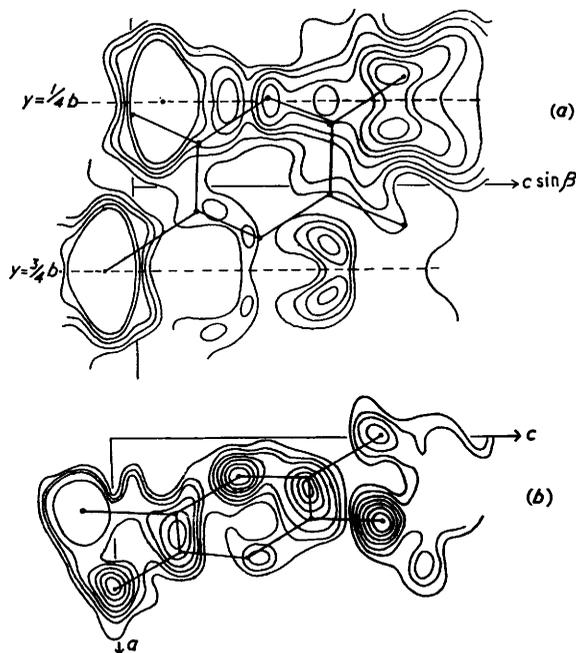


FIG. 1. Part of the electron-density projections (a) along [100], (b) along [010], calculated with the bromine phase constants. The positions of the first nine atoms selected for the hkl structure-factor calculations are shown. The first contour is at $2e/\text{\AA}^2$, the contour interval being approximately $1e/\text{\AA}^2$. The higher contours are omitted over the bromine atom.

six-membered ring fused to a second ring. Attached to the first ring was an atom, *ortho* to the bromine atom, which seemed likely to be the oxygen of a phenolic hydroxyl group. Little attempt was made to interpret the remaining peak pattern at this stage. The complete analysis shows that it includes peaks of over $3e/\text{\AA}^2$ at every atomic position finally found but, in addition, six peaks of this magnitude or greater which are spurious.

The a -axis projection shows the expected false symmetry, a centre of symmetry at the origin which introduces a mirror plane through the bromine-atom positions parallel with the b plane. In addition, the peaks are drawn out parallel to b . By comparison with the b -axis projection, x , y , and z parameters were assigned to the nine atoms of the molecule shown in (I) and Fig. 1a and b, one of the two mirror-image distributions being arbitrarily selected in the a -axis projection.

3. The contributions of the bromine atoms and the nine light atoms in (I) to the hkl structure factors were calculated, and the resulting phases used in the computation of the three-dimensional electron-density distribution, ρ_1 . The suspected oxygen atom at position 3 was weighted as oxygen, the remaining light atoms as carbon atoms.

The inclusion of the nine light-atom contributions in the phase calculation only partly destroyed the false symmetry of the electron-density distribution in ρ_1 . High peaks

occurred at each of the chosen atomic positions and there were no corresponding image peaks. For each of the other possible atomic positions, with the exception of those lying on or very near the false mirror planes, two mirror-image peaks occurred with co-ordinates x, y, z , and $x, \frac{1}{2} - y, z$. The peak-height relations are shown in Table 1, $\rho 1$.

A spoke model was built of the maxima in $\rho 1$, and their stereochemical relations were examined. In general it was found that the higher of the two image peaks was more favourably placed stereochemically and the sites of twenty light atoms were selected. These are indicated in (II) and Table 1, $\rho 2$. In fact, in only two of the remaining positions were the image peaks in $\rho 1$ slightly higher than the real peaks selected later.

TABLE 1. Peak heights in $e/\text{\AA}^3$ in $\rho 1$, $\rho 2$, and $\rho 3$.

Atom no.	$\rho 1$	$\rho 2$	$\rho 3$	Atom no.	$\rho 1$	$\rho 2$	$\rho 3$
O 3	8.2	9.3	9.3	C 9	7.4	6.7	7.6
O 6	8.4	8.2	9.9	C 10	7.4	6.8	7.3
O 14	3.3 (2.9)	7.6	10.4	C 11	2.9 (1.8)	7.1	7.3
O 15	3.1 (3.3)	4.0 (3.2)	10.4	C 12	* 3.3	5.1	6.4
O 17	3.7 (3.4)	5.0 (3.2)	11.0	C 13	* 3.6	5.8	6.9
O 18	4.0 (3.3)	4.5 (3.3)	11.3	C 14	* 4.7	7.0	7.6
C 1	8.7	8.5	8.2	C 15	3.2 (2.7)	7.6	7.4
C 2	7.9	6.9	7.6	C 16	2.9 (3.1)	3.5 (2.7)	7.9
C 3	6.3	6.6	6.6	C 17	3.5 (2.8)	4.1 (2.6)	8.1
C 4	6.2	6.2	6.9	C 18	* 4.4	7.2	7.5
C 5	6.0	6.5	6.9	C 19	2.9 (2.6)	3.0 (1.8)	7.2
C 7	3.7 (2.8)	7.6	8.0	C 20	3.1 (1.8)	7.1	6.8
C 8	3.3 (2.3)	7.4	7.7	C 21	* 2.9	6.2	7.0

Numbers in italics indicate peak heights of atoms not placed in the preceding phasing calculations, the height of the image peaks being given in parentheses.

