Crystal and Molecular Structure of 17β-Hydroxyandrost-4-en-3-one (Testosterone)

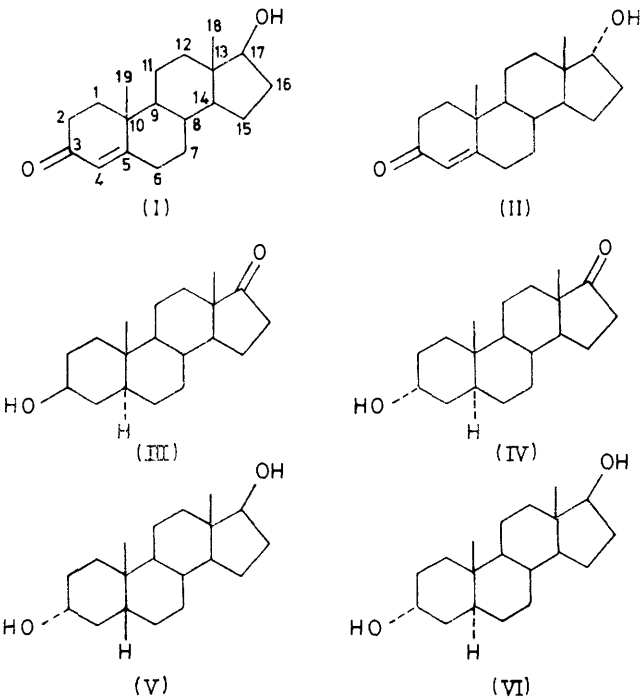
By **Paul J. Roberts, Roger C. Pettersen, George M. Sheldrick, Neil W. Isaacs,** and **Olga Kennard,*†**
University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW

Anhydrous testosterone is monoclinic, space group $P2_1$, a = 14·720(3), b = 11·080(2), c = 10·868(2) Å, β = 113·34(1)°, Z = 4. The structure was solved from diffractometer data by direct methods and refined to R 0·038 for 2498 observed (of a total of 2749) reflexions. The two independent molecules in the asymmetric unit have significantly different conformations. This finding is discussed in conjunction with results from studies on epitestosterone and two hydrated forms of testosterone in terms of the relationships between molecular structure and biological activity.

TESTOSTERONE (I) is the principal hormone of the testes, produced by the interstitial cells. It acts as a powerful androgen and is responsible for the male sex characteristics. It differs from epitestosterone [1] (II) only in the orientation of the 17-hydroxy-group but has about 25 times its biological activity. The preference for the β-orientation of the 17-hydroxy-group is a common feature of the biologically active estrogens and androgens [2] and we hoped that a comparison of the two structures, analysed with the same accuracy, would indicate whether the differences in biological activity are due to subtle changes in molecular structure or to the gross orientation of the C(17) substituent.

When we began work on the two structures, no comparable analysis of a pair of epimeric steroids was available. Since then, Weeks *et al.* have reported the structure of epiandrosterone [3] [5α-androstan-3β-ol-17-one (III)] and compared their results with androsterone [4] (IV). A third example of a pair of steroids with differing biological activities is 5β-androstan-3α,17β-diol (V) [3] and its 5α-epimer (VI).[5] This case is particularly interesting since the molecules differ not only in the orientation of a substituent but in the overall geometry of the steroid skeleton since the A/B ring junction is *cis*- in the 5β-compound but *trans*- in its 5α-epimer.

In this paper we present the crystal and molecular structure of anhydrous testosterone and compare the detailed molecular geometry with two forms of hydrated testosterone.[6,7] In conjunction with this we also present some comparisons between the other two epimeric steroid structures [(III) and (IV), (V) and (VI)] for which X-ray results are available.



(I) (II) (III) (IV) (V) (VI)

EXPERIMENTAL

Crystal Data.—$C_{19}H_{28}O_2$, M = 288. Monoclinic, a = 14·720(3), b = 11·080(2), c = 10·868(2) Å, β = 113·34(1)°, U = 1627·5(2) Å^3, D_m = 1·20 g cm^{-3}, Z = 4, D_x = 1·24 g cm^{-3}, $F(000)$ = 632. Space group $P2_1$ (C_2^2, No. 4) from

† External staff, Medical Research Council.

[1] N. W. Isaacs, W. D. S. Motherwell, J. C. Coppola, and O. Kennard, *J.C.S. Perkin II*. 1972, 2335.
[2] S. G. Korenman, H. Wilson, and M. B. Lipsett, *J. Biol. Chem.*, 1964, **239**, 1004.
[3] C. M. Weeks, A. Cooper, D. A. Norton, H. Hauptman, and J. Fisher, *Acta Cryst.*, 1971, *B*, **27**, 1562.
[4] D. F. High and J. Kraut, *Acta Cryst.*, 1966, **21**, 88.
[5] G. Precigoux, B. Busetta, C. Courseille, and M. Hospital, *Cryst. Struct. Comm.*, 1972, **1**, 265.
[6] B. Busetta, C. Courseille, F. Leroy, and M. Hospital, *Acta Cryst.*, 1972, *B*, **28**, 3293.
[7] M. Hospital, personal communication.

systematic absences. Cu-K_α radiation, $\lambda = 1{\cdot}5418$ Å; μ(Cu-K_α) = 5·8 cm^{-1}.

