PREPARATION OF STEROID STRUCTURAL DATA FOR THE CONSIDERATION OF POSSIBLE STRUCTURAL-FUNCTIONAL RELATIONSHIPS

DORITA A. NORTON

From the Department of Biophysics, Roswell Park Memorial Institute, Buffalo, New York

ABSTRACT A format has been established for the presentation of steroid structural data which may prove to be useful in the consideration of the possible structural-functional relationships of steroids in biological systems. In preparing this format the detailed structural information for all steroids with known crystallographic structures has been computed and summarized, including bond lengths and angles, non-planarity coefficients, dihedral angles, length of the steroid nucleus, and side group inclinations. From the structural point of view it is interesting to note that (*a*) bond lengths do not differ significantly from the theoretical values, (*b*) bond angles do differ significantly from the theoretical values, (*c*) the non-planarity coefficient of the steroid molecule seems to be a function of the number of angular methyl groups but is unaffected by the presence of one heavy atom, (*d*) except where aromatic bonds are present the non-planarity coefficients of the two-ring and one-ring subdivisions of the steroid nucleus is 0.25 A, (*e*) the average non-planarity coefficient of the ring joins and "seat" portions of the chairs is 0.03 A, (*f*) dihedral angles deviate from the theoretically expected values and also show wide variation in the steroids studied, (*g*) the length of the steroid nucleus is variable, (*h*) analysis of side group inclination data must be deferred until more steroid structural information is available. From the biological point of view it is currently impossible to estimate the importance of each steroid structural feature in determining physiological function.

INTRODUCTION

One of the major reasons there has been so little progress in explaining biological function in terms of molecular structure is that crystallographers seem content to report atomic coordinates, and possibly bond lengths and angles, as the culmination of often beautiful and sometimes lengthy structure determinations. Whereas these data are informative to those interested in the study of deviations of bond lengths and angles from the normally accepted values, they are not complete enough to be of much use in the search for structural-functional relationships. What we need

most at this point in the evolution of the study of structural-functional relationships on the molecular level is a complete description of the pertinent details of molecular configuration.

This is so because the modern view of the cell is one of molecular organization. In this setting physiological function depends on the reactions and interactions which occur during the transport of various molecules, originating in mobile phases, through highly organized arrangements of other molecules comprising the stationary phases of cells. When we search for structural-functional relationships, therefore, we are really looking for some aspect or aspects of the molecular geometry (or electronic geometry) which can be held responsible for the behavior of the stationary phase–mobile phase system. Both the stationary phase and mobile phase molecules must be inspected for possible function-determining features. These features are the ones which determine the ability of molecules to get close enough to each other to react or interact. Such features include the quantitative details of molecular size and shape, which are defined for the steroids, at least, by bond lengths and angles, non-planarity coefficients, dihedral angles, length of the steroid nucleus, and side group inclinations.

The purposes of this paper are to summarize the detailed structural information currently known about steroids, and to establish a format for the presentation of steroid structural data which may prove to be important in the consideration of the possible structural-functional relationships of steroids in biological systems.

PRESENTATION OF STEROID STRUCTURAL DATA

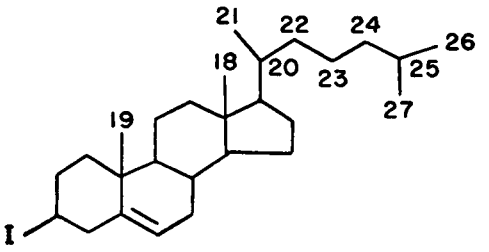
A. *Bond Lengths*

The crystal and molecular structures of six steroids have been determined by x-ray crystallographic techniques (1–6) (Fig. 1). The bond lengths found in these molecules have been reviewed and are discussed below.