* Peaks which fall so close to the false mirror plane that real and image peaks are unresolved.

Other than the maxima listed in Table 1 there were only three peaks in $\rho 1$ of height $2e/\text{\AA}^3$ or greater; all three were found to be spurious. There were two peaks of height 2.2 and $2.3e/\text{\AA}^3$ falling almost on the $z = \frac{1}{2}$ plane (see Fig. 7a), and a peak of height $3.6e/\text{\AA}^3$ on the $y = \frac{1}{4}$ plane. The last of these and a trough of approximately $-6e/\text{\AA}^3$ on the $y = \frac{3}{4}$ plane are found to occur at Br-Br vector repeat positions.*

4. A set of (hkl) structure factors was calculated from the contributions of the bromine and the twenty light atoms chosen from $\rho 1$. Here all light atoms were weighted as carbon except O 3, oxygen. With the resulting phase constants, a second electron-density distribution, $\rho 2$, was calculated. The peak heights observed are listed in Table 1. None of the atoms included in the structure-factor calculations had image peaks. For the six atoms omitted, there were again two peaks, one of which was now about $1e/\text{\AA}^3$ higher than the other. In their interpretation it was clear that in each case the higher peak was the real atom. The skeleton structure (III) was defined.

5. Direct evidence as to which were the oxygen atoms in the molecule was obtained from the (010) projection. A set of $(h0l)$ structure factors was calculated in which all the light atoms were weighted as carbon atoms. The reliability index R was found to be 0.20 (cf. Table 2). The electron-density projection shown in Fig. 2 was calculated, the resulting signs and the observed F 's being used. Compared with the first calculation there were 12 reversals of sign, the most important of which was for the term $(\bar{1}02)$ which is relatively large, and a further 23 small terms could be included in the calculation. The spurious

* This effect can be explained as follows. A Fourier series with phases given by the contributions of atoms at (x, y, z) and $(\bar{x}, \bar{y}, \bar{z})$ only can be expected to have a spurious minimum at $(3x, 3y, 3z)$. If $hx + ky + lz = \theta$, then the (hkl) term used is $\frac{F_o(hkl) \cos \theta}{|\cos \theta|}$, and the effect of this term at $(3x, 3y, 3z)$ is $\frac{F_o(hkl) \cos \theta \cos 3\theta}{|\cos \theta|}$. The phase of the term is most likely to be wrong when $|\cos \theta|$ is small and, when this is so, $\cos \theta \cos 3\theta$ is always ≤ 0 . Similarly $\cos \theta \cos 5\theta$ is ≥ 0 when $|\cos \theta|$ is small, but the range of θ values for which this is true is smaller ($54-126^\circ$ instead of $30-150^\circ$). It is interesting that the effect persists even when the phases used are based on nine out of twenty-six light atoms as well as the bromine atom.

FIG. 2. The electron density in bromomiræstrol projected along [010]. The first contour is at $2e/\text{\AA}^2$, the contour interval being $1e/\text{\AA}^2$ except over the bromine atom where the interval is $5e/\text{\AA}^2$.