Commercial testosterone (Koch–Light Laboratories) was recrystallized from ethanol–ethyl acetate as long needles. Unit cell dimensions and space group were determined photographically. Cell parameters were refined by least-squares analysis of θ values measured for 27 reflexions on a Picker diffractometer by use of graphite-monochromatized Cu-K_α radiation. Intensities were measured from a crystal of dimensions 0·10 × 0·10 × 0·05 mm. The diffractometer was operated in the θ—2θ scan mode with a range in 2θ of

TABLE 1

Reflexions used in starting set with allowed phases (rad)

−3	0	5	0
−11	1	7	0
1	0	8	0
−9	2	7	π/4, 3π/4
0	2	1	±π/4, ±3π/4
1	3	7	±π/4, ±3π/4
−10	0	8	0, π
−10	0	7	0, π

(2·0 + 0·260 tan θ)° and a speed of 1° min^{-1}. The standard deviation of an intensity was calculated from counting statistics using $\sigma^2(I) = S + B + (dS)^2$ where S = scan count, B = background, corrected to scan time, $I = S - B$, d = an empirical constant to allow for instrumental errors, calculated to be 0·019 from monitor reflexions. Of 2749 independent reflexions with 2θ(Cu-K_α) ⩽ 127° (minimum interplanar spacing 0·86 Å), 251 had $I/\sigma(I) < 2{\cdot}98$ and were classified unobserved. No absorption correction was made. Lorentz and polarization factors were applied and the structure amplitudes and normalized structure amplitudes (E values) were derived.

The structure was solved by weighted multisolution tangent refinement,[8] for 585 reflections with $|E| > 1{\cdot}2$. Difficulties were experienced in selecting a suitable starting set of reflexions; these were eventually resolved by finding the set which gave the most rapid phase expansion in a pseudotangent refinement, *i.e.* a weighted tangent refinement in which the weights w were based on estimated α values, rather than on summations involving numerical phases: $w = \tanh[\langle\alpha^2\rangle^{\frac{1}{2}}/2]$, where $\langle\alpha^2\rangle$ was calculated as in ref. 8, with the substitution of $w_{h'}w_{h-h'}\kappa_{h,h'}$ for $\kappa_{h,h'}$. This led to the starting set of eight reflexions given in Table 1. The first three reflexions were used to define the origin and the fourth to fix the enantiomorph.

An index $R_\alpha = [\Sigma_h g(\alpha - k\langle\alpha^2\rangle^{\frac{1}{2}})^2/\Sigma_h g\alpha^2]^{\frac{1}{2}}$ was calculated after 12 cycles of weighted tangent refinement for each of the 128 phase permutations. The weights[8] were $w = \tanh[\alpha/2]$, g was set at $\langle\alpha^2\rangle^{-\frac{1}{2}}$, and the scale factor k was chosen to minimise R_α. We have found that the correct solution normally possesses the lowest value of R_α for space groups containing translational symmetry operators. E Maps were calculated for the solutions having low values of R_α. The resulting peak co-ordinates were used to calculate E values (assuming atomic numbers proportional to the square roots of the peak heights) and a 'reliability index' $R_A = [\Sigma(E_{obs} - kE_{calc})^2/\Sigma(E_{obs})^2]^{\frac{1}{2}}$, where k was chosen to minimise R_A. The best solution had R_α 0·197 and R_A 0·218, and the highest 42 independent peaks could be assembled to give two complete molecules of testosterone. In fact the lowest 14 R_α values (<0·217) all corresponded to substantially corrrect E maps, differing mainly in the distribution of peak heights; the corresponding R_A values were in the range 0·218—0·375.

Subsequent analysis indicated that a very acceptable solution (R_α 0·234, R_A 0·254) could have been obtained from only 32 phase permutations (by omitting 0 2 1 from the starting set), and that the optimum minimum E and κ values for the tangent refinement were *ca.* 1·2 and 1·0 respectively.

Two cycles of full-matrix least-squares refinement treating 42 carbon and oxygen atoms isotropically reduced R to 0·123. The function minimised was $\Sigma w|F_o - kF_c|^2$ with weights given by $w = \{\sigma(F_o)[2F_{min.} + F_o + 2F_o^2/F_{max.}]\}^{-1}$.

TABLE 2

Analysis of variance

(*a*) As a function of sinθ

sinθ	0·00—0·39	0·50—0·58	0·64—0·69	0·74—0·78	0·82—0·86	0·89
N	254	251	266	244	252	278	222	267	266	197
V	52	41	30	30	37	35	36	40	43	43

(*b*) As a function of $\sqrt{(F/F_{max.})}$

$\sqrt{(F/F_{max.})}$	0·00—0·16	0·18	0·20	0·22	0·24	0·27	0·30	0·34	0·41	1·00
N	340	238	228	245	233	278	223	223	243	246
V	51	41	35	35	34	39	31	31	31	49

(*c*) By zones and parity groups

	0*kl*	*h*0*l*	*hk*0	All	*ggg*	*ugg*	*gug*	*uug*	*ggu*	*ugu*	*guu*	*uuu*
N	116	254	154	2497	330	312	299	300	337	326	292	301
V	40	49	45	39	38	37	39	35	37	46	39	41

Scattering factors were from ref. 9. The positions of 54 hydrogen atoms (hydroxy-hydrogens excluded) were calculated and added to a structure-factor calculation (R 0·096). Because of the larger number (602) of variables, calculations were continued by refining the two molecules alternately by anisotropic full-matrix least-squares for two cycles each with the hydrogen positions fixed. Two cycles of isotropic refinement were next applied to the 54 hydrogen atoms giving R 0·043 and R_W 0·045 [$R_W = \Sigma w^{\frac{1}{2}}|F_o - F_c|/\Sigma w^{\frac{1}{2}}|F_o|$]. The remaining two hydrogen atoms of the hydroxy-groups were located at this stage from a difference-Fourier map. The calculations were completed with two cycles of isotropic refinement of all 56 hydrogen atoms, followed by two cycles of anisotropic refinement of the carbon and oxygen atoms. Parameter shifts in the last cycles were all small fractions of their standard deviations. Final values of R and R_W were 0·038 and 0·040 respectively. Table 2 shows an analysis of variance computed at this stage. Observed and

[8] G. Germain, P. Main, and M. M. Woolfson, *Acta Cryst.*, 1970, *B*, **26**, 274.

[9] D. T. Cromer and J. Mann, *Acta Cryst.*, 1968, *A*, **24**, 321.