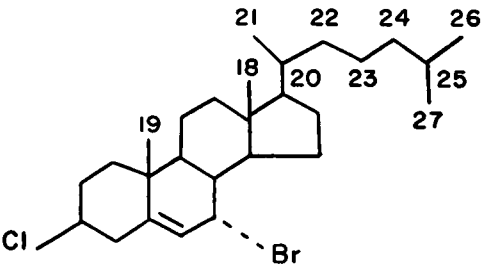
In general, steroid bond lengths fall into the following categories:

1. *Carbon-Carbon Single Bonds.* Theoretically, carbon-carbon single bonds should have a length of 1.54 A. It is somewhat surprising, therefore, to find that a number of such bonds (both nuclear and extranuclear) seem to be either longer or shorter than the theoretically expected value. When the carbon-carbon single bond length deviations are examined more closely, however, it is found that the maximum total percentage deviation is not significantly different from the statistically allowable distribution of one-third of the bond-length population outside of the standard deviation range. For this reason, it can be concluded that there is no evidence for real nuclear or extranuclear carbon-carbon single bond length deviations.

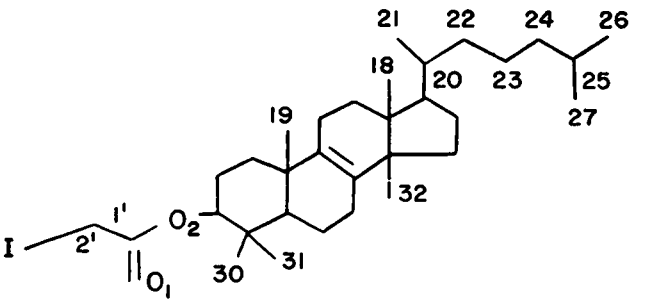
2. *Carbon-Carbon Double Bonds.* The carbon-carbon double bonds in steroids fall into the following two categories: (*a.*) isolated or simple double bonds whose theoretical length is 1.34 A; and (*b.*) bonds in aromatic systems whose



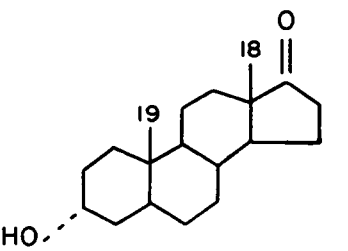
Cholesteryl iodide



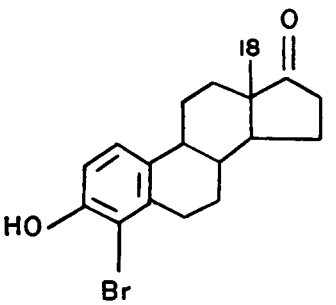
7α-Bromocholesteryl chloride



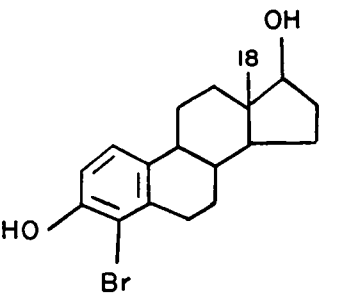
Lanostenyl iodoacetate



Androsterone



4-Bromoestrone



4-Bromoestradiol

FIGURE 1 Steroid structural formulas.

theoretical length is 1.39 A. In the steroids studied thus far, no significant deviations have been found in the lengths of either the simple or the aromatic double bonds present.

3. *Carbon-R Bonds.* (*a.*) *Carbon-halogen bonds.* The carbon-halogen bonds in the six steroids described in this paper are of the paraffinic type. As such, they should have the theoretical lengths of 2.13 A for iodine, 1.77 A for chlorine, and 1.94 A for bromine. No significant deviations from these theoretical values were observed.

(*b*). *Carbon-oxygen bonds.* C_3-O_3 *bonds.* The C_3-O_3 bond in androsterone is paraffinic and has a theoretical value of 1.43 A. The C_3-O_3 bonds in 4-bromoestrone and 4-bromoestradiol are aromatic and have a theoretical value of 1.36 A. No deviations were observed.

C_{17}-O_{17} *Bonds.* The C_{17}-O_{17} bonds in androsterone and 4-bromoestrone are of the keto type and have a theoretical value of 1.23 A. Androsterone shows a slight deviation from this (1.19 ± 0.01 A), whereas 4-bromoestrone shows no deviation. The C_{17}-O_{17} bond in 4-bromoestradiol (1.48 ± 0.02 A) is paraffinic and deviates from the theoretical value of 1.43 A. These deviations are not believed to be significant.