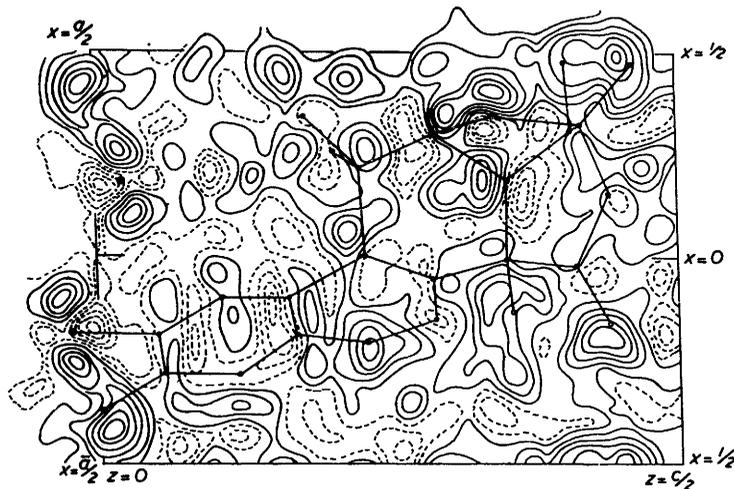
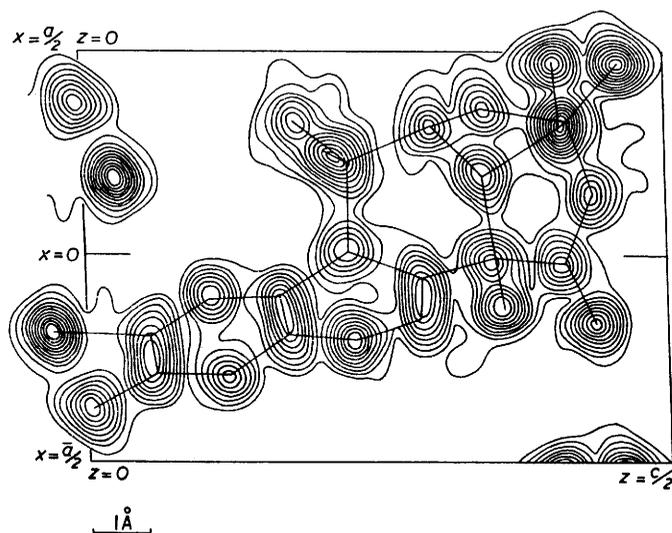
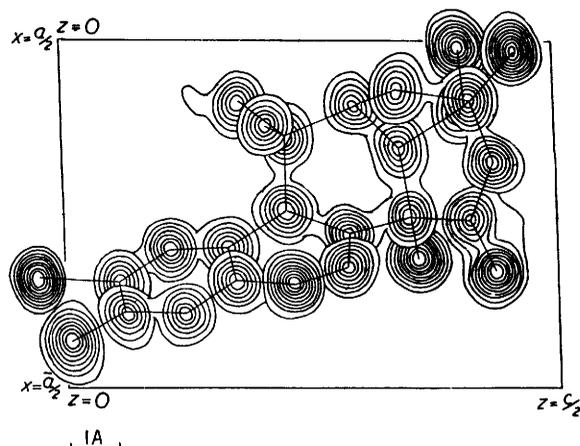


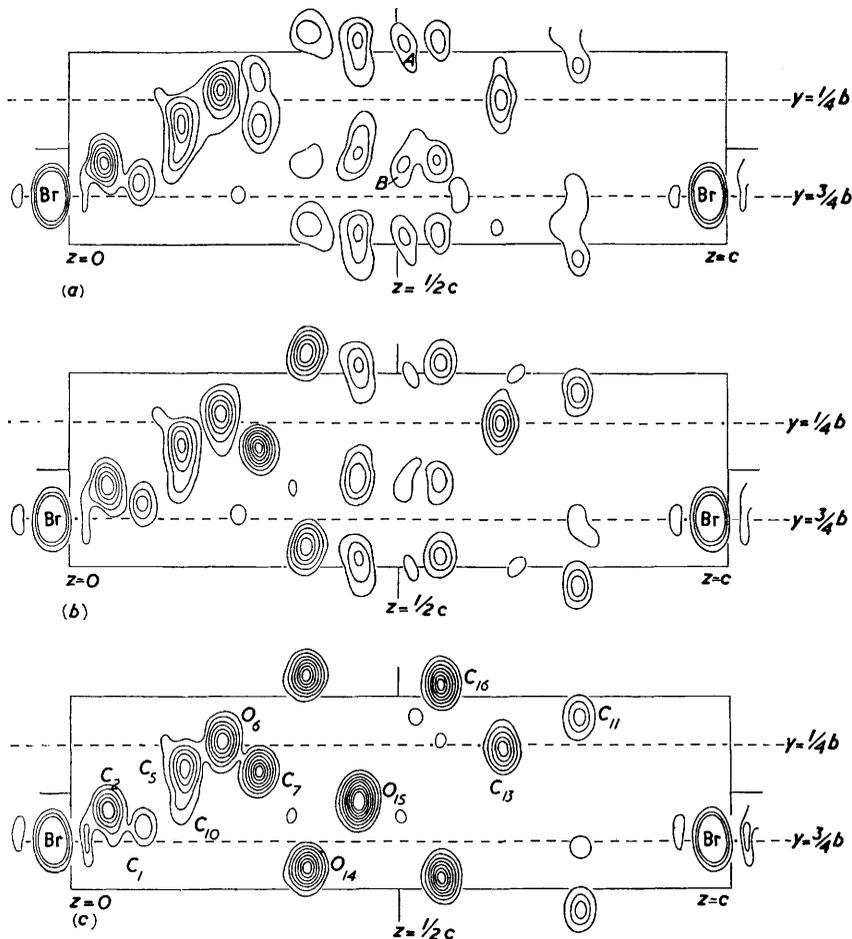
FIG. 3. The $[F_o(\text{hol}) - F_c(\text{hol})]$ synthesis. All light atoms were weighted as carbon atoms in the calculation of $F_c(\text{hol})$. Zero and positive contours full line, negative contours broken line; contour interval $0.5e/\text{\AA}^2$ except in the region of the bromine atom where it is $1e/\text{\AA}^2$.