TABLE 3

Final fractional co-ordinates ($\times 10^3$ for hydrogen, 10^4 for others) and isotropic temperature factors ($Å^2 \times 10^2$) for hydrogen

	Molecule A					Molecule B			
	x/a	*y/b*	*z/c*	*U*		*x/a*	*y/b*	*z/c*	*U*
C(1)	1974(2)	7201	9348(3)		C(101)	474(2)	774(3)	8731(3)	
C(2)	1913(2)	6967(3)	10695(3)		C(102)	518(2)	537(3)	10127(3)	
C(3)	2252(2)	5727(3)	11187(3)		C(103)	1521(2)	166(3)	11077(3)	
C(4)	3040(2)	5227(3)	10851(2)		C(104)	2351(2)	629(3)	10812(2)	
C(5)	3409(2)	5786(3)	10050(2)		C(105)	2253(2)	1354(2)	9774(2)	
C(6)	4289(2)	5260(3)	9864(3)		C(106)	3147(2)	1921(3)	9687(3)	
C(7)	4162(2)	5273(3)	8400(3)		C(107)	3161(2)	1795(3)	8295(3)	
C(8)	3911(2)	6530(2)	7805(2)		C(108)	2211(2)	2286(2)	7227(2)	
C(9)	2953(2)	6970(2)	7907(2)		C(109)	1304(2)	1647(2)	7297(2)	
C(10)	3016(2)	6984(2)	9383(2)		C(110)	1250(2)	1687(2)	8695(2)	
C(11)	2580(2)	8175(3)	7178(3)		C(111)	335(2)	2074(3)	6154(2)	
C(12)	2524(2)	8188(3)	5735(3)		C(112)	368(2)	1981(3)	4756(3)	
C(13)	3509(2)	7819(3)	5696(3)		C(113)	1258(2)	2649(3)	4719(3)	
C(14)	3778(2)	6567(3)	6338(2)		C(114)	2193(2)	2125(3)	5834(3)	
C(15)	4624(2)	6143(4)	5953(3)		C(115)	3040(2)	2585(4)	5495(3)	
C(16)	4370(3)	6715(4)	4562(3)		C(116)	2612(3)	2585(4)	3955(4)	
C(17)	3480(2)	7534(3)	4299(3)		C(117)	1497(3)	2351(3)	3486(3)	
C(18)	4302(3)	8782(4)	6368(3)		C(118)	1157(2)	4008(3)	4812(3)	
C(19)	3735(2)	7977(3)	10219(3)		C(119)	992(2)	2964(3)	9025(3)	
O(3)	1930(2)	5182(3)	11906(2)		O(103)	1646(2)	−447(2)	12070(2)	
O(17)	3531(2)	8562(3)	3524(2)		O(117)	909(2)	2979(3)	2292(2)	
H(1)	183(2)	808(4)	913(3)	3(1)	H(101)	−23(3)	105(4)	808(3)	2(1)
H(2)	149(2)	665(3)	860(3)	2(1)	H(102)	64(3)	5(4)	830(4)	3(1)
H(3)	125(3)	710(4)	1070(4)	6(1)	H(103)	7(3)	−9(4)	1014(4)	3(1)
H(4)	236(3)	752(4)	1138(4)	3(1)	H(104)	31(3)	126(4)	1061(3)	4(1)
H(5)	336(2)	446(4)	1125(3)	2(1)	H(105)	301(2)	47(3)	1148(3)	2(1)
H(6)	447(3)	444(4)	1029(4)	3(1)	H(106)	377(3)	156(4)	1037(3)	3(1)
H(7)	489(2)	578(4)	1040(3)	3(1)	H(107)	309(3)	281(4)	984(4)	4(1)
H(8)	363(2)	472(3)	788(3)	1(1)	H(108)	324(2)	95(4)	809(3)	3(1)
H(9)	478(2)	499(4)	832(3)	2(1)	H(109)	378(2)	225(3)	827(3)	2(1)
H(10)	453(2)	710(4)	831(3)	2(1)	H(110)	215(2)	316(3)	741(3)	2(1)
H(11)	248(2)	633(3)	744(3)	1(1)	H(111)	139(2)	77(4)	716(3)	3(1)
H(12)	315(3)	878(4)	778(4)	4(1)	H(112)	22(2)	299(3)	635(3)	3(1)
H(13)	192(3)	838(4)	722(3)	2(1)	H(113)	−24(2)	161(3)	615(3)	2(1)
H(14)	227(3)	899(4)	532(4)	5(1)	H(114)	−29(3)	231(4)	405(4)	4(1)
H(15)	198(2)	759(3)	512(3)	2(1)	H(115)	41(3)	109(4)	454(4)	3(1)
H(16)	318(2)	598(3)	585(3)	2(1)	H(116)	219(2)	125(3)	570(3)	2(1)
H(17)	468(3)	520(4)	593(4)	5(1)	H(117)	365(3)	209(4)	591(3)	3(1)
H(18)	527(3)	647(4)	655(3)	3(1)	H(118)	328(3)	335(4)	592(4)	4(1)
H(19)	423(4)	614(5)	378(5)	9(2)	H(119)	294(3)	200(5)	357(4)	9(1)
H(20)	493(3)	728(5)	457(4)	6(1)	H(120)	277(3)	337(4)	360(4)	6(1)
H(21)	287(2)	706(3)	383(3)	3(1)	H(121)	132(2)	144(4)	329(3)	4(1)
H(22)	445(3)	892(4)	733(4)	3(1)	H(122)	112(3)	425(4)	571(4)	6(1)
H(23)	499(3)	851(4)	639(4)	4(1)	H(123)	174(3)	441(4)	471(4)	4(1)
H(24)	407(3)	954(4)	594(4)	6(1)	H(124)	49(3)	430(5)	398(4)	6(1)
H(25)	388(3)	789(5)	1127(4)	5(1)	H(125)	96(3)	298(4)	992(4)	5(1)
H(26)	441(3)	789(4)	1024(4)	4(1)	H(126)	141(3)	360(4)	892(4)	4(1)
H(27)	351(3)	875(5)	991(4)	5(1)	H(127)	27(3)	312(4)	840(4)	3(1)
H(28)	280(3)	885(4)	312(3)	8(1)	H(128)	124(2)	377(4)	228(3)	8(1)

