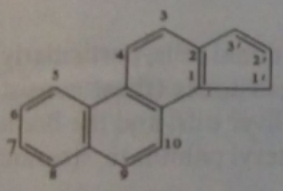
11 Steroids

§1. Introduction

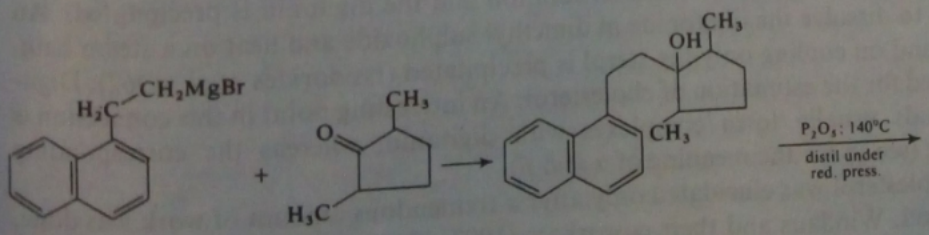
The steroids form a group of structurally related compounds which are widely distributed in animals and plants. Included in the steroids are the sterols (from which the name *steroid* is derived), vitamin D, the bile acids, a number of sex hormones, the adrenal cortex hormones, some carcinogenic hydrocarbons, certain sapogenins, etc. The structures of the steroids are based on the 1,2-cyclopentenophenanthrene skeleton (Rosenheim and King, 1932; Wieland and Dane, 1932). All the steroids give, among other products, Diels' hydrocarbon on dehydrogenation with selenium at 360°C (Diels, 1927). In fact, a steroid could be defined as any compound which gives Diels' hydrocarbon when distilled with selenium. When the distillation with selenium is carried out at 420°C, the steroids give mainly chrysene (**10** §4b) and a small amount of picene (**10** §4c).

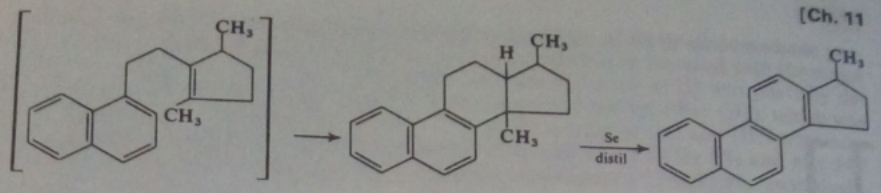


1,2-cyclopentenophenanthrene

In the earlier work, the various steroids were designated by trivial names, but the tendency now is to discard these in favour of systematic names, which may be applied when the structure is known (see §7).

Diels' hydrocarbon is a solid, m.p. 126–127°C. Its molecular formula is $C_{18}H_{16}$, and the results of oxidation experiments, X-ray crystal analysis and absorption spectrum measurements showed that the hydrocarbon is probably 3′-methyl-1,2-cyclopentenophenanthrene. This structure was definitely established by synthesis, *e.g.*, that of Harper, Kon and Ruzicka (1934), who used the Bogert–Cook method [**10** §2vi], starting from 2-(1-naphthyl)-ethylmagnesium bromide and 2,5-dimethylcyclopentanone.





Diels' hydrocarbon

Sterols

§2

Sterols occur in animal and plant oils and fats. They are crystalline compounds, and contain an alcoholic group; they occur free or as esters of the higher fatty acids, and are isolated from the unsaponifiable portion of oils and fats. Cholesterol, 5α-cholestan-3β-ol (cholestanol) and 5β-cholestan-3β-ol (coprostanol) are the animal sterols; ergosterol and stigmasterol are the principal plant sterols. The sterols that are obtained from animal sources are often referred to as the *zoosterols*, and those obtained from plant sources as the *phytosterols*. A third group of sterols, which are obtained from yeast and fungi, are referred to as the *mycosterols*. This classification, however, is not rigid, since some sterols are obtained from more than one of these groups.

§3. **Cholesterol,** $C_{27}H_{46}O$, m.p. 149°C.

This is the sterol of the higher animals, occurring free or as fatty esters in all animal cells, particularly in the brain and spinal cord. Cholesterol was first isolated from human gallstones (these consist almost entirely of cholesterol). The main sources of cholesterol are the fish-liver oils, and the brain and spinal cord of cattle. Lanoline, the fat from wool, is a mixture of cholesteryl palmitate, stearate and oleate.

Cholesterol is a white crystalline solid which is optically active, ($[\alpha]_D$ 39°). Cholesterol (and other sterols) gives many colour reactions, *e.g.*,

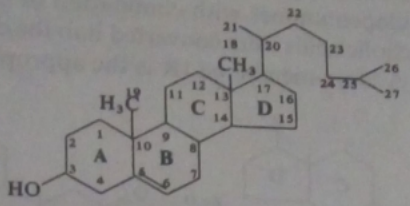
(i) *The Salkowski reaction* (1908). When concentrated sulphuric acid is added to a solution of cholesterol in chloroform, a red colour is produced in the chloroform layer.

(ii) *The Liebermann–Burchard reaction* (1885, 1890). A greenish colour is developed when a solution of cholesterol in chloroform is treated with concentrated sulphuric acid and acetic anhydride.

When an ethanolic solution of cholesterol is treated with an ethanolic solution of digitonin (a saponin; see §32), a large white precipitate of cholesterol digitonide is formed. This is a molecular complex containing one molecule of cholesterol and one of digitonin, from which the components may be recovered by dissolving the complex in pyridine (which brings about complete dissociation) and then adding ether (the cholesterol remains in solution and the digitonin is precipitated). An alternative method is to dissolve the digitonide in dimethyl sulphoxide and heat on a steam bath. Dissociation occurs, and on cooling only the sterol is precipitated (Issidorides *et al.*, 1962). Digitonide formation is used for the estimation of cholesterol. An interesting point in this connection is that 3β-hydroxysteroids usually form complexes with digitonin, whereas the corresponding 3α-compounds do not (see §5 for the meaning of α and β).

The structure of cholesterol was elucidated only after a tremendous amount of work was done, particularly by Wieland, Windaus and their coworkers (1903–1932). Only a very bare outline is

given here, and in order to appreciate the evidence that is going to be described, it is necessary to have the established structure of cholesterol at the beginning of our discussion. (I) is the structure of



(I)

cholesterol, and shows the method of numbering. The molecule consists of a *side-chain* and a *nucleus* which is composed of four rings; these rings are usually designated A, B, C and D (or (I), (II), (III) and (IV)), beginning from the six-membered ring on the left (see also (iii) below). It should be noted that the nucleus contains two angular methyl groups, one at C-10 and the other at C-13.

(i) **Structure of the ring system.** Under this heading we shall deal with the nature of the ring system present in cholesterol; the problem of the angular methyl groups is dealt with later [see (iv)].

The usual tests for functional groups showed that cholesterol contains one double bond and one hydroxyl group. Now let us consider the following set of reactions.

$$\underset{C_{27}H_{46}O\ (I)}{\text{Cholesterol}} \xrightarrow{H_2-Pt} \underset{C_{27}H_{48}O\ (II)}{\text{Cholestanol}} \xrightarrow{CrO_3} \underset{C_{27}H_{46}O\ (III)}{\text{Cholestanone}} \xrightarrow[HCl]{Zn-Hg} \underset{C_{27}H_{48}\ (IV)}{\text{Cholestane}}$$

The conversion of cholesterol into cholestanol (II) shows the presence of one double bond in (I) and the oxidation of (II) to the ketone cholestanone (III) shows that cholesterol is a secondary alcohol. Cholestane (IV) is a saturated hydrocarbon, and corresponds to the general formula C_nH_{2n-6}, and consequently is tetracyclic; thus cholesterol is tetracyclic. [D.B.E. of cholestane is $27 + 1 - 48/2 = 4$.]

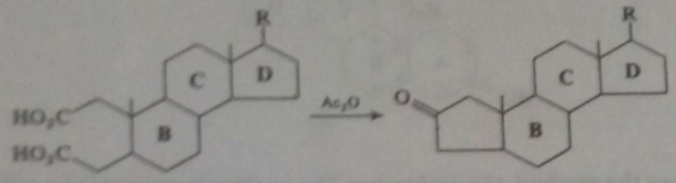
When cholesterol is distilled with selenium at 360°C, Diels' hydrocarbon is obtained (see §1). The formation of this compound could be explained by assuming that this nucleus is present in cholesterol. The yield of this hydrocarbon, however, is always poor, and other products are always formed at the same time, particularly chrysene (see §1). Thus, on the basis of this dehydrogenation, the presence of the cyclopentenophenanthrene nucleus must be accepted with reserve. Rosenheim and King (1932) thought that chrysene was the normal product of the selenium dehydrogenation, and so proposed (on this basis and also on some information obtained from X-ray analysis work of Bernal, 1932; see §5) that the steroids contained the chrysene skeleton. Within a few months, however, Rosenheim and King (1932) modified this suggestion, as did also Wieland and Dane (1932). These two groups of workers proposed that the cyclopentenophenanthrene nucleus is the one present in cholesterol (*i.e.*, in steroids in general). This structure fits far better all the evidence that has been obtained from a detailed investigation of the oxidation products of the sterols and bile acids, and has now been confirmed by the synthesis of cholesterol (see §9).

(*a*) The nature of the *nucleus* in sterols and bile acids was shown to be the same, since 5β-cholanic acid (cholanic acid) or 5α-cholanic acid (allocholanic acid) is one of the oxidation products (see §5).

(*b*) The oxidation of the bile acids led to the formation of products in which various rings were opened. The examination of these products showed that the positions of the hydroxyl groups were limited mainly to three positions 3, 7 and 12, and further work showed that the hydroxyl groups behaved differently towards a given reagent (see also §5).