FIG. 4. A representation of the electron-density distribution, ρ_3 , over the region of one molecule viewed along [010]. The contours are taken from the calculated section nearest to each atomic centre. The contour interval is $1e/\text{\AA}^2$ with the first contour drawn at $1e/\text{\AA}^2$.



maximum. An examination of the interatomic distances defined by these positions showed that these were nearer expected values in the region of the first nine atoms placed than among those selected later. This seemed very reasonable since atomic positions in the first group introduced had been effectively submitted to two refining cycles, in the second group to one, and in the third group to none. A preferred list of co-ordinates shown

FIG. 7. The electron-density distribution in the section at $x = 25/30a$ in ρ_1 (a), ρ_2 (b), and ρ_3 (c). The contour interval is $1e/\text{\AA}^3$, starting with the $1e/\text{\AA}^3$ contour, except over the bromine atom where the contours are at 5, 10, and $15e/\text{\AA}^3$. The false mirror planes at $y = \frac{1}{4}b$ and $y = \frac{3}{4}b$ are indicated by broken lines.



in Table 4 was accordingly drawn up; here the positions of the first nine atoms introduced are those found in ρ_3 while the remaining positions are those of ρ_3 modified by an n shift where $n_x = n_z = 1.3$ and $n_y = 1.5$ for atoms of the second group, $n_x = n_z = 1.6$, $n_y = 2.0$ for those of the third group. A view of the arrangement of the atoms in the molecule which follows from these positions is shown in Fig. 5; Fig. 6 records the bond lengths involved. The latter are, in general, in good agreement with accepted values but show individual deviations of up to 0.17 \AA which are indications only of the present state of approximation of the structure analysis. It is noticeable that all the larger bond-length deviations, four of $>0.15 \text{ \AA}$, involve one atom near the original mirror plane away from which movement is likely to be slowed down by mirror image effects.

The course of the analysis as a whole is illustrated by Fig. 7a, b, and c. In this Figure a section in the three-dimensional electron-density distribution at $x = 25/30a$ is shown

for ρ_1 , ρ_2 , and ρ_3 . This section passes at or near the centres of ten atoms, and illustrates the way in which the image peaks vanish as the corresponding atoms are included in the phase-angle calculations. Thus in ρ_1 , image peaks occur for all the atoms not included in the phasing calculations (C 7, O 14, O 15, C 16, and C 11) except C 13 which lies close to the false mirror plane. The peaks *A* and *B* are two small spurious peaks (see p. 3688). In ρ_2 , the only atoms omitted were O 15 and C 16, and two image peaks occur of which the real peak is the higher. In ρ_3 there are no image peaks.

Fig. 7 shows that in ρ_3 the electron-density peaks are drawn out parallel with the *b* axis. This is almost certainly a consequence of the disordered character of the crystals studied. The observed peaks represent the superposition of the atoms of two molecules which differ a little in their *y* parameters. The deviation between the mean structure postulated in the structure-factor calculations preceding ρ_3 and the true structure is

TABLE 3. Comparison of observed and calculated structure amplitudes.