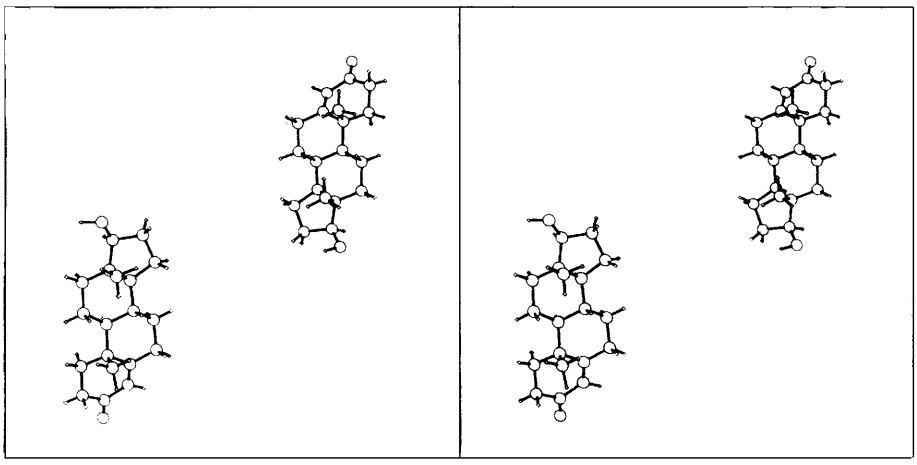


FIGURE 1 General stereoview of the two molecules in the asymmetric unit

calculated structure factors are listed in Supplementary Publication No. SUP 20755 (13 pp., 1 microfiche).* There were no peaks on the final difference Fourier $>0{\cdot}1$ eÅ^{-3}. Final positional and thermal parameters are given in Tables 3 and 4 together with their standard deviations

TABLE 4

Anisotropic temperature factors (Å^2 × 10^3)

	U_{11}	U_{22}	U_{33}	U_{23}	U_{13}	U_{12}
C(1)	45(1)	49(2)	47(1)	4(1)	25(1)	6(1)
C(2)	55(2)	49(2)	57(2)	−5(1)	36(1)	−4(1)
C(3)	56(2)	51(2)	43(1)	−3(1)	24(1)	0(1)
C(4)	56(2)	39(1)	36(1)	0(1)	19(1)	0(1)
C(5)	37(1)	38(1)	33(1)	−2(1)	9(1)	0(1)
C(6)	41(1)	49(2)	43(1)	8(1)	16(1)	9(1)
C(7)	45(1)	43(1)	44(1)	4(1)	17(1)	12(1)
C(8)	33(1)	36(1)	37(1)	0(1)	14(1)	0(1)
C(9)	33(1)	34(1)	33(1)	−1(1)	11(1)	−1(1)
C(10)	33(1)	32(1)	38(1)	−1(1)	12(1)	−3(1)
C(11)	60(2)	44(2)	52(2)	12(1)	27(1)	17(1)
C(12)	53(2)	50(2)	44(1)	11(1)	17(1)	13(1)
C(13)	43(1)	44(1)	38(1)	2(1)	13(1)	−2(1)
C(14)	38(1)	43(2)	38(1)	−1(1)	14(1)	5(1)
C(15)	59(2)	88(3)	58(2)	11(2)	35(1)	24(2)
C(16)	72(2)	105(3)	55(2)	14(2)	36(2)	25(2)
C(17)	55(2)	63(2)	39(1)	5(1)	18(1)	1(1)
C(18)	72(2)	63(2)	47(2)	−1(2)	17(1)	−24(2)
C(19)	50(1)	45(2)	47(2)	−10(1)	19(1)	−13(1)
O(3)	76(1)	73(1)	74(1)	14(1)	47(1)	−7(1)
O(17)	75(1)	86(2)	47(1)	18(1)	26(1)	−5(1)
C(101)	42(1)	52(2)	43(1)	3(1)	16(1)	−7(1)
C(102)	52(2)	60(2)	51(2)	9(1)	22(1)	−4(1)
C(103)	62(2)	40(1)	36(1)	1(1)	21(1)	0(1)
C(104)	48(1)	50(2)	36(1)	3(1)	10(1)	10(1)
C(105)	38(1)	42(1)	32(1)	−4(1)	10(1)	2(1)
C(106)	38(1)	62(2)	43(1)	0(1)	9(1)	−1(1)
C(107)	33(1)	54(2)	51(1)	4(1)	16(1)	2(1)
C(108)	40(1)	37(1)	41(1)	−1(1)	19(1)	2(1)
C(109)	35(1)	39(1)	35(1)	0(1)	14(1)	3(1)
C(110)	33(1)	40(1)	32(1)	1(1)	10(1)	1(1)
C(111)	39(1)	64(2)	37(1)	6(1)	13(1)	0(1)
C(112)	49(1)	68(2)	38(1)	0(1)	12(1)	−8(1)
C(113)	62(2)	41(2)	37(1)	0(1)	22(1)	1(1)
C(114)	49(1)	38(1)	47(1)	−2(1)	26(1)	3(1)
C(115)	63(2)	74(2)	65(2)	5(2)	41(2)	2(2)
C(116)	95(3)	89(3)	65(2)	4(2)	54(2)	6(2)
C(117)	89(2)	57(2)	47(2)	−1(2)	38(2)	0(2)
C(118)	76(2)	43(2)	49(2)	6(1)	23(1)	12(1)
C(119)	58(2)	49(2)	43(1)	−1(1)	24(1)	9(1)
O(103)	85(1)	55(1)	49(1)	16(1)	29(1)	11(1)
O(117)	107(2)	82(2)	39(1)	3(1)	27(1)	−14(1)

Coefficients in the temperature factor expression:

$\exp[-2\pi^2(U_{11}h^2a^{*2} + U_{22}k^2b^{*2} + U_{33}l^2c^{*2} + 2U_{12}hka^*b^* + 2U_{13}hla^*c^* + 2U_{23}klb^*c^*)]$.

calculated from the inverse matrix of the last refinement cycle.