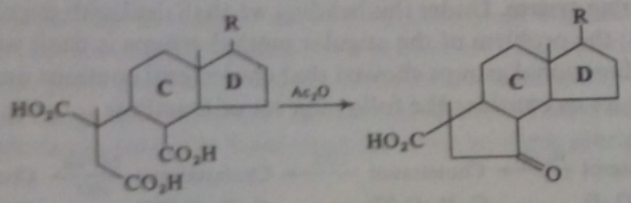
(*c*) The rings in the steroid nucleus were opened to give a dicarboxylic acid and the relative positions of the two carboxyl groups with respect to each other were determined by the application of **Blanc's rule**: On heating with acetic anhydride, 1,5-dicarboxylic acids form cyclic anhydrides, and 1,6-dicarboxylic acids form cyclopentanones with elimination of carbon dioxide (see also Vol. I).

Ring A. Cholesterol and the cholic acids were converted into the dicarboxylic acid (A) which gave a cyclopentanone, and so ring A is six-membered (R is the appropriate side-chain).



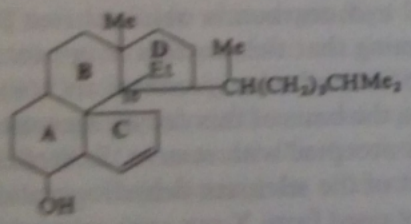
(A)

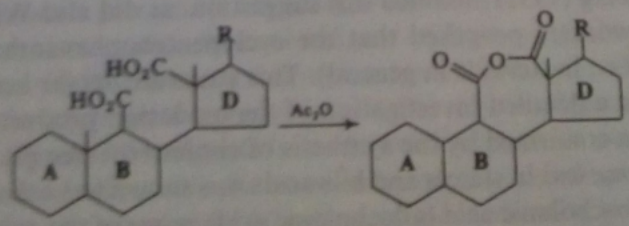
Ring B. Cholesterol was converted into the tricarboxylic acid (B) which gave the cyclopentanone derivative shown. Hence ring B is six-membered.



(B)

Ring C. Deoxycholic acid was converted into a dicarboxylic acid which gave a cyclic anhydride. It was therefore assumed that ring C was five-membered, and this led Windaus and Wieland (1928) to propose the following formula for cholesterol, and the uncertain point (at that time) was the nature of the two extra carbon atoms. These were *assumed* to be present as an ethyl group at position 10, but Wieland *et al.* (1930) finally proved that there was no ethyl group at this position. These two 'homeless' carbon atoms were not placed until Rosenheim and King first proposed that steroids contained the chrysene nucleus and then proposed the cyclopentenophenanthrene nucleus (see above). Bernal (1932) also showed, from the X-ray analysis of cholesterol, ergosterol, etc., that the molecule was thin, whereas the above structure for the steroid nucleus would be rather thick.

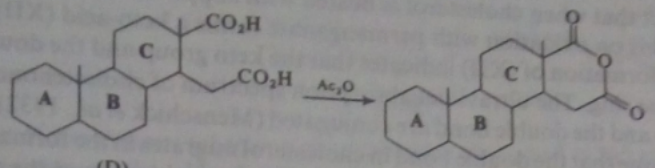




(C)

If we use the correct structure of cholesterol, the cyclisation reaction results in the formation of a *seven-membered cyclic anhydride*. Thus, in this case (and in some others), the Blanc rule fails and leads to erroneous conclusions.

Ring D. 5β-Cholestane (Coprostane) was converted into etiobilianic acid (see (iii), below), and this gave a cyclic anhydride. Hence ring D is five-membered.



(D)

(ii) **Positions of the hydroxyl group and double bond.** Let us consider the following reactions:

$$\underset{C_{27}H_{46}O\,(III)}{\text{Cholestanone}} \xrightarrow{HNO_3} \underset{C_{27}H_{46}O_4\,(V)}{\text{Dicarboxylic acid}} \xrightarrow{300^\circ C} \underset{C_{26}H_{44}O\,(VI)}{\text{Ketone}}$$

Since the dicarboxylic acid (V) contains the same number of carbon atoms as the ketone (III) from which it is derived, the keto group in (III) must therefore be in a ring. Also, since pyrolysis of the dicarboxylic acid (V) produces a ketone with the loss of one carbon atom, it therefore follows from Blanc's rule that (V) is either a 1,6- or 1,7-dicarboxylic acid. Now we have seen that the nucleus contains three six-membered rings and one five-membered ring. Thus the dicarboxylic acid (V) must be obtained by the opening of ring A, B or C, and consequently it follows that the hydroxyl group in cholesterol (which was converted into the keto group in cholestanone; see (i) above) is in ring A, B or C.

Actually *two* isomeric dicarboxylic acids are obtained when cholestanone is oxidised. The formation of these two acids indicates that the keto group in cholestanone is flanked on either side by a methylene group, *i.e.*, the grouping $-CH_2COCH_2-$ is present in cholestanone. Examination of the reference structure (I) of cholesterol shows that such an arrangement is possible only if the hydroxyl group is in ring A.

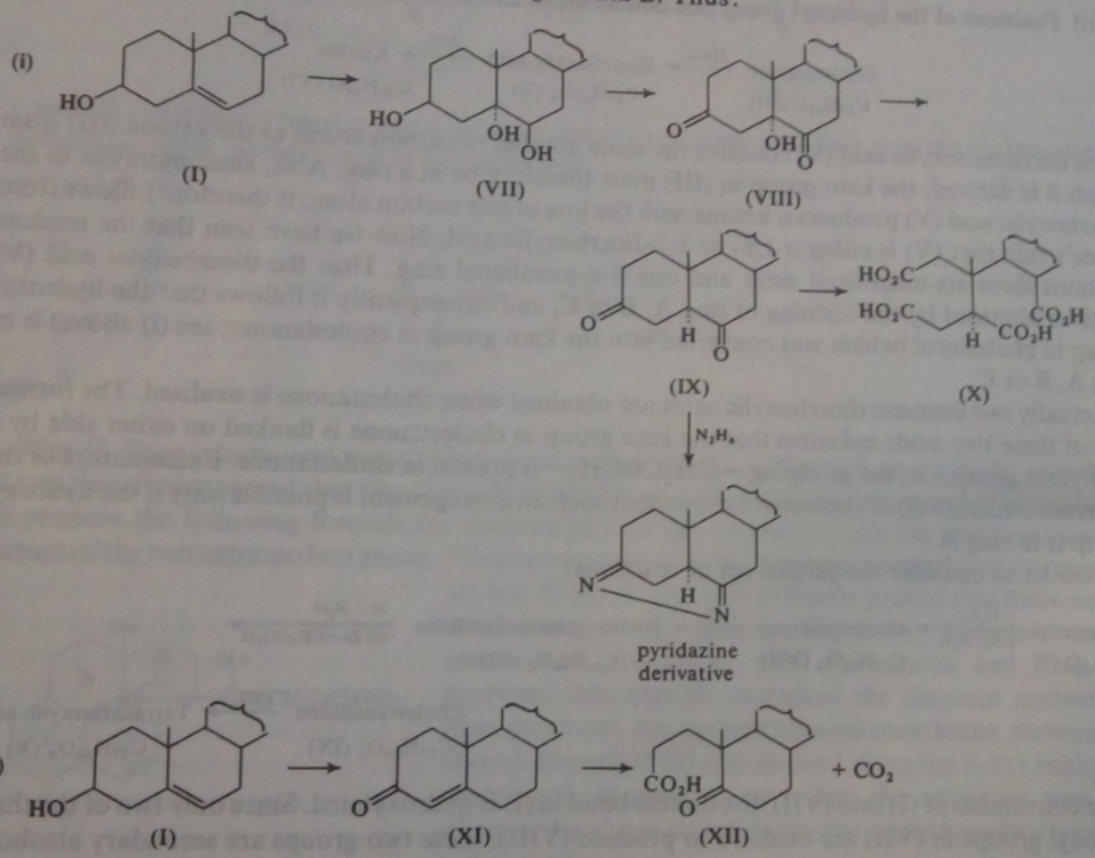
Now let us consider the further set of reactions:

$$\underset{C_{27}H_{46}O\,(I)}{\text{Cholesterol}} \xrightarrow[CH_3CO_3H]{H_2O_2} \underset{C_{27}H_{48}O_3\,(VII)}{\text{Cholestanetriol}} \xrightarrow{CrO_3} \underset{C_{27}H_{44}O_3\,(VIII)}{\text{Hydroxycholestanedione}} \xrightarrow[(ii)\ Zn-CH_3CO_2H]{(i)\ -H_2O}$$

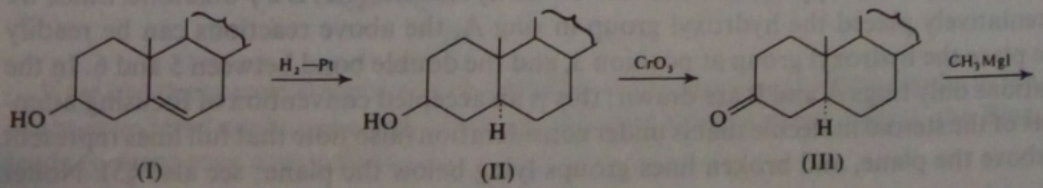
$$\underset{C_{27}H_{44}O_2\,(IX)}{\text{Cholestanedione}} \xrightarrow{CrO_3} \underset{C_{27}H_{44}O_8\,(X)}{\text{Tetracarboxylic acid}}$$