C-O bonds in the C-3 side chain of lanostenyl iodoacetate. The C-O bonds in this side chain are esteric. The double bond (C_1'-O_1) should have a length of 1.23 A, C_1'-O_2 should be 1.30 A, and O_2-C_3 should be 1.35 A. No significant deviations from these theoretical values were observed.

B. Bond Angles

Theoretically, the bond angles in steroids can be divided into four groups: (1) those which equal the tetrahedral angle (110°); (2) those in aromatic rings which equal

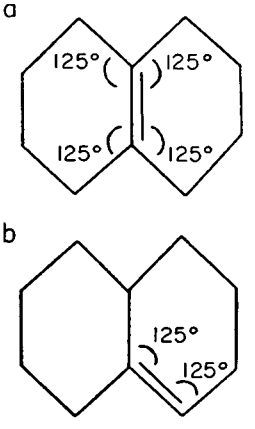


FIGURE 2 (*a.*) Double bond common to two fused rings. (*b.*) Double bond adjacent to two fused rings.

120°; (3) those in ring D which equal 105°; and (4) those which equal 125°, namely angles between a single and a double bond. In the case of condensed ring systems such as the steroids, a double bond common to two fused rings implies that four angles equal 125°, whereas a double bond adjacent to the ring fusion implies that two angles equal 125° (Fig. 2). In the six steroids studied thus far, there are numerous bond angle deviations from the theoretical values (Tables I and

TABLE I

NUCLEAR BOND ANGLES (DEGREES)*

		1	2	3	4	5	6
Angle	Standard deviation	6	2	6	1	2	2
10–1–2					113		123
1–2–3			115		112		
2–3–4		98	102				
3–4–5			117		113		123
4–5–10		120	113	121	113	116	
5–10–1			113	120	106		
10–5–6		114		103	112	123	123
5–6–7			114			117	
6–7–8		124	119	120	113	107	
7–8–9			104	102			
8–9–10			106		113		
9–10–5		123		118	107		
11–9–8		100	106	107	112		
9–8–14		119	106	135		106	105
8–14–13			124	102	113		
14–13–12		102	96				114
13–12–11						106	
12–11–9			127		113		113
17–13–14				90	99		98
13–14–15		93	102	120			
14–15–16		121			103	102	101
15–16–17		97					
16–17–13			101	116	108	108	

* Only those angles which differ significantly from the theoretically expected values have been listed.

II). When these deviations are summarized (Table III), it can be seen that many total percentage deviations are greater than the statistically permissible distribution of one-third of the bond angle population outside of the standard deviation range. For this reason, it can be concluded that some of the observed bond angle deviations are real. Unfortunately, there is presently no way of deciding which of the deviations are real and which are statistically allowable. Further consideration of bond angle deviations must be deferred until a greater number of steroid structures

TABLE II

EXTRANUCLEAR BOND ANGLES (DEGREES)*

Angle	Standard deviation	1	2	3	4	5	6
		6	2	6	1	2	2
12–13–18		91	127				
14–13–18			112			117	117
17–13–18		127	107	92			
1–10–19			107				
5–10–19		102	104		112		
9–10–19			114	83			
5–4–31				122			
30–4–31				97			
8–14–32				86			
13–14–32				121			
13–17–20			119	118			
16–17–20		100		118			
17–20–21			103				
17–20–22		100	114	100			
21–20–22		99	97	116			
20–22–23			114				
23–24–25		96	119	122			
24–25–26				121			
24–25–27		87	115				
26–25–27			117	90			
4–3–O_2				121			
3–O_2–1′				126			
O_2–1′–O_1				132			
O_1–1′–2′				111			
4–3–I_3		101					
3–4–Br_4							117
6–7–Br_7			92				
8–7–Br_7			113				
2–3–O_3					107		123
16–17–O_{17}					125	124	

* Only those angles which differ significantly from the theoretically expected values have been listed.

are determined with low strandard deviations (*e.g.* approximately ±0.005 A), and until a study of steroid molecular strain has been made.