DISCUSSION

The chemical structure and numbering scheme used are shown in (I). Oxygen atoms are assigned the same number as the carbon atoms of the steroid nucleus to which they are bonded. Hydrogen atoms are numbered consecutively, beginning with those bonded to C(1). The same numbering scheme is used for each of the two molecules of the asymmetric unit except that the numbers used for molecule B are increased by adding 100. Figure 1 is a stereo view of the two molecules in the asymmetric unit. Tables 5—7 give the intramolecular geometry in terms of bond lengths, valency angles, and torsion angles for the two molecules of anhydrous testosterone. These Tables also list, for comparison, the values for two forms of hydrated testosterone [6,7] and for epitestosterone.[1]

The following symbols are used in Tables 5—8: T1 and T2 to represent molecules A and B of testosterone, H1 and H2 for the monoclinic and orthorhombic forms of hydrated testosterone, and E for epitestosterone.

The agreement in bond length and valency angles of the two crystallographically independent molecules is

TABLE 5

(*a*) Bond lengths (Å), with standard deviations in parentheses

	T1	T2	H1 *	H2 *	E *
C(1)–C(2)	1·524(5)	1·516(5)	1·538	1·512	1·530
C(1)–C(10)	1·538(4)	1·538(4)	1·546	1·541	1·529
C(2)–C(3)	1·487(5)	1·486(4)	1·492	1·503	1·504
C(3)–C(4)	1·456(5)	1·455(5)	1·451	1·449	1·436
C(3)–O(3)	1·222(5)	1·225(4)	1·229	1·233	1·228
C(4)–C(5)	1·346(4)	1·345(4)	1·343	1·353	1·345
C(5)–C(6)	1·505(5)	1·495(4)	1·493	1·490	1·509
C(5)–C(10)	1·513(4)	1·523(3)	1·518	1·532	1·521
C(6)–C(7)	1·527(5)	1·528(5)	1·535	1·543	1·526
C(7)–C(8)	1·518(4)	1·521(3)	1·525	1·532	1·532
C(8)–C(9)	1·537(4)	1·539(4)	1·535	1·543	1·539
C(8)–C(14)	1·528(3)	1·514(4)	1·527	1·522	1·524
C(9)–C(10)	1·570(3)	1·553(4)	1·561	1·566	1·564
C(9)–C(11)	1·539(4)	1·548(3)	1·531	1·555	1·538
C(10)–C(19)	1·547(4)	1·544(4)	1·572	1·540	1·544
C(11)–C(12)	1·538(5)	1·543(4)	1·551	1·539	1·557
C(12)–C(13)	1·523(5)	1·519(5)	1·529	1·528	1·504
C(13)–C(14)	1·533(4)	1·542(4)	1·533	1·543	1·532
C(13)–C(17)	1·535(5)	1·550(5)	1·532	1·553	1·540
C(13)–C(18)	1·536(5)	1·520(5)	1·545	1·523	1·513
C(14)–C(15)	1·536(5)	1·521(5)	1·543	1·540	1·555
C(15)–C(16)	1·542(5)	1·537(5)	1·546	1·532	1·558
C(16)–C(17)	1·525(5)	1·536(6)	1·532	1·558	1·546
C(17)–O(17)	1·436(5)	1·424(4)	1·439	1·424	1·426

* Calculated from published co-ordinates; for definition of headings see text.

Bond lengths involving hydrogen atoms

	T1	T2		T1	T2
C(1)–H(1)	1·00(4)	1·04(3)	C(12)–H(15)	1·05(3)	1·02(4)
C(1)–H(2)	1·04(3)	1·01(5)	C(14)–H(16)	1·05(3)	0·98(4)
C(2)–H(3)	0·98(5)	0·96(4)	C(15)–H(17)	1·05(5)	0·99(4)
C(2)–H(4)	0·99(4)	1·07(4)	C(15)–H(18)	0·99(3)	0·96(4)
C(4)–H(5)	0·98(4)	0·97(3)	C(16)–H(19)	1·01(5)	0·99(5)
C(6)–H(6)	1·00(4)	1·01(3)	C(16)–H(20)	1·04(5)	1·02(5)
C(6)–H(7)	1·02(3)	1·01(5)	C(17)–H(21)	0·99(3)	1·04(4)
C(7)–H(8)	0·98(3)	0·98(4)	C(18)–H(22)	1·00(4)	1·03(5)
C(7)–H(9)	1·00(4)	1·05(4)	C(18)–H(23)	1·04(4)	1·02(5)
C(8)–H(10)	1·07(3)	1·00(3)	C(18)–H(24)	0·96(4)	1·09(4)
C(9)–H(11)	0·98(3)	1·00(4)	C(19)–H(25)	1·08(5)	0·99(5)
C(11)–H(12)	1·07(4)	1·07(4)	C(19)–H(26)	0·98(4)	0·97(5)
C(11)–H(13)	1·01(4)	0·99(4)	C(19)–H(27)	0·93(6)	1·02(3)
C(12)–H(14)	1·00(4)	1·04(3)	O(17)–H(28)	1·04(4)	1·01(4)

well within experimental error. The average difference in bond length is only 0·008 Å and the maximum is 0·017 Å for the C(9)–C(10) bond. The mean C(sp^3)–C(sp^3) is 1·534 compared with the theoretical value of 1·533 Å°.[10] The mean C(sp^2)–C(sp^3) bond is 1·501 (calc. 1·505), the C(sp^2)–C(sp^2) bond length is 1·346 (calc. 1·336) and the single-bond C(3)–C(4) in the 4-en-3-one system is 1·456 (1·465 ± 0·005 Å, ref. 11). There is a tendency for the C(9)–C(10) bond to be rather longer than the theoretical value in all the compounds listed in Table 5.