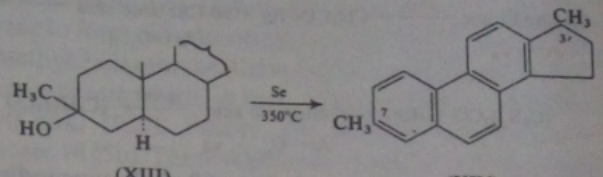
In the conversion of (I) into (VII), the double bond in (I) is hydroxylated. Since only two of the three hydroxyl groups in (VII) are oxidised to produce (VIII), these two groups are secondary alcoholic groups (one of these being the secondary alcoholic group in cholesterol), and the third, being resistant to oxidation, is probably a tertiary alcoholic group. Dehydration of (VIII) (by heating *in vacuo*) and subsequent reduction of the double bond forms (IX), and this, on oxidation, gives a tetracarboxylic acid *without loss of carbon atoms*. Thus the two keto groups in (IX) must be in *different* rings; had they been in the *same* ring, then carbon would have been lost and (X) not obtained. It therefore follows that the hydroxyl group and double bond in cholesterol must be in *different* rings. Furthermore, since (IX) forms a pyridazine derivative with hydrazine, (IX) is a γ-diketone. Since we have already tentatively placed the hydroxyl group in ring A, the above reactions can be readily explained if we place the hydroxyl group at position 3, and the double bond between 5 and 6. In the following equations only rings A and B are drawn; this is an accepted convention of focusing attention on any part of the steroid molecule that is under consideration (also note that full lines represent groups lying above the plane, and broken lines groups lying below the plane; see also §5). Noller

(1939) has shown that the pyridazine derivative is a polymer, and so the interpretation that (IX) is a γ-diketone is rendered uncertain. Supporting evidence, however, for the above interpretation is afforded by the fact that when cholesterol is heated with copper oxide at 290°C, cholestenone (XI) is produced, and this on oxidation with permanganate forms a keto-acid (XII) with the loss of one carbon atom. The formation of (XII) indicates that the keto group and the double bond in cholestenone are in the *same* ring. The ultraviolet absorption spectrum of cholestenone, λ_{max} 240 nm, shows that the keto group and the double bond are conjugated (Menschick *et al.*, 1932). These results can be explained if we assume that the double bond in cholesterol migrates in the formation of cholestenone, the simplest explanation being that the hydroxyl group is in position 3 and the double bond between 5 and 6, *position 5 being common to both rings A and B*. Thus:



The position of the hydroxyl group at position 3 is definitely proved by the experiments of Kon *et al.* (1937, 1939). These authors reduced cholesterol (I) to cholestanol (II), oxidised this to cholestanone (III), treated this with methylmagnesium iodide and dehydrogenated the product, a tertiary alcohol (XIII), to 3′,7-dimethylcyclopentenophenanthrene (XIV) by means of selenium. The structure of (XIV) was proved by synthesis, and so the reactions may be formulated as follows, with the hydroxyl at position 3.

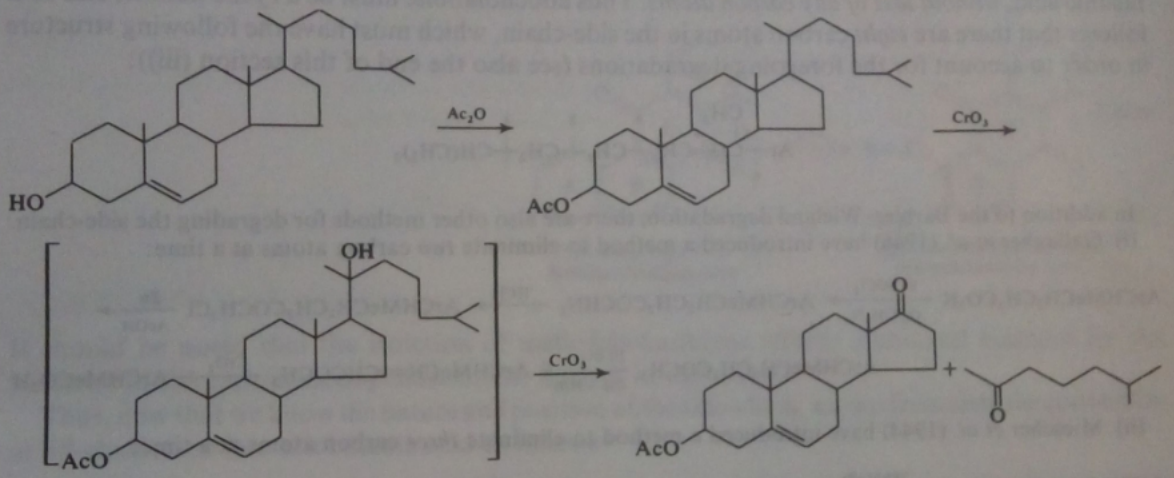




(XIII) (XIV)

The stereochemistry of the various reactions given above is discussed in §§5 and 8.

(iii) **Nature and position of the side-chain.** Acetylation of cholesterol produces cholesteryl acetate and this, on oxidation with chromium trioxide, forms a steam-volatile ketone and the acetate of a hydroxyketone (which is not steam volatile). The ketone was shown to be isohexyl methyl ketone, $CH_3CO(CH_2)_3CH(CH_3)_2$. Thus this ketone is the side-chain of cholesterol, the point of attachment of the side-chain being at the carbon of the keto group. These results do not show where the side-chain is attached to the nucleus of cholesterol, but if we accept that the position is at 17, then we may formulate the reactions as follows:



The nature of the side-chain has also been shown by the application of the Barbier–Wieland degradation. Since this method also leads to evidence that shows *which ring* of the nucleus is attached to the side-chain, we shall consider the problem of the nature of the side-chain again.

The Barbier–Wieland degradation offers a means of 'stepping down' an acid one carbon atom at a time as follows:

$$RCH_2CO_2H \xrightarrow[HCl]{CH_3OH} RCH_2CO_2CH_3 \xrightarrow{2C_6H_5MgBr} RCH_2C(OH)(C_6H_5)_2 \xrightarrow{-H_2O} RCH{=}C(C_6H_5)_2 \xrightarrow{CrO_3} RCO_2H + (C_6H_5)_2CO$$

Methylmagnesium bromide may be used instead of phenylmagnesium bromide, and the alcohol so obtained may be directly oxidised:

$$RCH_2C(OH)(CH_3)_2 \xrightarrow{CrO_3} RCO_2H + (CH_3)_2CO$$

In the following account, only phenylmagnesium bromide will be used to demonstrate the application of the method to the steroids.

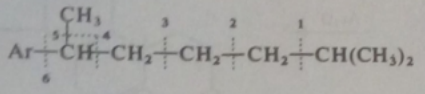
Cholesterol was first converted into 5β-cholestane (coprostane). If we represent the nucleus of 5β-cholestane as Ar, and the side-chain as C_n, then we may formulate the degradation of 5β-cholestane as follows (B–W represents a Barbier–Wieland degradation):

$$5\beta\text{-Cholestane}\ (Ar{-}C_n) \xrightarrow{CrO_3} CH_3COCH_3 + 5\beta\text{-Cholanic acid}\ (Ar{-}C_{n-3}) \xrightarrow{B-W} (C_6H_5)_2CO + \text{Nor-}5\beta\text{-cholanic acid}\ (Ar{-}C_{n-4}) \xrightarrow{B-W}$$

$$(C_6H_5)_2CO + \text{Bisnor-}5\beta\text{-cholanic acid}\ (Ar{-}C_{n-5}) \xrightarrow{B-W} (C_6H_5)_2CO + \text{Etiocholyl methyl ketone}\ (Ar{-}C_{n-6}) \xrightarrow{CrO_3} 5\beta\text{-Etianic acid}\ (Ar{-}C_{n-7})$$

The formation of acetone from 5β-cholestane indicates that the side-chain terminates in an isopropyl group. The conversion of bisnor-5β-cholanic acid into a ketone shows that there is an alkyl group on the α-carbon atom in the former compound. Furthermore, since the ketone is oxidised to 5β-etianic acid (formerly known as aetiocholanic acid) with the loss of one carbon atom, the ketone must be a methyl ketone, and so the alkyl group on the α-carbon atom in bisnor-5β-cholanic acid is a methyl group.

Now the carboxyl group in etianic acid is directly attached to the nucleus; this is shown by the following fact. When etianic acid is subjected to one more Barbier–Wieland degradation, a ketone, etiocholanone, is obtained and this, on oxidation with nitric acid, gives a dicarboxylic acid, etiobilianic acid, *without loss of any carbon atoms*. Thus etiocholanone must be a *cyclic* ketone, and so it follows that there are *eight* carbon atoms in the side-chain, which must have the following structure in order to account for the foregoing degradations (see also the end of this section (iii)):



In addition to the Barbier–Wieland degradation, there are also other methods for degrading the side-chain:

(i) Gallagher *et al.* (1946) have introduced a method to eliminate *two* carbon atoms at a time:

$$ArCHMeCH_2CH_2CO_2H \xrightarrow[(ii)\ CH_2N_2]{(i)\ SOCl_2} ArCHMeCH_2CH_2COCHN_2 \xrightarrow{HCl} ArCHMeCH_2CH_2COCH_2Cl \xrightarrow[AcOH]{Zn}$$

$$ArCHMeCH_2CH_2COCH_3 \xrightarrow[(ii)\ -HBr]{(i)\ Br_2} ArCHMeCH{=}CHCOCH_3 \xrightarrow{CrO_3} ArCHMeCO_2H$$

(ii) Miescher *et al.* (1944) have introduced a method to eliminate *three* carbon atoms at a time:

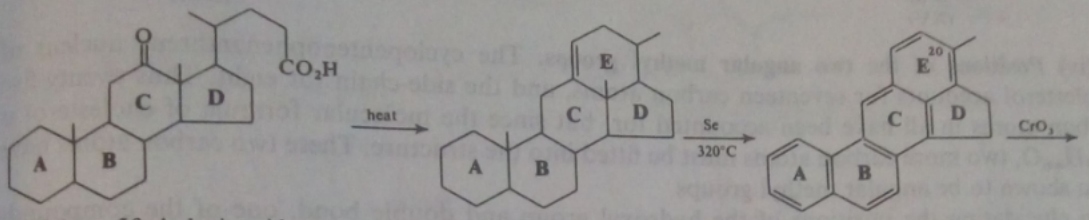
$$ArCHMeCH_2CH_2CO_2Me \xrightarrow{2PhMgBr} ArCHMeCH_2CH_2C(OH)Ph_2 \xrightarrow{-H_2O} ArCHMeCH_2CH{=}CPh_2 \xrightarrow[succinimide]{N\text{-bromo-}}$$

$$ArCHMeCHBrCH{=}CPh_2 \xrightarrow{-HBr} ArCMe{=}CHCH{=}CPh_2 \xrightarrow{CrO_3} ArCOMe$$

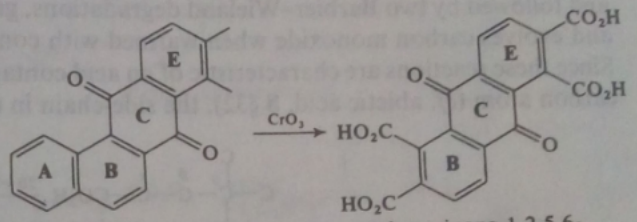
(iii) Jones *et al.* (1958) have carried out the fission of a steroid side-chain with an acid catalyst and have then subjected the volatile products to chromatography. This method has been used with as little as 30 mg of material.