Plane	F_o	F_c	Plane	F_o	F_c	Plane	F_o	F_c	Plane	F_o	F_c	Plane	F_o	F_c
100	55	+64	2,0,24	16	+11	$\bar{5},0,22$	15	+15	$\bar{5},0,10$	22	-26	$\bar{5},0,12$	22	-17
101	9	+16	2,0,25	18	+13	$\bar{5},0,23$	11	+10	$\bar{5},0,11$	10	-10	$\bar{5},0,13$	10	-13
102	8	+9				$\bar{5},0,24$	13	+15	$\bar{5},0,12$	10	-13	$\bar{5},0,14$	10	-8
103	14	+11	$\bar{5},0,1$	127	-131				$\bar{5},0,13$	13	-17	$\bar{5},0,15$	8	+10
104	10	+3	$\bar{5},0,2$	21	+24	400	9	+1	$\bar{5},0,15$	14	-20	$\bar{5},0,16$	6	+16
105	30	-24	$\bar{5},0,3$	17	-23	401	7	-11	$\bar{5},0,16$	10	-13			
106	49	-47	$\bar{5},0,4$	14	-23	403	14	+13	$\bar{5},0,17$	16	-16			
107	59	-51	$\bar{5},0,5$	9	+3	404	34	+36	$\bar{5},0,19$	12	-16	$\bar{5},0,2$	8	-12
108	30	-32	$\bar{5},0,6$	17	+19	405	23	+23				$\bar{5},0,4$	10	-15
109	19	-13	$\bar{5},0,7$	21	+23	407	9	+7				$\bar{5},0,7$	6	-13
1,0,10	71	-68	$\bar{5},0,8$	26	+25	408	17	+18	600	31	+42			
1,0,11	43	-37	$\bar{5},0,9$	40	+34	409	35	+34	602	9	+8			
1,0,12	37	-36	$\bar{5},0,10$	50	+45	4,0,10	21	+25	603	15	+16	001	13	+22
1,0,13	57	-51	$\bar{5},0,11$	44	+41	4,0,11	9	+4	604	13	-12	002	36	+39
1,0,15	70	-63	$\bar{5},0,12$	23	+23	4,0,12	19	+12	605	18	+22	003	61	+66
1,0,16	34	-29	$\bar{5},0,13$	33	+30	4,0,13	25	+23	606	12	-9	004	29	+26
1,0,17	11	-11	$\bar{5},0,14$	20	+16	4,0,15	19	+20	609	18	-18	005	88	+93
1,0,23	12	+8	$\bar{5},0,15$	28	+24	4,0,17	11	+5	6,0,10	16	-18	006	16	+7
1,0,25	14	+10	$\bar{5},0,16$	26	+25	4,0,23	14	-13	6,0,11	11	-8	007	52	+29
			$\bar{5},0,17$	23	+22				6,0,12	15	-15	008	18	-14
			$\bar{5},0,19$	17	+19	$\bar{5},0,1$	7	-4	6,0,13	14	-14	009	39	+28
T01	38	+31	$\bar{5},0,20$	10	+8	402	24	-27	6,0,14	10	-8	0,0,10	11	-6
T02	58	-62	$\bar{5},0,25$	7	-5	403	35	-45	6,0,15	14	-13	0,0,11	16	-10
T03	31	+33				404	24	-23	6,0,16	13	-13	0,0,12	11	-8
T05	32	+32				405	7	-16	6,0,18	10	-8	0,0,13	26	-24
T06	50	+52	300	51	-51	406	33	-39	6,0,19	16	-16	0,0,14	8	-1
T07	38	+37	301	10	+1	407	38	-41				0,0,15	36	-30
T08	36	+34	302	46	-46	408	14	-21				0,0,16	40	-34
T09	26	+23	303	20	-20	409	37	-42	$\bar{5},0,2$	27	+32	0,0,17	23	+22
1,0,10	33	+28	304	24	-24	4,0,10	18	-22	603	18	+28	0,0,18	11	-10
1,0,11	34	+37	305	26	-26	4,0,11	45	-53	604	23	+28	0,0,19	29	-22
1,0,12	16	+14	306	7	+2	4,0,12	23	-30	605	11	+16	0,0,20	15	-10
1,0,13	20	+17	308	22	+16	4,0,13	9	-10	606	13	+19	0,0,21	25	-17
1,0,14	16	+15	3,0,10	40	+36	4,0,14	11	-14	607	17	+21	0,0,22	10	-7
1,0,15	10	-12	3,0,11	19	+16	4,0,16	20	-24	608	10	-6	0,0,23	12	-7
1,0,16	9	+9	3,0,12	40	+40	4,0,17	9	+6	609	20	+25	0,0,24	9	-7
1,0,19	27	-24	3,0,13	17	+15	4,0,22	13	+11	6,0,12	14	+17	0,0,25	12	-5
1,0,20	20	-15	3,0,14	36	+31	4,0,24	7	+10	6,0,13	11	-11	0,0,26	4	-4
1,0,21	16	-12	3,0,15	42	+42				6,0,18	12	-13			
1,0,22	24	-18	3,0,16	18	+15									
1,0,23	17	-13	3,0,17	28	+24	500	20	+24				011	68	81
1,0,24	17	-13	3,0,18	10	+5	501	23	+28	701	15	-19	012	54	66
1,0,25	11	-9	3,0,19	21	+12	502	17	+20	703	14	-19	013	12	21
1,0,26	15	-17	3,0,20	18	+14	503	33	+36	704	12	-17	014	43	39
			3,0,21	13	+10	504	45	+51	705	10	-14	015	72	77
			3,0,22	11	+7	505	21	+27	706	13	-18	016	44	32
200	73	-77				506	20	+22	707	12	-11	017	57	56
201	7	-10				508	23	+26	708	14	-17	018	39	34
202	49	-48	$\bar{5},0,1$	34	-39	509	15	+17	709	11	-13	019	58	52
203	31	-27	$\bar{5},0,2$	46	-43	5,0,10	9	-3	7,0,10	9	-13	0,1,10	47	38
204	68	-70	$\bar{5},0,3$	15	-11	5,0,12	9	+7	7,0,11	11	-10	0,1,11	44	38
205	67	-60	$\bar{5},0,4$	46	-44	5,0,15	13	-13	7,0,12	10	-14	0,1,12	26	20
206	44	-44	$\bar{5},0,5$	13	-16	5,0,16	17	-12	7,0,13	16	-19	0,1,13	26	21
207	23	-18	$\bar{5},0,6$	21	-20	5,0,17	12	-8				0,1,14	30	23
208	7	-5	$\bar{5},0,7$	19	-18	5,0,19	18	-14				0,1,15	9	4
209	52	-47	$\bar{5},0,8$	11	-13	5,0,20	15	-15	703	10	+13	0,1,16	31	24
2,0,10	29	-27	$\bar{5},0,9$	10	-8				707	11	+14	0,1,17	15	12
2,0,14	9	-8	$\bar{5},0,10$	18	+19				709	12	+14	0,1,18	14	5
2,0,15	17	+15	$\bar{5},0,11$	19	-18	501	25	+29	7,0,11	11	+14	0,1,19	17	13
2,0,16	27	-25	$\bar{5},0,12$	9	+10	502	8	+8	7,0,13	11	+15	0,1,20	10	7
2,0,17	9	+2	$\bar{5},0,15$	10	+14	503	19	+23	7,0,14	12	+15	0,1,21	9	8
2,0,18	17	+12	$\bar{5},0,16$	17	+20	504	17	+20				0,1,22	12	6
2,0,19	11	+9	$\bar{5},0,17$	23	+25	505	24	+26				0,1,23	18	15
2,0,20	25	+18	$\bar{5},0,18$	9	+13	506	8	+8	800	12	-15	0,1,24	16	10
2,0,21	12	+9	$\bar{5},0,19$	20	+21	507	14	-15	802	15	-19	0,1,25	11	7
2,0,22	22	+17	$\bar{5},0,20$	14	+14	508	8	-7	804	9	-11	0,1,26	2	7
2,0,23	10	+7	$\bar{5},0,21$	10	+8	509	11	+8	807	8	-9			