* For details see Notice to Authors No. 7 in *J.C.S. Dalton*, 1972, Index issue.

[10] L. S. Bartell and R. A. Bonhan, *J. Chem. Phys.*, 1960, **32**, 824.

[11] *Chem. Soc. Special Publ.*, No. 12, 1958.

An examination of other published steroid structures seems to suggest that this lengthening is related to the presence of a double bond in ring A. The length of the C(9)–C(10) bond in some 11 structures ranges from 1·553—1·570 Å with a mean value of 1·563 Å. In steroids with a fully saturated or fully unsaturated ring A, the C(9)–C(10) bond does not show such a systematic lengthening.

TABLE 6

Valency angles (deg.), with standard deviations in parentheses

	T1	T2	H1*	H2*	E*
C(2)–C(1)–C(10)	112·8(2)	113·7(2)	112·4	114·1	113·4
C(1)–C(2)–C(3)	111·3(2)	112·3(3)	110·4	111·4	109·6
C(2)–C(3)–C(4)	116·6(3)	116·5(2)	116·3	117·4	117·2
C(2)–C(3)–O(3)	121·8(3)	121·8(3)	121·9	120·9	120·5
C(4)–C(3)–O(3)	121·4(3)	121·6(3)	121·6	121·7	122·2
C(3)–C(4)–C(5)	123·6(3)	123·8(2)	123·6	123·6	122·4
C(4)–C(5)–C(6)	119·9(3)	120·0(2)	119·9	120·3	120·4
C(4)–C(5)–C(10)	122·5(3)	122·7(2)	123·1	122·0	123·9
C(6)–C(5)–C(10)	117·5(2)	117·3(2)	116·9	117·6	115·7
C(5)–C(6)–C(7)	112·3(2)	112·5(2)	111·6	111·7	114·1
C(6)–C(7)–C(8)	111·2(2)	110·7(2)	111·4	110·7	112·5
C(7)–C(8)–C(9)	109·6(2)	110·6(2)	109·8	109·7	110·6
C(7)–C(8)–C(14)	112·2(2)	111·6(2)	111·2	111·0	110·8
C(9)–C(8)–C(14)	109·0(2)	108·9(2)	108·5	109·0	109·0
C(8)–C(9)–C(10)	113·1(2)	114·6(2)	113·5	113·3	112·9
C(8)–C(9)–C(11)	112·9(2)	111·2(2)	113·0	112·7	112·6
C(10)–C(9)–C(11)	112·5(2)	112·8(2)	113·0	112·1	112·9
C(1)–C(10)–C(5)	109·7(2)	108·9(2)	109·7	110·3	109·3
C(1)–C(10)–C(9)	108·8(2)	109·4(2)	108·0	108·2	108·9
C(1)–C(10)–C(19)	110·3(2)	110·1(2)	111·1	110·6	110·2
C(5)–C(10)–C(9)	109·7(2)	109·7(2)	108·8	107·0	107·3
C(5)–C(10)–C(19)	107·4(2)	107·2(2)	108·6	108·0	109·4
C(9)–C(10)–C(19)	110·9(2)	111·5(2)	110·4	112·7	111·6
C(9)–C(11)–C(12)	113·4(2)	113·0(2)	113·1	113·2	112·3
C(11)–C(12)–C(13)	111·0(2)	110·9(2)	110·5	110·9	111·1
C(12)–C(13)–C(14)	107·9(2)	107·8(2)	107·8	108·1	108·7
C(12)–C(13)–C(17)	115·4(2)	114·4(2)	114·9	115·0	116·8
C(12)–C(13)–C(18)	110·7(3)	111·9(2)	110·9	112·2	111·5
C(14)–C(13)–C(17)	99·5(2)	99·3(2)	99·9	99·1	101·8
C(14)–C(13)–C(18)	113·7(2)	113·6(2)	113·7	113·2	112·2
C(17)–C(13)–C(18)	109·4(3)	109·3(3)	109·2	108·7	105·5
C(8)–C(14)–C(13)	113·7(2)	113·4(2)	113·8	113·3	114·3
C(8)–C(14)–C(15)	119·6(2)	120·8(2)	119·7	119·2	118·8
C(13)–C(14)–C(15)	104·1(3)	104·2(2)	103·4	104·3	104·2
C(14)–C(15)–C(16)	103·5(3)	104·0(3)	104·0	104·0	102·5
C(15)–C(16)–C(17)	106·6(3)	106·5(3)	105·9	106·6	107·9
C(13)–C(17)–C(16)	104·6(2)	104·8(3)	104·8	103·5	103·9
C(13)–C(17)–O(17)	115·4(2)	114·8(3)	115·7	115·0	113·0
C(16)–C(17)–O(17)	110·0(3)	113·8(3)	109·5	109·5	107·7

* Calculated from published co-ordinates.

The valency angles in the two molecules are also equal within experimental error. As is usual in steroid molecules, the valency angles in the six-membered rings are larger, and those in the five-membered ring smaller, than the tetrahedral value. The external valency angles at C(13) and C(14) are substantially larger than tetrahedral, counteracting the effect of the small interior angles in ring D and thus maintaining the constancy of the sum of the torsion angles C(12)–C(13)–C(14)–C(8) and C(17)–C(13)–C(14)–C(15) at the C/D ring junction.[12] The most notable difference in valency angles is 3·8° between the values of the C(16)–C(17)–O(17) angle in the two molecules. This may be related to the differences in hydrogen bonding discussed in the section on the extended crystal structure.