The problem now is: Where is the position of this side-chain? This is partly answered by the following observation. The dicarboxylic acid, etiobilianic acid, forms an anhydride when heated with acetic anhydride. Thus the ketone (etiocholanone) is probably a five-membered ring ketone (in accordance with Blanc's rule), and therefore the side-chain is attached to the five-membered ring D. The actual point of attachment to this ring, however, is not shown by this work. The formation of Diels' hydrocarbon (§1) from cholesterol suggests that the side-chain is at position 17, since selenium dehydrogenations may degrade a side-chain to a methyl group (see **10** §2vii). Position 17 is also supported by evidence obtained from X-ray photographs and surface film measurements. Finally, the following chemical evidence may be cited to show that the position of the side-chain is 17. As we have seen above, 5β-cholanic acid may be obtained by the oxidation of 5β-cholestane. 5β-Cholanic acid may also be obtained by the oxidation of deoxycholic acid (a bile acid; see §14) followed by a Clemmensen reduction. Thus the side-chains in cholesterol and deoxycholic acid are in the same

position. Now deoxycholic acid can also be converted into 12-keto-5β-cholanic acid which, on heating to 320°C, loses water and carbon dioxide to form dehydronorcholene (Wieland *et al.*, 1930). This, when distilled with selenium, forms 20-methylcholanthrene, the structure of which is indicated by its oxidation to 5,6-dimethyl-1,2-benzanthraquinone which, in turn, gives on further oxidation, anthraquinone-1,2,5,6-tetracarboxylic acid (Cook, 1933). Finally, the structure of 20-methylcholanthrene has been confirmed by synthesis (see **10** §5b). The foregoing facts can be explained only if the side-chain in cholesterol is in position 17; thus:



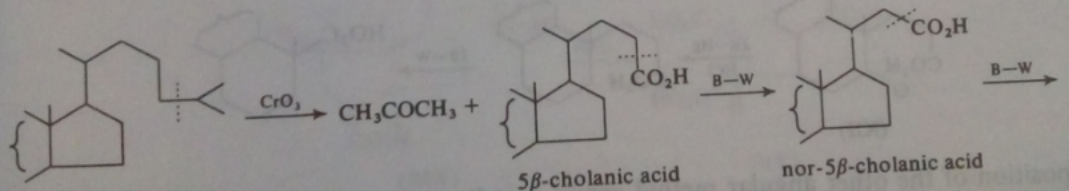
12-keto-5β-cholanic acid — dehydronorcholene — 20-methylcholanthrene



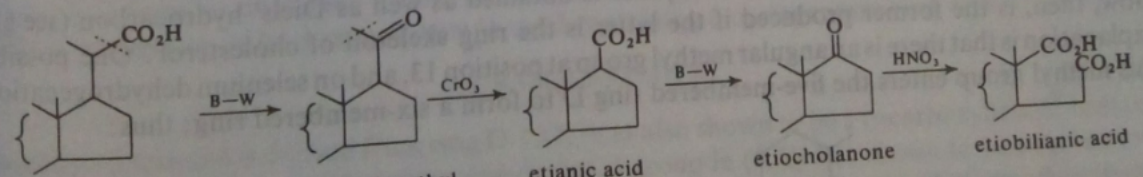
5,6-dimethyl-1,2-benzanthraquinone — anthraquinone-1,2,5,6-tetracarboxylic acid

It should be noted that the isolation of methylcholanthrene affords additional evidence for the presence of the cyclopentenophenanthrene nucleus in cholesterol.

Thus, now that we know the nature and position of the side-chain, we can formulate the conversion of 5β-cholestane into etiobilianic acid as follows:

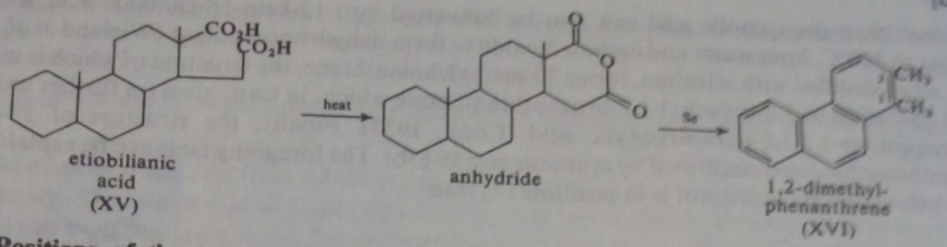


5β-cholestane — 5β-cholanic acid — nor-5β-cholanic acid



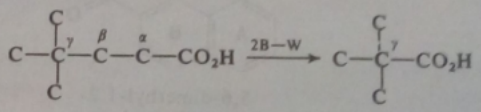
bisnor-5β-cholanic acid — etiocholyl methyl ketone — etianic acid — etiocholanone — etiobilianic acid

A point of interest in this connection is that when the anhydride of etiobilianic acid is distilled with selenium, 1,2-dimethylphenanthrene is obtained (Butenandt *et al.*, 1933). This also provides proof for the presence of the phenanthrene nucleus in cholesterol, and also evidence for the position of the C-13 angular methyl group (see (iv)).

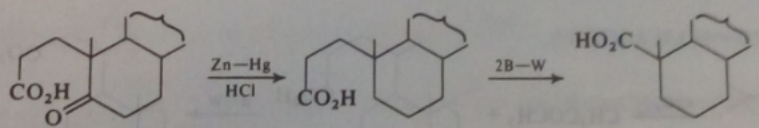


(iv) **Positions of the two angular methyl groups.** The cyclopentenophenanthrene nucleus of cholesterol accounts for seventeen carbon atoms, and the side-chain for eight. Thus twenty-five carbon atoms in all have been accounted for, but since the molecular formula of cholesterol is $C_{27}H_{46}O$, two more carbon atoms must be fitted into the structure. These two carbon atoms have been shown to be angular methyl groups.

In elucidating the positions of the hydroxyl group and double bond, one of the compounds obtained was the keto-acid (XII). This compound, when subjected to the Clemmensen reduction and followed by two Barbier–Wieland degradations, gives an acid which is very difficult to esterify, and evolves carbon monoxide when warmed with concentrated sulphuric acid (Tschesche, 1932). Since these reactions are characteristic of an acid containing a carboxyl group attached to a tertiary carbon atom (*cf.* abietic acid, **8** §32), the side-chain in (XII) must be of the type

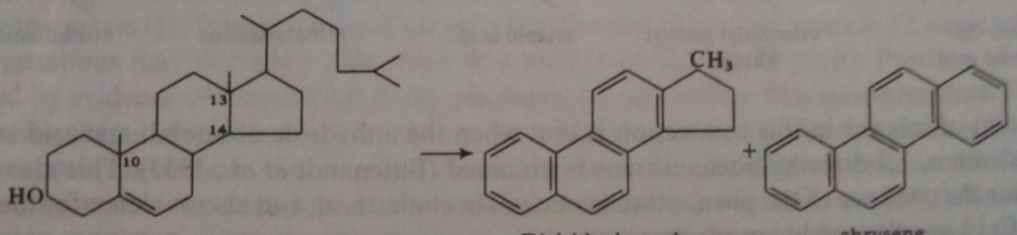


Thus there must be an alkyl group at position 10 in (XII). This could be an ethyl group (as originally believed by Windaus and Wieland) or a methyl group, provided that in the latter case the second 'missing' carbon atom can be accounted for. As we shall see later, there is also a methyl group at position 13, and so the alkyl group at position 10 must be a methyl group. On this basis, the degradation of (XII) may be formulated:



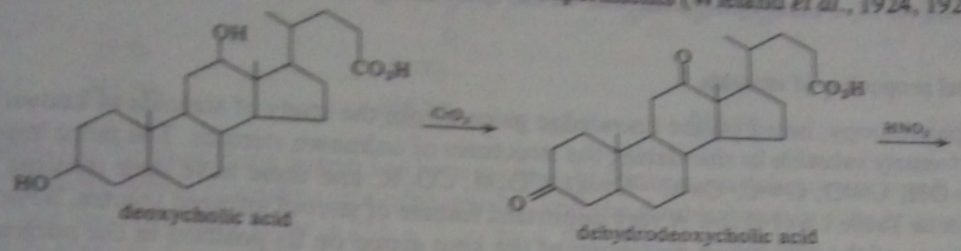
(XII)