C. *Non-Planarity Coefficients*[1]

1. *Entire Molecule and Nucleus.* For essentially planar molecules, such as the steroids discussed in this paper, additional quantitative description of

[1] This discussion does not take hydrogen atoms into consideration, since, in general, they were not observed in the structure determinations.

TABLE III

BOND ANGLE DEVIATIONS

	Carbon–Carbon–Carbon					
	Nuclear			Extranuclear		
	Small	Large	Total	Small	Large	Total
	per cent	*per cent*	*per cent*	*per cent*	*per cent*	*per cent*
Cholesteryl iodide	26	22	48	41	1	42
7α-Bromocholesteryl chloride	39	30	69	27	53	82
Lanostenyl iodoacetate	22	30	52	24	28	52
Androsterone	17	48	65	—	—	—
4-Bromoestrone	22	13	35	—	—	—
4-Bromoestradiol	13	22	35	—	—	—

	Carbon–Carbon–R		
	Small	Large	Total
	per cent	*per cent*	*per cent*
	21	**21**	**42**

molecular detail can be obtained by referring atoms to a series of best planes through the molecule (7–10). Best plane equations were calculated for the entire molecule and its various subdivisions (Fig. 3) using the IBM 1620 program of Harris and Harker (11). Each atom was referred to the various best planes by calculation of its perpendicular distance from the plane. A non-planarity coefficient for each plane considered was obtained by calculating the root mean square distances of atoms from the plane.

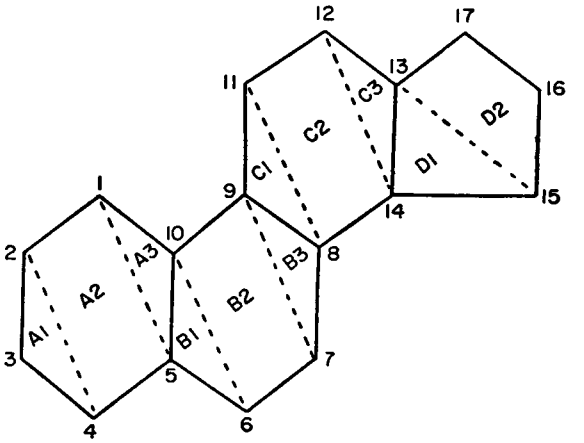


FIGURE 3 Various subdivisions of the steroid nucleus.

Table IV shows the non-planarity coefficients for the entire molecule plus and minus heavy atoms, and for the steroid nucleus with and without angular methyl groups. With the expection of 7α-bromocholesteryl chloride and lanostenyl iodoacetate the non-planarity coefficients of the best planes through the entire molecule including heavy atoms are uniform (0.49 to 0.57 A).

In the case of 7α-bromocholesteryl chloride the non-planarity coefficient of the best plane through the entire molecule including heavy atoms is high (0.67 A), presumably because of the presence of two heavy atoms. This explanation is supported by the observation that the non-planarity coefficient of the best plane through the entire molecule minus heavy atoms shows a significant decrease (0.57 A).

TABLE IV

NON-PLANARITY COEFFICIENTS ENTIRE MOLECULE AND NUCLEUS

	1*	2	3	4	5	6
Standard deviation	0.09	0.03	0.09	0.01	0.04	0.02
Entire molecule	0.57	0.67	0.81	—	0.49	0.53
Entire molecule minus heavy atoms	0.58	0.57	0.81	0.77	0.49	0.51
Nucleus plus five angular methyls	—	—	0.77	—	—	—
Nucleus plus two angular methyls	0.59	0.65	0.61	0.62	—	—
Nucleus plus C-18 angular methyl	0.47	0.52	0.49	0.49	0.50	0.51
Nucleus	0.28	0.28	0.30	0.27	0.26	0.24

* Numerals in box heading: (1) Cholesteryl iodide; (2) 7α-Bromocholesteryl chloride; (3) Lanostenyl iodoacetate; (4) Androsterone; (5) 4-Bromoestrone; (6) 4-Bromoestradiol.