[12] H. J. Geise, C. Altona, and C. Romers, *Tetrahedron*, 1967, **23**, 439.

Molecular Conformation.—It is evident from an examination of the torsion angles (Table 7) that there are

TABLE 7

Torsion angles (deg.)

	T1	T2	H1*	H2*	E*
Ring A					
C(5)–C(10)–C(1)–C(2)	47·8	47·3	45·9	44·9	44·7
C(10)–C(1)–C(2)–C(3)	−55·7	−54·1	−57·5	−53·5	−57·6
C(1)–C(2)–C(3)–C(4)	33·4	29·2	37·6	31·1	36·2
C(2)–C(3)–C(4)–C(5)	−5·2	0·0	−8·1	−1·3	−4·0
C(3)–C(4)–C(5)–C(10)	−1·8	−5·1	−3·2	−7·1	−9·5
C(4)–C(5)–C(10)–C(1)	−19·6	−18·5	−16·2	−14·7	−11·4
Ring B					
C(9)–C(10)–C(5)–C(6)	44·9	45·0	48·9	50·2	52·4
C(10)–C(5)–C(6)–C(7)	−48·1	−50·1	−51·3	−52·0	−49·8
C(5)–C(6)–C(7)–C(8)	53·6	54·4	53·8	53·2	47·4
C(6)–C(7)–C(8)–C(9)	−58·7	−56·5	−56·7	−56·9	−51·1
C(7)–C(8)–C(9)–C(10)	57·6	54·5	56·6	59·0	57·6
C(8)–C(9)–C(10)–C(5)	−49·4	−46·8	−51·0	−52·9	−56·6
Ring C					
C(14)–C(8)–C(9)–C(11)	−50·1	−53·2	−51·3	−50·7	−51·1
C(8)–C(9)–C(11)–C(12)	49·9	52·2	51·1	49·8	51·6
C(9)–C(11)–C(12)–C(13)	−53·4	−54·4	−53·7	−53·2	−54·6
C(11)–C(12)–C(13)–C(14)	57·1	56·4	56·9	57·3	56·6
C(12)–C(13)–C(14)–C(8)	−61·5	−61·4	−61·9	−62·1	−60·0
C(13)–C(14)–C(8)–C(9)	57·6	59·7	58·2	58·3	56·3
Ring D					
C(17)–C(13)–C(14)–C(15)	46·0	45·8	46·3	46·7	44·9
C(13)–C(14)–C(15)–C(16)	−32·7	−35·2	−33·2	−32·9	−33·7
C(14)–C(15)–C(16)–C(17)	6·2	10·1	6·8	5·7	10·0
C(15)–C(16)–C(17)–C(13)	22·4	18·5	22·0	23·2	17·0
C(16)–C(17)–C(13)–C(14)	−41·6	−38·9	−41·9	−42·2	−37·6

* Calculated from published co-ordinates.

substantial differences in the conformations of the two molecules and that these differences are as great as, or greater than, differences between the hydrated and anhydrous testosterone molecules, or between any of the testosterone molecules and epitestosterone. It is fortunate indeed that the existence of two molecules in the asymmetric unit of anhydrous testosterone has provided a check for any inference about conformational differences which might be attributed to the change in orientation of the C(17) substituent.

The most substantial difference is in ring A. It has been observed before [13] that in steroids with a 4-en-3-one system there is a tendency for the four atoms [O(3), C(3)—(5)] comprising this system to deviate from planarity by *ca.* 0·03—0·04 Å. This deviation was attributed to a compromise between a sofa configuration for ring A which would required the torsion angle C(4)–C(5)–C(10)–C(1) to be *ca.* 27° and the need to relieve undue flattening of ring B (a sofa configuration for ring A would require the torsion angle C(9)–C(10)–C(5)–C(6) to approach 33°). In only two steroids (epitestosterone and 17α-hydroxyprogesterone [14]) was this system found to be strictly planar. It is interesting that in testosterone the four atoms C(3)—(5) and O(3) are coplanar in molecule B (root-mean-square deviation from the mean plane 0·004 Å) and significantly displaced from the mean plane in molecule A (0·016 Å). The two conformations thus

[13] P. J. Roberts, J. C. Coppola, N. W. Isaacs, and O. Kennard, *J.C.S. Perkin II*, 1973, 774.

[14] J. P. Declerq, G. Germain, and M. van Meerssche, *Cryst. Struct. Comm.*, 1972, **1**, 9.

appear equally probable and the molecules adopt one or other in the crystal structure either by chance or in consequence of packing forces.

Rings B and C are in the distorted chair form with the degree and position of the distortion varying in the five compounds. Ring D is intermediate between a C(13) envelope and a half-chair conformation. The variations in torsion angles among the five compounds reflect the flexibility of the five-membered ring system rather than