In the structure-factor calculation, temperature-factor-corrected scattering factors were used for carbon and oxygen based on the calculations of Berghuis, Haanappel, Potters, Loopstra, Macgillivray, and Veenendaal (*Acta Cryst.*, 1955, 8, 478). For bromine the scattering curve given in the International Tables, Vol. II, 1955, was used, with a further correction of -1.55 electrons for anomalous dispersion in the *K*-shell.

reflected in the character of the agreement between the observed and calculated F 's (cf. Table 3). The overall reliability index for the 1150 observed spectra was 19.8%, but the values for the individual layers listed in Table 2 show that R increases markedly with h . At the same time the average intensity of the X-ray reflections falls with increasing h , as shown in the section of the structure-factor calculations given in Table 3. The individual y -parameters of the atoms can only be considered to be known with rather low accuracy at present. It is clearly unlikely that they can be known with any great precision unless the ordered crystal is examined.

DISCUSSION

The positions of the atoms listed in Table 4 define the skeleton structure of miroestrol as consisting of four six-membered rings and one five-membered ring. As described above, these positions have not been further refined and the bond lengths and bond angles they indicate are subject to quite large limits of error, particularly when they have a major component in the y direction. However, as they stand, they appear to define the chemical structure of bromomiroestrol as represented by $C_{20}H_{21}O_6Br$, with formula (III).

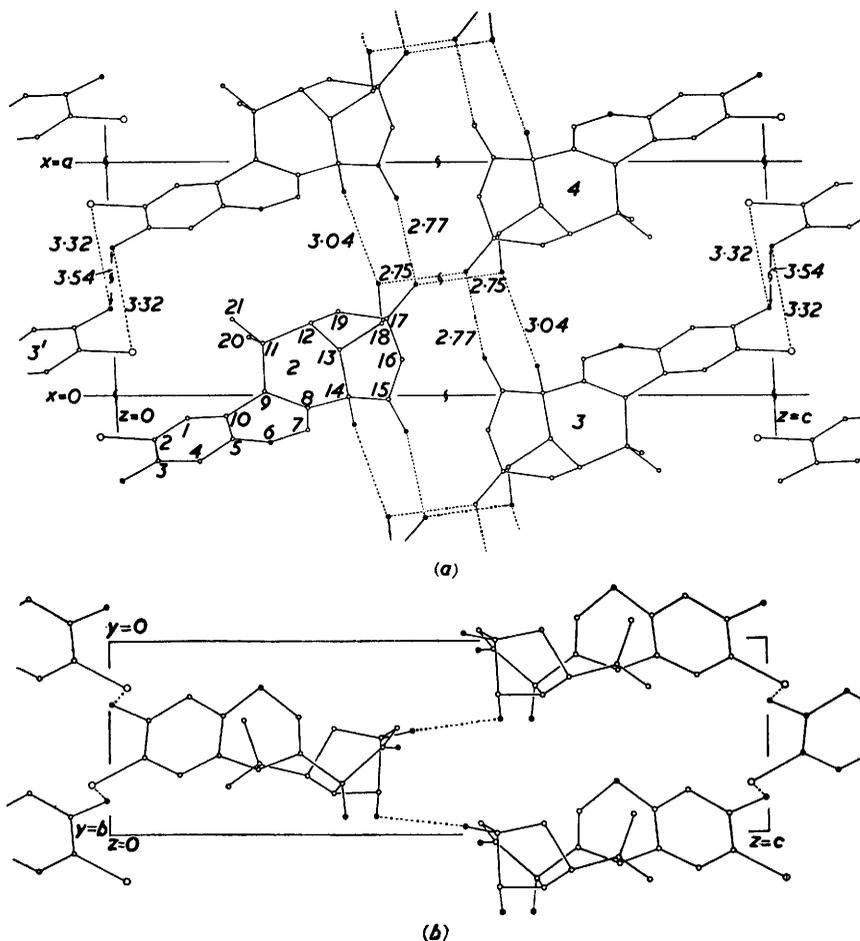
TABLE 4. Atomic co-ordinates.