In the case of lanostenyl iodoacetate the exceedingly high non-planarity coefficient (0.81 A) of the best plane through the entire molecule including the heavy atom seems to be due primarily to the presence of the five angular methyl groups, rather than to the presence of the heavy atom or the C_3 side chain. The evidence for this is that (1) the non-planarity coefficient of the best plane through the entire molecule minus the heavy atom is also 0.81 A, and (2) the non-planarity coefficient of the best plane through the nucleus plus five angular methyls (but minus the C_3 side chain) is 0.77 A. It is interesting to note that in the case of lanostenyl iodoacetate the non-planarity coefficients gradually decrease as the number of angular methyl groups is reduced (0.77 A for five, 0.61 A for two, 0.49 A for one, and 0.30 A for none). This suggests that the over-all planarity of the entire steroid molecule is a function of the number of angular methyl groups present. Additional support for this hypothesis is that a similar decrease is observed in compounds 1 through 4 where non-planarity coefficient data are available for two, one, and zero angular methyl groups.

With the exception of 7α-bromocholesteryl chloride whose high non-planarity coefficient (0.67 A) seems to be due to the presence of two heavy atoms, and

lanostenyl iodoacetate whose high non-planarity coefficient (0.81 A) is caused by the five angular methyl groups, the non-planarity coefficients of the best planes through entire steroid molecules minus heavy atoms are nearly constant (0.49 to 0.58 A) and equal to the non-planarity coefficients of the best planes through entire steroid molecules plus heavy atoms. This indicates that the presence or absence of a single heavy atom, such as a halogen, has no effect on the planarity of the entire molecule.

2. *Two-Ring and One-Ring Subdivisions* (Table V). The average non-planarity of both the two-ring and one-ring subdivisions of the steroid nucleus is 0.25 A. Notable exceptions to this are found in 4-bromoestrone and 4-bromoestra-

TABLE V

NON-PLANARITY COEFFICIENTS TWO-RING AND ONE-RING SUBDIVISIONS

	1*	2	3	4	5	6
Standard deviation	0.09	0.03	0.09	0.01	0.04	0.02
AB	0.28	0.29	0.29	0.23	0.19	0.19
BC	0.22	0.29	0.26	0.23	0.26	0.27
CD	0.24	0.23	0.23	0.21	0.24	0.23
A	0.29	0.22	0.21	0.23	0.02	0.02
B	0.17	0.26	0.24	0.23	0.27	0.23
C	0.26	0.23	0.25	0.23	0.27	0.26
D	0.22	0.21	0.16	0.20	0.18	0.22

* Numerals in box heading: (1) Cholesteryl iodide; (2) 7α-Bromocholesteryl chloride; (3) Lanostenyl iodoacetate; (4) Androsterone; (5) 4-Bromoestrone; (6) 4-Bromoestradiol.

diol. In the AB two-ring subdivision, the aromatic bonds of ring A reduce the non-planarity coefficient to 0.19 A in both cases. As expected, the A-rings of these compounds are planar (non-planarity coefficients both equal 0.02 A).

3. *Ring Joins and "Seat" Portions of Chairs* (Table VI). The ring joins and "seat" portions of chairs are planar (average non-planarity coefficients equal 0.03 A). Several deviations from the expected planarity of ring joins are apparent.

D. *Dihedral Angles*

The ten dihedral angles of steroids (Fig. 4), *i.e.*, angles between best planes, may be divided into the following three groups on the basis of theoretical expectations: (1) those equal to 120 degrees, (2) those equal to 180 degrees (coplanar), and (3) those equal to some value between 120 and 180 degrees. Group (1) includes all dihedral angles except those at ring joins, those involving a double bond, and those in ring D. Group (2) includes the dihedral angles of ring-joins between six-