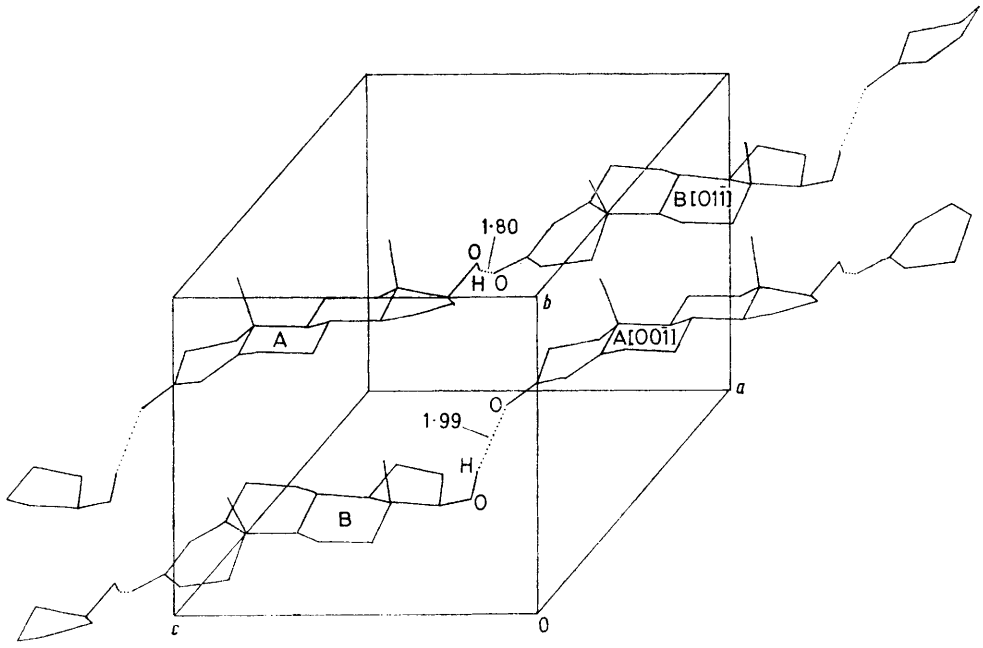


FIGURE 2 ' Head-to-tail ' hydrogen-bonding pattern between molecules A and B

any differences in non-bonded interactions due to a change in the orientation of the C(17) substituent. The conformation of ring D in molecule B resembles epitestosterone while in molecule A it resembles the hydrated forms of testosterone. These similarities are clearly indicated by the values of the pseudo-rotation constant [15] Δ (Table 8). Values of ϕ_m,[15] the phase angle

TABLE 8

Values of pseudorotation constant Δ and the phase angle of pseudorotation ϕ_m

Calc.	Δ(deg.)	ϕ_m(deg.)
C(14) envelope	−36	
Half-chair	0	
C(13) envelope	36	
T1	20·9	46·7
T2	9·8	46·0
H1	19·0	46·9
H2	21·1	47·5
E	9·1	45·0

of pseudorotation, cluster around 46° as is common in most steroids.

The Extended Crystal Structure.—Figure 2 shows a view of the unit cell and the hydrogen bonds linking the two molecules A and B in the extended structure. As is common in many steroids, there is a ' head-to-tail ' linking between the C(17) hydroxy-group attached to ring D and the O(3) function of ring A. However, in the present structure, the hydrogen bonding is between molecules of type A and B and gives rise to two OH · · · O distances of length 1·99 and 1·80 Å. The corresponding angles are OH · · · O 163° (A→B) and 167° (B→A). There are no intermolecular H · · · H contacts <2·35 Å.

Strong van der Waals interactions (Table 9) bind molecules A and B together. These were calculated by use of empirical atom-pair potentials [16] which have been shown to be sufficiently accurate to determine the position of minimum energy in several crystal structures.[17] The strongest associations in the structure are

TABLE 9

Intermolecular van der Waals interactions < −1·0 kcal. Symmetry operators (NS) are (1) x,y,z and (2) $-x, \frac{1}{2}+y, -z$

(*a*) Interactions between molecule A and symmetry generated molecules A′

NS	t_x	t_y	t_z	E
1	0	0	1	−3·18
2	1	0	1	−2·06

(*b*) Interactions between molecule A and symmetry-generated molecules B′

NS	t_x	t_y	t_z	E
1	0	0	0	−8·07
1	0	1	0	−8·40
2	0	0	1	−4·26
2	1	0	1	−2·09

(*c*) Interactions between molecule B and symmetry-generated molecules B′

NS	t_x	t_y	t_z	E
1	0	0	1	−3·03
2	0	0	1	−4·10

[15] C. Altona, H. J. Geise, and C. Romers, *Tetrahedron*, 1968, **24**, 13.

[16] E. Giglio, *Nature*, 1969, **222**, 339.

[17] W. D. S. Motherwell and N. W. Isaacs, *J. Mol. Biol.*, 1972, **71**, 231.

clearly those between the two molecules of the asymmetric unit and those relating molecule A to molecule B one cell translation in the *b* direction. The extended crystal structure can thus be said to consist of two parallel ribbons of molecules A and B (Figure 2) extended into a sheet in the *b* direction with considerably weaker links to an identical system related by the two-fold screw axis.

Structure and Biological Activity.—Our results suggest that observed differences in the steroid nucleus cannot be directly attributed to changes in the orientation of the C(17) substituent, but represent energetically equally favourable conformations or the effect of intermolecular forces. Thus any explanation of differences in the biological activity of the epimers is more likely to be related to the gross spatial orientation (α or β) of the substituents than to subtle conformational changes in the steroid skeleton. This conclusion is supported by a comparison of the structures of epiandrosterone and androsterone [3] where the only substantial difference between the two was the packing of the molecules in the extended crystal structure.

A further example of a pair of epimeric steroids is 5α-androstane-3α,17β-diol (VI) [6] and its 5β-epimer (V).[3] These two steroids differ in the orientation of the hydrogen attached to C(5) which results in an A/B *cis* function in the 5β-compound. Although the shape of the entire nucleus is radically altered in an A/B *cis*-steroid, the bond lengths and valency angles agree closely in these two compounds. As anticipated by Weeks *et al.*[3] from an examination of the cell constants, the α and β compounds are arranged in a remarkably similar manner with 'head-to-tail' hydrogen bonding about two-fold screw axes. The difference in activity is clearly to be attributed to the gross change in the steroid nucleus and it is perhaps significant that for this pair of epimers, in some biological situations it is the α-epimer which is more active while in others it is the β-epimer. In the other two pairs of epimers discussed, the activity is always less in the α- than in the β-epimer.

We thank the M.R.C. for financial support, the University of Cambridge Computing Service for computing facilities, and Dr. M. Hospital for providing data on the two hydrated forms of testosterone.

[3/750 *Received, 9th April,* 1973]