Atom	x	y	z	Group	Atom	x	y	z	Group
Br	0.8147	0.7535	0.9725	1	C 9	0.0112	0.8448	0.2265	1
O 3	0.6312	0.1737	0.0035	1	C 10	0.9051	0.9155	0.1659	1
O 6	0.7961	0.2641	0.2322	1	C 11	0.2272	0.8700	0.2267	2
O 14	0.8712	0.6013	0.3595	2	C 12	0.3135	0.8099	0.3005	2
O 15	0.8353	0.9525	0.4348	3	C 13	0.1913	0.7220	0.3431	2
O 17	0.4663	0.0373	0.4612	3	C 14	0.9945	0.7709	0.3525	2
O 18	0.4765	0.5987	0.4053	3	C 15	0.9839	0.9406	0.4130	2
C 1	0.8990	0.8127	0.1066	1	C 16	0.1502	0.0510	0.4376	3
C 2	0.8041	0.9045	0.0547	1	C 17	0.3293	0.9556	0.4156	3
C 3	0.7160	0.0976	0.0603	1	C 18	0.3084	0.7297	0.4092	2
C 4	0.7136	0.2193	0.1222	1	C 19	0.3588	0.0521	0.3423	3
C 5	0.8082	0.1197	0.1738	1	C 20	0.3255	0.7546	0.1807	2
C 7	0.8470	0.1082	0.2876	2	C 21	0.2549	0.0936	0.2056	2
C 8	0.9435	0.9231	0.2897	2					

In outline, the evidence for this structure is as follows. Ring A, to which the bromine is attached, is planar and evidently aromatic. The atom attached at position 3, identified as oxygen, must accordingly be formulated as belonging to a phenolic hydroxyl group, the presence of which is supported by chemical evidence. Ring B too is essentially planar, but atom 6 is identified as oxygen and the ring is therefore heterocyclic. Also, of the substituent atoms, C 14 lies closely in the plane of ring B, C 11 markedly out of it so that the angles 11, 9, 10 and 11, 9, 8 are nearly tetrahedral. These features lead to the deduction that a double bond should be placed between C 7 and C 8, leaving C 9 fully substituted. The atoms 20 and 21 attached at C 11 are both carbon, and the group has the form of a *gem*-dimethyl group. The identification is supported by the existence of small peaks near the expected sites of hydrogen atoms attached to C 20 and C 21 and arranged in the most probable staggered conformation. Ring C, from its shape and substituents, is fully reduced; ring D is also largely reduced; the planar configuration of atoms 14, 15, and 16 and of the oxygen atom attached at C 15 is consistent with the formulation of this atom as ketonic oxygen. The existence of the five-membered ring, ring E, was not suspected until late in the analysis. As atomic positions were selected from the electron-density peaks in this region in ρ_2 , it was realised that the peaks representing C 19 and C 12 were 1.98 Å apart, intermediate between a bonded (1.54 Å) and a non-bonded distance (2.5 Å or more). In the first case, the structure would be as represented, in the second, atom 19 would presumably be formulated as a methyl group attached to ring D, corresponding with the alternative possible miroestrol formula, $C_{20}H_{24}O_6$. To test these alternatives, the atoms were placed for the structure-factor calculation, set 3, 1.86 Å apart

at positions slightly displaced from the peak maxima of ρ_2 . It was expected that, if there were a bond between C 19 and C 12, the peaks representing these atoms would move together and this proved to be the case. Their separation is 1.75 Å in ρ_3 , and drops to 1.70 Å on applying the n shifts as in Table 4 co-ordinates; though this value is still a little high, there is no doubt that C 12 and C 19 are bonded. It is presumably the existence of ring E which forces a markedly boat-shaped conformation on ring D and the slightly

FIG. 8. The arrangement of the molecules in the crystal structure of bromomiræstrool projected (a) on (010), (b) on (100).