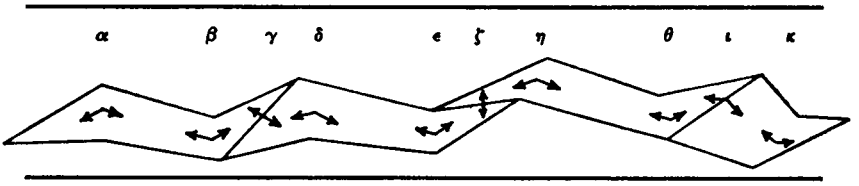
TABLE VI

NON-PLANARITY COEFFICIENTS RING JOINS AND "SEAT" PORTIONS

	1*	2	3	4	5	6
Standard deviation	0.09	0.03	0.09	0.01	0.04	0.02
A3B1	0.24	0.23	0.00	0.03	0.03	0.02
B3C1	0.03	0.09	0.05	0.02	0.02	0.02
C3D1	0.12	0.17	0.00	0.08	0.16	0.08
A2	0.11	0.03	0.01	0.00	0.01	0.01
B2	0.10	0.17	0.16	0.02	0.02	0.09
C2	0.07	0.02	0.03	0.01	0.00	0.04
D2	0.17	0.14	0.03	0.06	0.01	0.12

* Numerals in box heading: (1) Cholesteryl iodide; (2) 7α-Bromocholesteryl chloride; (3) Lanostenyl iodoacetate; (4) Androsterone; (5) 4-Bromoestrone; (6) 4-Bromoestradiol.

FIGURE 4 Dihedral angles (degrees).



	1	2	3	4	5	6
Standard deviation	6	2	6	1	2	2
α	117	127^{L}	128^{L}	133^{L}	180*	174*
β	136^{L}	136^{L}	144^{L}	131^{L}	180*	170*
γ	137‡	132‡	179*	175^{*S}	175^{*S}	174^{*S}
δ	180‡	163‡	139^{L}	127^{L}	174^{*S}	175^{*S}
ϵ	146^{L}	129^{L}	163‡	132^{L}	122	124^{L}
ζ	180*	166^{*S}	173^{*S}	176^{*S}	171^{*S}	180*
η	132^{L}	142^{L}	155‡	132^{L}	127^{L}	127^{L}
θ	128^{L}	129^{L}	131^{L}	129^{L}	119	127^{L}
ι	161‡	153‡	174‡	168‡	161‡	168‡
κ	147‡	143‡	140‡	137‡	139‡	136‡

* Theoretical value is 180°.
‡ Theoretical value not known.
S = Small
L = Large

membered rings (whether or not they have a double bond at their apex), dihedral angles in aromatic rings, and those adjacent to aromatic rings (*e.g.*, δ in 4-bromoestrone and 4-bromoestradiol). Group (3) includes dihedral angles between ring subdivisions which contain a double bond and the κ-dihedral angle in ring D. Assignment of theoretical values to the dihedral angles in group three must be deferred

until further study is made of the strained ring systems of steroids; namely, rings containing one double bond, and particularly the D-ring.

A comparison of the experimentally observed dihedral angles with known theoretically expected values shows that a large number of deviations occur (55 per cent large, 2 per cent small, 57 per cent total).

It should also be noted that each of the dihedral angles shows great variation from steroid to steroid. This implies that the length of the steroid nucleus is variable. Table VII lists the experimentally determined lengths (C_3-C_{16}) of the six steroid nuclei studied herein.

TABLE VII

LENGTH OF THE STEROID NUCLEUS (C_3-C_{16})

	(A)	
Cholesteryl iodide	9.35	±0.09
7α-Bromocholesteryl chloride	8.92	±0.03
Lanostenyl iodoacetate	9.21	±0.09
Androsterone	9.02	±0.01
4-Bromoestrone	8.97	±0.04
4-Bromoestradiol	8.99	±0.02

E. *Side Group Inclination*

The angles of inclination which the various side groups make with neighboring best planes have been calculated. Table VIII gives the angles of inclination with the best plane of the ring to which the side group is attached or with the best plane of the ring to the left of side groups attached at ring fusions. (Similar calculations have been made with respect to all the best planes of the molecules in question, and are available upon request.) Certain irregularities are apparent in Table VIII, such as the surprisingly low angles of inclination of the C_{18} angular methyl groups of cholesteryl iodide and 7α-bromocholesteryl chloride, and the surprisingly high angles of inclination of the C_{31} and C_{32} angular methyl groups of lanostenyl iodoacetate. Analyses of these data, however, must be deferred until structural information is available for a larger number of steroids.