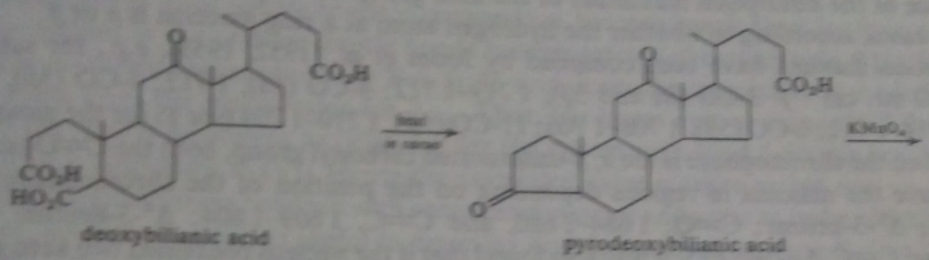
The position of the other angular methyl group is indicated by the following evidence. When cholesterol is distilled with selenium, chrysene is obtained as well as Diels' hydrocarbon (see §1). How, then, is the former produced if the latter is the ring skeleton of cholesterol? One possible explanation is that there is an angular methyl group at position 13, and on selenium dehydrogenation, this methyl group enters the five-membered ring D to form a six-membered ring; thus:

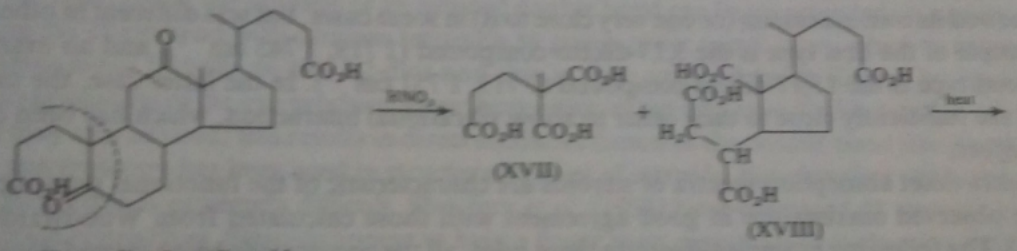


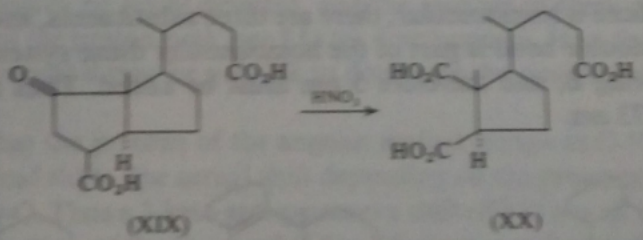
cholesterol — Diels' hydrocarbon — chrysene

This evidence, however, is not conclusive, since ring expansion could have taken place had the angular methyl group been at position 14. Further support for the positions of the two angular methyl groups is given by the following degradative experiments (Wieland *et al.*, 1924, 1928, 1933):









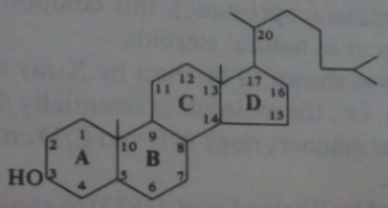
(XVII) was shown to be butane-2,2,4-tricarboxylic acid; thus there is a methyl group at position 10. (XVIII) was shown to be a tetracarboxylic acid containing a cyclopentane ring with a side-chain

$$—CH(CH_3)CH_2CH_2CO_2H.$$

Thus this compound is derived from ring D. (XX) was also shown to be a tricarboxylic acid containing a cyclopentane ring. Furthermore, one carboxyl group in (XX) was shown to be attached to a tertiary carbon atom, and so it follows that there is a methyl group at 13 or 14. (XX) was then shown to have the *trans* configuration, *i.e.*, the two carboxyl groups are *trans*. Thus its precursor (XIX) must have its two rings in the *trans* configuration (the methyl group and hydrogen atom at the junction of the rings are thus *trans*). Theoretical considerations of the strain involved in the *cis*- and *trans*-forms of (XIX) suggest that the *cis*-form of (XIX) would have been obtained had the methyl group been at position 14. Thus the position of this angular methyl group appears (from this evidence)

§5. Stereochemistry of the steroids

If we examine the fully saturated sterol, we find that there are eight dissimilar chiral centres in the nucleus (3, 5, 8, 9, 10, 13, 14 and 17). Thus there are $2^8 = 256$ optical isomers possible. If we also include the chiral centre in the side-chain (20), then there are 512 optical isomers possible.



The stereoisomerism of the steroids is conveniently classified into two types, one dealing with the way in which the rings are fused together, and the other with the configurations of substituent groups, particularly those at C-3 and C-17.

Configuration of the nucleus. There are six chiral centres in the nucleus (5, 8, 9, 10, 13 and 14), and therefore there are $2^6 = 64$ optically active forms theoretically possible. In practice, however, many of these cannot exist because of steric limitations.

§9. Synthesis of cholesterol

Before describing the synthesis of cholesterol, we shall discuss the problem of the synthesis of complex molecules in general. Many examples of these syntheses have already been described (see Ch. 8, Terpenoids. Ch. 9, Carotenoids). Two difficulties of the classical chemists were the isolation of pure compounds from natural sources and the separation of isomers (usually geometrical and optical) formed in the various steps of a synthesis. Modern methods of separation, particularly chromatography, have overcome these problems. Also, recent syntheses have been more successful and more elegant due to the increased knowledge of reaction mechanisms and to the introduction of selective reagents.

An interesting development in the presentation of recent syntheses is the *discussion* of the reasons that led to the adoption of the sequence of steps for carrying out the synthesis. Classical chemists obviously also had their reasons for carrying out their syntheses in a particular way, but these are not often described or are only briefly mentioned in their publications.

A characteristic feature of recent syntheses is the use of control elements. These may be divided into two types: **regiospecific** or **regioselective control elements**, and **stereospecific** or **stereoselective control elements**. The terms 'specific' and 'selective' are used in the sense described in **4** §5k. Regiospecific control elements are groups which have been deliberately introduced to cause reactions to occur at a specific site in a molecule and, if necessary, can be readily removed without affecting the rest of the molecule. Stereospecific control elements are those which cause a reaction to proceed in such manner that the product has one particular type of geometry rather than another. Control elements were used by the classical chemists, but many more of these elements have now been introduced. Some examples of their application have already been described, *e.g.*, regiospecific: protecting groups, activating of a methylene group by an adjacent oxo group; stereospecific: asymmetric synthesis (more correctly this is an example of stereoselectivity), stereochemical control by steric effects, addition and elimination reactions.

A simple molecule may be described as one which is small and whose total synthesis requires a relatively small number of steps. Very often, such a synthesis may be readily achieved by 'working backwards'. On the other hand, a complex molecule may be described as a large molecule whose total synthesis requires a large number of steps. Furthermore, the synthesis of a complex molecule usually involves problems of stereochemistry. It is important to note, however, that success in achieving a synthesis, be it of a simple or a complex molecule, ultimately depends on a very good knowledge of organic reactions and their application.

Some points that may be noted for the general approach to the synthesis of complex molecules are (see the appropriate reading references):

(i) The recognition of structural units within the molecule which can be formed and/or assembled by known chemical methods. Starting materials should be readily accessible. The first objective is assisted by examination of the molecule (to be synthesised) for any type of symmetry. Recognition of symmetry will lead to a shorter route. Structural units within a molecule are termed 'synthons', and their recognition may suggest routes for the synthesis. Furthermore, recognition of a relationship of the molecule to some other *known* compound may permit the use of a complicated synthon if the known compound is readily available.

(ii) The necessity of obtaining the best yields of the products is of paramount importance, and to achieve this may require the use of control elements.

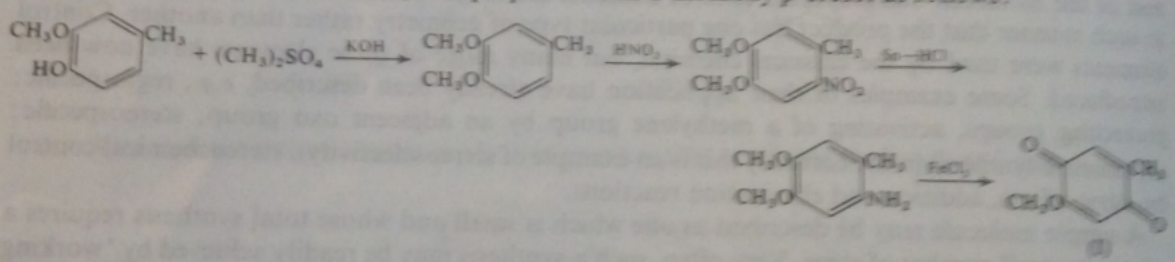
(iii) The relative positions of chiral centres (when present) may give information on the type of control elements required to give the desired configurations.

(iv) The presence of reactive functional groups which can give rise to neighbouring group participation may suggest steps that lead to a desired intermediate, *e.g.*, by temporary cyclisation and so controlling the stereochemical course of the reaction.

We shall now discuss the synthesis of cholesterol and consider it in the light of the above discussion. Basically, the synthesis of steroids involves the construction of the steroid nucleus in the form of the required conformation. The early methods started with ring A or rings A,B, and the other rings were then built up as follows: A → AB → ABC → ABCD. However, as the number of selective reagents increased, different starting points and different orders of fusion were developed, *e.g.*, (i) AB → ABCD; (ii) AC → ABC → ABCD; (iii) AD → ABCD; (iv) BC → BCD → ABCD; (v) CD → ACD → ABCD; (vi) CD → BCD → ABCD.