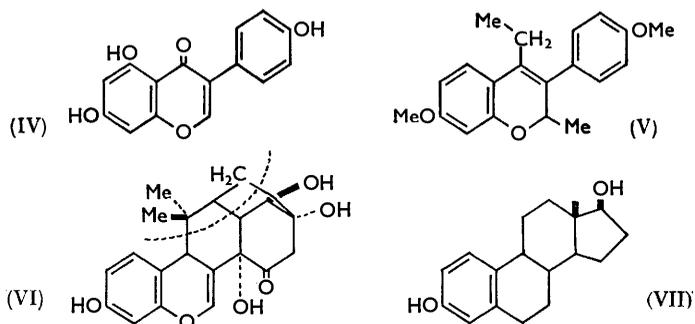


boat-shaped conformation on ring c. The substituent atoms at C 14, C 17, and C 18 are all identified as oxygen; from their geometry they belong to hydroxyl groups, three or more of which are indicated by the chemical evidence.

The arrangement of the molecules in the crystal provides some confirmation for the character of the different active groups present. A series of short contacts, corresponding in length with hydrogen bonds, can be traced running between the 15-keto-group and the 17-hydroxyl group (2.77), then OH 17-OH 18 (2.75 Å) and OH 18-OH 14 (perhaps a little long) (3.04 Å), as shown in Fig. 8. Small peaks occur in these bonds near the suspected positions of the hydroxyl-hydrogen atoms. At the opposite end of the molecules, however, the distance apart of the phenolic hydroxyl groups of neighbouring molecules, required by the parameters of Table 4, places these groups 3.54 Å apart, too far for any

hydrogen bond to be formed between them. Their distance from neighbouring bromine atoms is rather shorter (3.32 Å) and could correspond with weak attraction. But one can see that the two hydroxyl groups might themselves be brought into hydrogen-bonded contact by a bodily translation of alternate molecules, $\frac{1}{2}$ — $\frac{3}{4}$ Å parallel with *b*. This would modify the character of the contacts at the opposite ends of each molecule; the distance between OH 17 groups in molecules 2 and 4 of Fig. 8 would decrease and could become the shortest in the structure. It seems possible that it is this situation, that there are alternative ways of making hydrogen-bonded contacts between the molecules, that is responsible for the disorder in crystals of type I and the consequent elongated peaks of ρ_3 . Clearly the preferred hydrogen-bonded arrangement could be found by a detailed analysis of the ordered crystals of type II.

Biogenetically, it seems possible to dissect the miræstrol nucleus into an isoprene unit consisting of CH₃ 20, CH₃ 21, C 11, C 12, and C 19 and a partly reduced isoflaven nucleus (VI). This dissection relates it, on the one hand, to the many plant products in which isoprene units are found linked in various ways to condensed ring systems (cf. Robinson⁵) and, on the other, to isoflavones such as genistein⁶ (IV), and isoflavenes, such as 4-ethyl-7,4'-dimethoxy-2-methylisoflav-3-en⁷ (V), which have œstrogenic properties. Both in these compounds and in miræstrol (VI) the biological potency is clearly not significant



of any structural relation of the molecular skeleton present to the steroids. It seems to be a consequence of an accidental geometrical similarity in the positions of the hydroxyl groups, or potential hydroxyl groups, at the two ends of each molecule. In miræstrol, the actual distance given by the co-ordinates in Table 4 between the 3-OH and the 18-OH is 10.57 Å, closely the distance to be expected from models of the 3-OH to the 17-OH of œstradiol. The disposition of the second terminal hydroxyl group at C 17 of miræstrol (10.91 Å from the 3-OH group) is rather more different in its relative position from the second, 16-OH group of œstriol; it is worth observing that miræstrol is the more potent œstrogen.

The calculations necessary to establish the structure of miræstrol were carried out partly on punched-card machines supplied by British Tabulating Machines Ltd., and partly on Deuce at the National Physical Laboratory with the help of the staff of the Mathematics Division. The research was assisted by a grant from the National Research and Development Corporation. N. E. T. is indebted to the Commonwealth Scientific and Industrial Research Organisation for the award of an overseas studentship, and J. S. R. thanks the Pressed Steel Company for a research fellowship.

CHEMICAL CRYSTALLOGRAPHY LABORATORY,
 OXFORD UNIVERSITY.

[Received, November 30th, 1959.]

⁵ Robinson, "Structural Relations of Natural Products," Oxford Univ. Press, 1953, p. 42.

⁶ Bradbury and White, *J.*, 1951, 3447; 1953, 871.

⁷ Lawson, *J.*, 1954, 4448.