SUMMARY

A format has been established for the concise presentation of steroid x-ray structural data. It is anticipated that this format will be useful in the search for possible steroid structural-functional relationships in biological systems. Inspection of these data for six steroids indicates that the following are salient from the structural point of view:

1. Bond lengths do not differ significantly from the theoretical values.
2. Bond angles differ significantly from the theoretical values.

TABLE VIII

SIDE GROUP INCLINATION* (DEGREES)

	1‡	2	3	4	5	6
Standard deviation	6	2	6	1	2	2
C_{18}	69	70	81	82	81	82
C_{19}	87	88	81	84		
C_{30}			24			
C_{31}			79			
C_{32}			80			
O_3				83	0	3
O_{17} (keto)				20	9	
O_{17} (hydroxyl)						28
Br_4					1	5
Br_7		84				
I_3	23					
Cl_3		19				

* With respect to the best plane to the left of the side group when the group is at the join of two best planes.

‡ Numerals in box heading: (1) Cholesteryl iodide; (2) 7α-Bromocholesteryl chloride; (3) Lanostenyl iodoacetate; (4) Androsterone; (5) 4-Bromoestrone; (6) 4-Bromoestradiol.

3. The average non-planarity coefficients of the steroid molecule and its various subdivisions are as follows:

	A
Entire molecule	0.61
Entire molecule minus heavy atoms	0.62
Nucleus plus two angular methyls	0.62
Nucleus	0.27
Two-ring subdivisions	0.24
One-ring subdivisions	0.23
Ring join portions[a]	0.04
"Seat" portions[a]	0.03

4. The dihedral angles deviate from the theoretically expected values and also show wide variation in the six steroids studied.

5. The length of the steroid nucleus is variable.

6. Analysis of side group inclination data must be deferred until more steroid structural information is available.

[a] Non-planarity coefficients attributable to strain have been omitted from the calculation of averages.

DISCUSSION

From the biological point of view, it is impossible even to estimate at this time the importance of each structural feature in determining physiological function. Although some of the biological specificity of steroids can be attributed to the chemistry of the functional side groups, it can be expected that some of the biological specificity will be explained by the structural subtleties of the steroid molecule.

This investigation was supported by a P.H.S. research grant (CA-06183) from the National Cancer Institute, Public Health Service.

Received for publication, May 26, 1964.

BIBLIOGRAPHY

1. CARLISLE, C. H., and CROWFOOT, D., *Proc. Roy. Soc.*, 1945, **A184**, 64.
2. BÜRKI, H., and NOWACKI, W., *Z. Krist.*, 1956, **108**, 206.
3. FRIDRICHSONS, J., and MATHIESON, A. McL., *Chem. Soc. London*, 1953, 2159.
4. HIGH, D. F., The Crystal Structure of Androsterone, 1962, University of Washington Ph.D. Thesis, Ann Arbor, University Microfilms Inc., Michigan.
5. NORTON, D. A., KARTHA, G., and LU, C. T., *Acta Cryst.*, 1963, **16**, 89.
6. NORTON, D. A., KARTHA, G., and LU, C. T., *Acta Cryst.*, 1964, **17**, 77.
7. NORTON, D. A., and OHRT, J. M., *Nature*, 1963, **197**, 450.
8. OHRT, J. M., HANER, B. A., and NORTON, D. A., *Nature*, 1963, **199**, 1180.
9. OHRT, J. M., and NORTON, D. A., *Nature*, 1963, **199**, 864.
10. NORTON, D. A., and OHRT, J. M., *Nature*, 1964, **203**, 754.
11. HARRIS, D., and HARKER, D., Best Planes through N Points, 1963, unpublished.