Two groups of workers, *viz.*, Robinson *et al.* (1951) and Woodward *et al.* (1951), have synthesised cholesterol. One of the outstanding difficulties in the synthesis of steroids is the stereochemical problem. The cholesterol nucleus contains eight chiral centres and so 256 optical isomers are possible (see also §4 for further details). Thus every step in the synthesis which produced a new chiral centre had to result in the formation of some (the more the better) of the desired stereoisomer, and at the same time resolution of racemic modifications also had to be practicable. Another difficulty was attacking a particular point in the molecule without affecting the other parts. This problem led to the development of specific reagents. The following is an outline of the Woodward synthesis. Some steps are not stereospecific or even stereoselective. Later syntheses of various steroids are superior in this respect (see, *e.g.*, aldosterone, §28b). The synthesis of cholesterol described here is of the type: C → CD → BCD → ABCD.

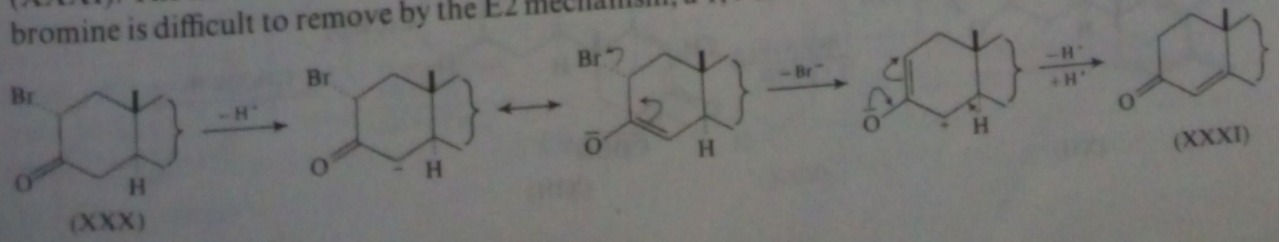
4-Methoxy-2,5-toluquinone (I) was prepared from 2-methoxy-*p*-cresol as follows:



(I)

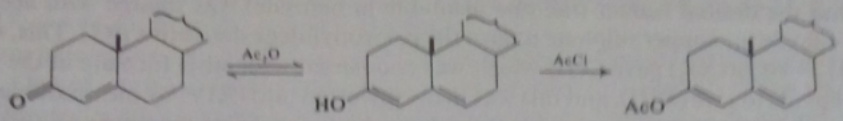
(I) was condensed with butadiene (Diels–Alder reaction) to give (II). This has the *cis* configuration and was isomerised (quantitatively) to the *trans*-isomer (III) by dissolving in aqueous alkali, adding a seed crystal of the *trans*-isomer and then acidifying. Isomerisation occurs *via* the enolate to give the more stable *trans*-isomer (see also §5; configuration of the nucleus). (III), on reduction with lithium aluminium hydride, gave (IV). (IV) is a vinyl ether of a glycol which, on treatment with aqueous acid, undergoes hydrolysis (demethylation) to give a β-hydroxyketone which is readily dehydrated to (V) in acid solution. Conversion of (V) to (VI) by removal of the hydroxyl group was carried out by a new technique: (V) was acetylated and the product, the ketol acetate, was heated with zinc in acetic anhydride to give (VI) [reduction with metal and acid usually reduces α,β-unsaturated bonds in ketones]. (VI), on treatment with ethyl formate in the presence of sodium methoxide, gave the hydroxymethylene ketone (VII) [Claisen condensation]. When this was treated with ethyl vinyl ketone in the presence of potassium t-butoxide, (VIII) was formed (Michael condensation). The object of the double bond in the ketone ring in (VI) is to prevent formylation occurring on that side of the keto group, and the purpose of the formyl group is to produce an active methylene group (this is now flanked on *both* sides by carbonyl groups). The necessity for this "activation" lies in the fact that ethyl vinyl ketone tends to self-condense, and consequently decrease the yield of (VIII). Both operations are examples of the introduction of regiospecific control elements. (VIII) was now cyclised quantitatively by means of potassium hydroxide in aqueous dioxan to the single product (IX). This is the desired compound; the other possible isomer ((IX) with the two hydrogens *cis* instead of *trans* as shown) is not formed since the *cis*-isomer is less stable than the *trans* due to greater steric interactions in the former, *i.e.*, the cyclisation is stereospecific (steric effect control). Also, the cyclisation

occurs by an intramolecular aldol condensation followed by dehydration. (IX) was then treated with osmium tetroxide to give two *cis*-glycols of structure (X) [one is *cis* with respect to the angular methyl group and the other is *trans*]. Glycol formation occurs readily at the isolated double bond (the other two double bonds are conjugated and so have less double bond character than an isolated double bond; the reaction with osmium tetroxide is very sensitive to this change). These glycols were separated and the desired isomer (the one insoluble in benzene) was treated with acetone in the presence of anhydrous copper sulphate to give the isopropylidene derivative (XI). This, on catalytic reduction (H_2—Pd/$SrCO_3$) gave (XII) which was condensed with ethyl formate in the presence of sodium methoxide to give (XIII), and this was then converted into (XIV) by means of methylaniline. The purpose of this treatment was to block undesired condensation reactions on this side of the keto group (at this position 3); this is another example of a regiospecific control element. When (XIV) was condensed with vinyl cyanide (cyanoethylation) and the product hydrolysed with alkali, the product was a mixture of two keto acids. These were separated and the stereoisomer (XV) [methyl group in front and propionic acid group behind the plane of the rings] was converted into the enol lactone (XVI) which, on treatment with methylmagnesium bromide, gave (XVII), and this, on ring closure by means of alkali, gave (XVIII). When this was oxidised with periodic acid in aqueous dioxan, the dialdehyde (XIX) was obtained (*via* hydrolysis of the diol), and this, when heated in benzene solution in the presence of a small amount of piperidine acetate, gave (XX) [and a small amount of an isomer]. This cyclisation occurs by an intramolecular aldol condensation under the influence of the base, piperidine acetate. Since either aldehyde group can be involved in the condensation, two products are possible. In (XIX), the upper methylene group is *cis* to the hydrogen atom at C-14, whereas the lower methylene group is *cis* to the 18-methyl group. Hence, the upper methylene group experiences less steric hindrance than the lower one and consequently it is the former that loses a proton to form the carbanion. Therefore (XX) is the predominant isomer. (XX) was oxidised to the corresponding acid which was then converted into the methyl ester (XXI) with diazomethane. (XXI), a racemate, was resolved by reduction of the keto group with sodium borohydride to the hydroxy esters [(±)-3α- and (±)-3β-]. The (+)-form of the 3β-alcohol was preferentially precipitated by digitonin, and this stereoisomer was now oxidised (Oppenauer oxidation) to give the desired stereoisomer (+)-(XXI). This was catalytically reduced (H_2—Pt) to (XXII), which was then oxidised to (XXIII) which was now a mixture of stereoisomers (from the mixture of (XXII); H at 17 behind and in front). These were separated, reduced (sodium borohydride), and hydrolysed. The β-isomer, (XXIV), was converted into the methyl ketone by first acetylating, then treating with thionyl chloride and finally with dimethylcadmium. This acetylated hydroxyketone, (XXV), on treatment with isohexylmagnesium bromide, gave (XXVI). This was a mixture of isomers (a new chiral centre has been introduced at position 20). (XXVI), on dehydration, gave one product, (XXVII), and this, on catalytic hydrogenation (H_2—Pt), gave a mixture of 5α-cholestanyl acetates (the chiral C-20 has been re-introduced). These acetates were separated and the desired isomer, on hydrolysis, gave 5α-cholestan-3β-ol, (XXVIII), which was identical with natural cholestanol. The conversion of cholestanol into cholesterol (XXXIII) is then carried out by a series of reactions introduced by various workers. Bromination of (XXIX) in acetic acid in the presence of hydrogen bromide (as catalyst) gives the 2α-bromo-derivative ((XXX); see §8). (XXX), on treatment with pyridine, gives (XXXI). The mechanism of this elimination is uncertain. A possibility is that because the *equatorial* bromine is difficult to remove by the E2 mechanism, a 1,4-elimination occurs by removal of a proton



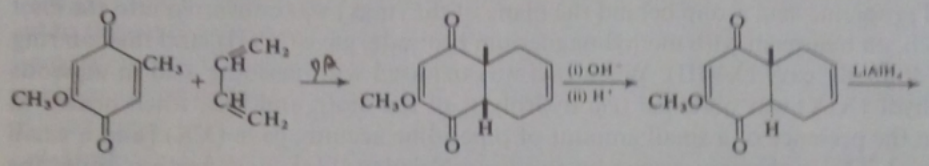
from position 4 by the base (the methylene group in this position is activated by the adjacent oxo group; *cf.* however, the bromination of acetone).

Heating (XXXI) with acetyl chloride in the presence of acetic anhydride gives the enol acetate (XXXII) which, on reduction with lithium aluminium hydride followed by acidification, gives cholesterol (XXXIII). The mechanism of this reaction is uncertain.

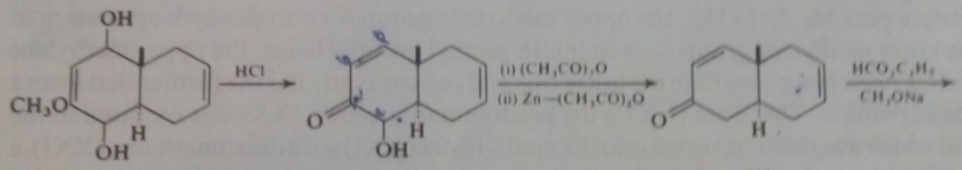


(XXXI) (XXXII)

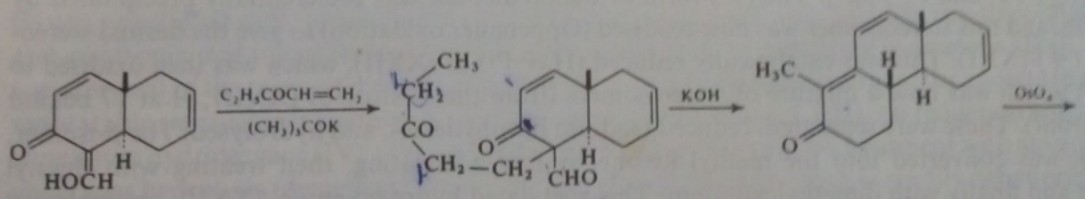
Possibly the electron-attracting effect of the acetoxy-group activates the 3,4-double bond to hydride transfer from the lithium aluminium hydride.



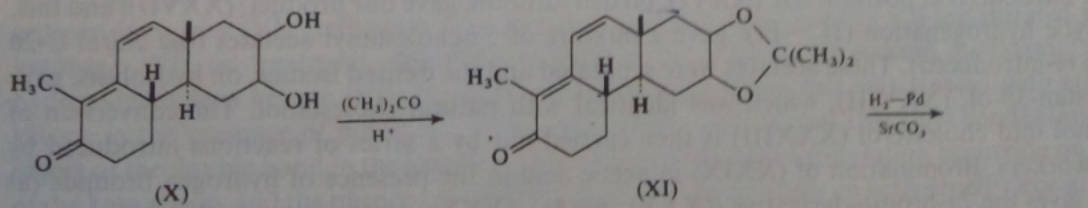
(I) (II) (III)



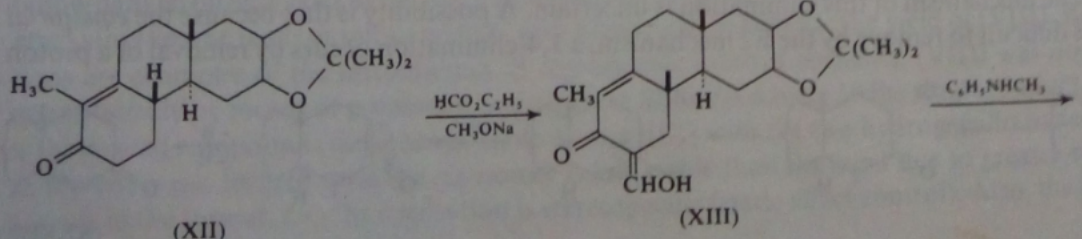
(IV) (V) (VI)



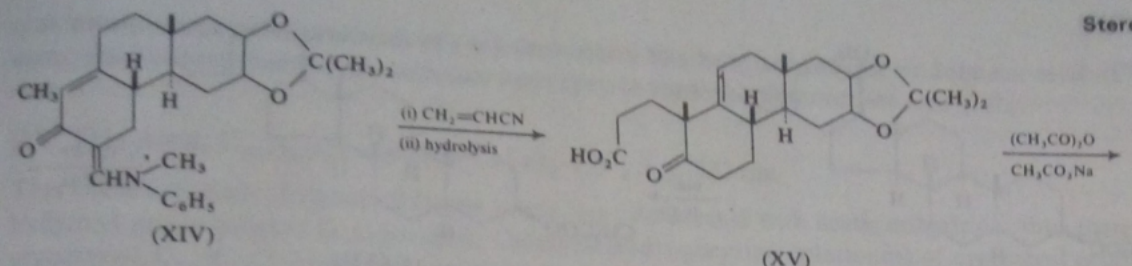
(VII) (VIII) (IX)



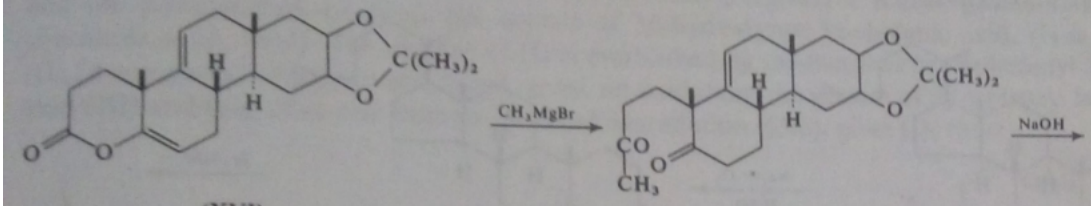
(X) (XI)



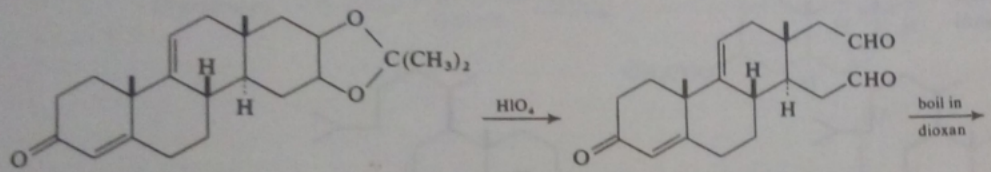
(XII) (XIII)



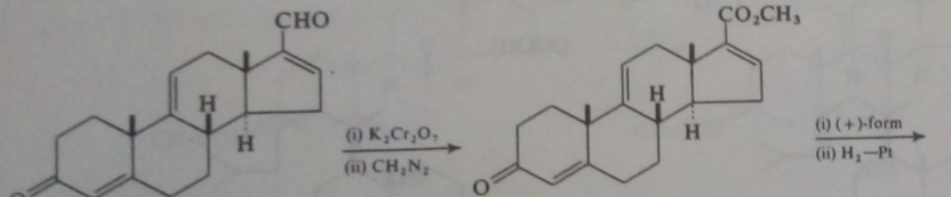
(XIV) (XV)



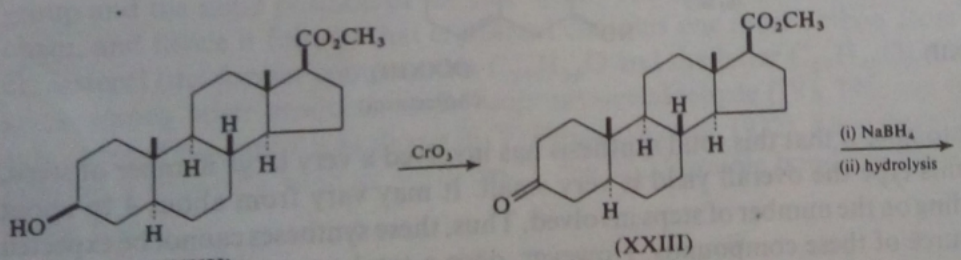
(XVI) (XVII)



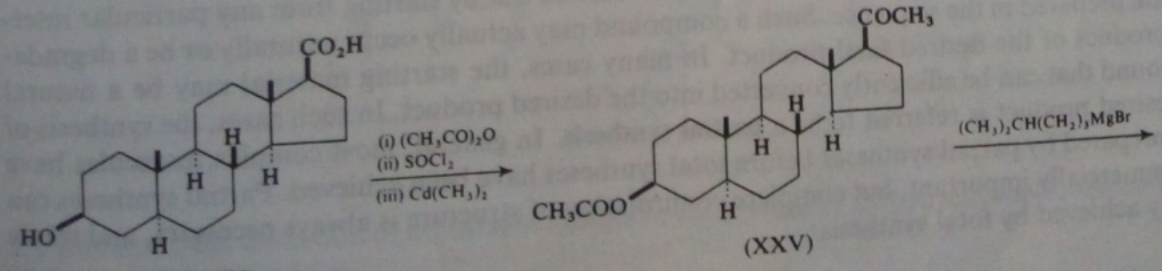
(XVIII) (XIX)



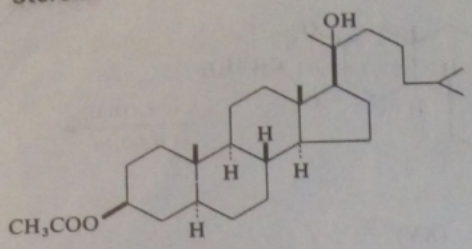
(XX) (XXI)



(XXII) (XXIII)

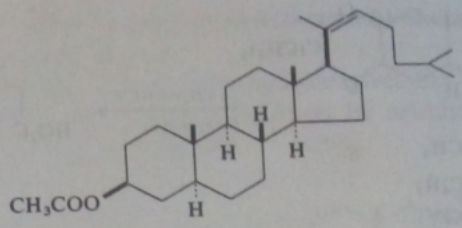


(XXIV) (XXV)



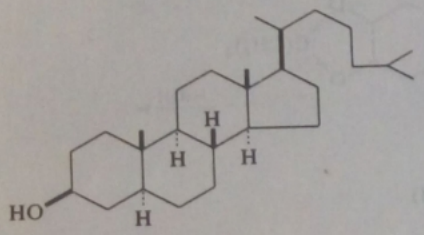
(XXVI)

heat
(−H2O)



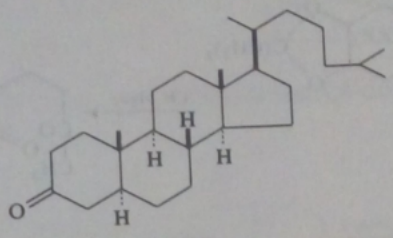
(XXVII)

(i) H2—Pt
(ii) NaOH



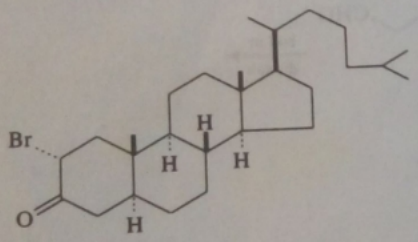
(XXVIII)
cholestanol

Na2Cr2O7
H2SO4



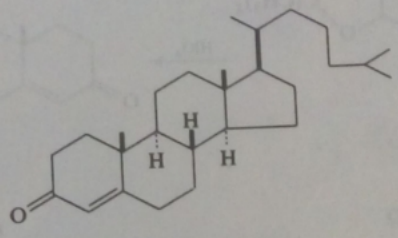
(XXIX)

Br2/HBr



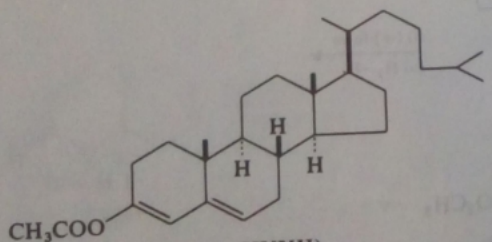
(XXX)

pyridine



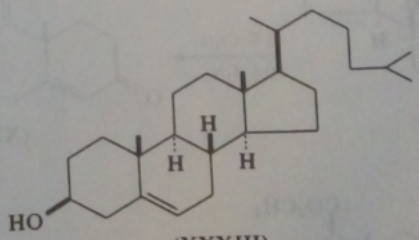
(XXXI)

CH3COCl
(CH3CO)2O



(XXXII)

(i) LiAlH4
(ii) HCl



(XXXIII)
cholesterol

involved a very large number of steps,

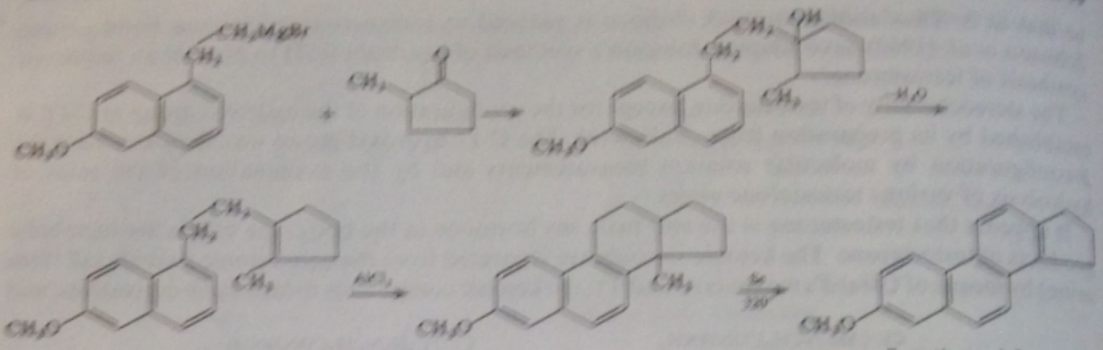
OESTROGENS

§20. Oestrone (estrone)

It has been known for a long time that there are hormones which control the uterine cycle, but it was not until 1929 that Butenandt and Doisy independently isolated the active substance **oestrone** from the urine of pregnant women. Oestrone is the first known member of the sex hormones, and soon after its discovery two other hormones were isolated, oestriol and oestradiol.

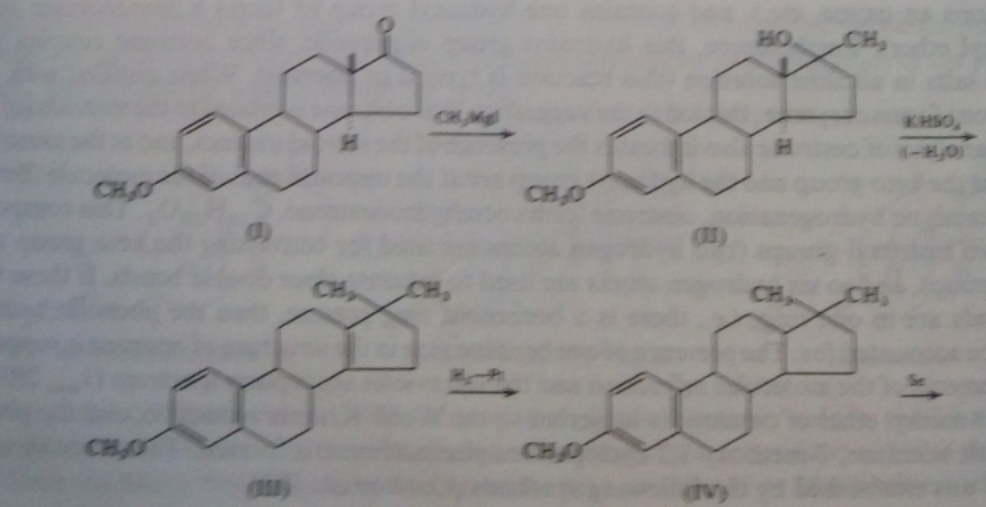
(+)-Oestrone, m.p. 259°C, $[\alpha]_D$ +170°, has the molecular formula $C_{18}H_{22}O_2$. It behaves as a ketone (forms an oxime, etc.), and contains one hydroxyl group (it forms a monoacetate and a monomethyl ether). Furthermore, this hydroxyl group is *phenolic*, since oestrone couples with diazonium salts in alkaline solution (this reaction is typical of phenols). When distilled with zinc dust, oestrone forms chrysene; this led to the suggestion that oestrone is related to the steroids (*cf.* §1). The X-ray analysis of oestrone also indicates the presence of the steroid nucleus, and at the same time showed that the keto group and the hydroxyl group are at the opposite ends of the molecule (Bernal, 1932). On catalytic hydrogenation, oestrone forms octahydrooestrone, $C_{18}H_{30}O_2$. This compound contains two hydroxyl groups (two hydrogen atoms are used for converting the keto group to an alcoholic group), and so six hydrogen atoms are used to saturate *three* double bonds. If these three double bonds are in one ring, *i.e.*, there is a benzenoid ring present, then the phenolic hydroxyl group can be accounted for. The presence of one benzene ring in the structure of oestrone is supported by measurements of the molecular refraction and the ultraviolet absorption spectrum (λ_{max} 280 nm).

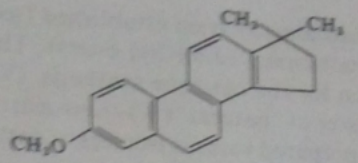
When the methyl ether of oestrone is subjected to the Wolff–Kishner reduction, and the product distilled with selenium, 7-methoxy-1,2-cyclopentenophenanthrene is formed. The structure of this compound was established by the following synthesis (Cook *et al.*, 1934):



7-methoxy-1,2-cyclopentenophenanthrene

Thus the benzene ring in oestrone is ring A, and the (phenolic) hydroxyl group is at position 3; hence the skeleton of oestrone is as shown. Into this skeleton we must fit the keto group, and since this skeleton contains only 17 carbon atoms, another carbon atom must also be placed. The position of the keto group was shown to be at 17, and the extra carbon atom was shown to be an angular methyl group at position 13, as follows (Cook *et al.*, 1935). When the methyl ether of oestrone (I) is treated with methylmagnesium iodide, compound (II) is obtained. When (II) is dehydrated with potassium hydrogen sulphate to (III), this catalytically reduced to (IV) and then (IV) distilled with selenium, the product is 7-methoxy-3′,3′-dimethyl-1,2-cyclopentenophenanthrene (V). The formation of (V) can be explained only if there is a keto group at position 17 and an angular methyl group at position 13. It should be noted that in the given equations, the dehydration is accompanied by the migration of the angular methyl group; this assumption is based on the analogy with known examples in which this occurs. Furthermore, this migration of a methyl group is characteristic of *trans*-fused hydrindanols of type (II), and so the configuration of rings C/D is *trans* (*cis*-C/D fusion leads to dehydration without rearrangement). In the *trans*-C/D fusion, the CH_3-18 group is in the axial position and so satisfies the stereoelectronic requirements for the 1,2-migration with loss of the hydroxyl group at C-17.

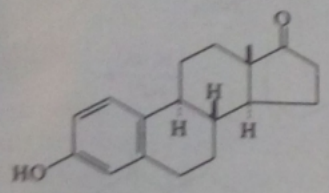




(V)

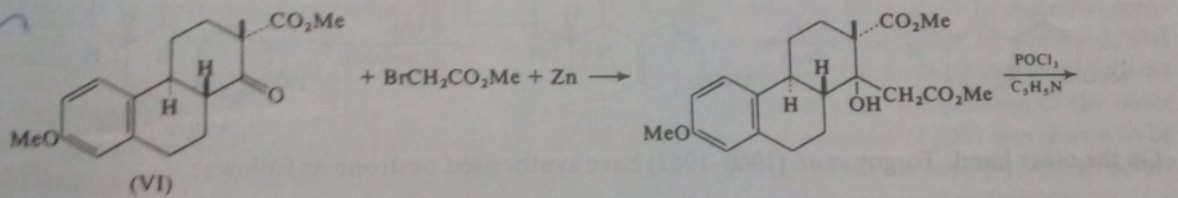
The structure of (V) has been confirmed by synthesis (Cook *et al.*, 1935). Thus the structure of oestrone is as shown (see also below).

This has been confirmed by the total synthesis of Anner and Miescher (1948). These authors started with the phenanthrene derivative (VI) which had been prepared previously by Robinson *et al.* (1938), and by Bachmann *et al.* (1942). The first step of the Anner–Miescher synthesis involves the Reformatsky reaction, and a later one the Arndt–Eistert synthesis.

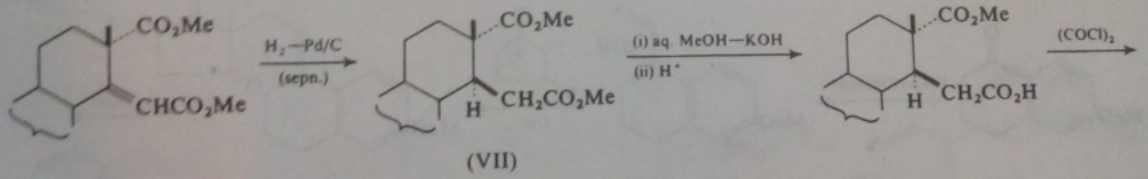


oestrone

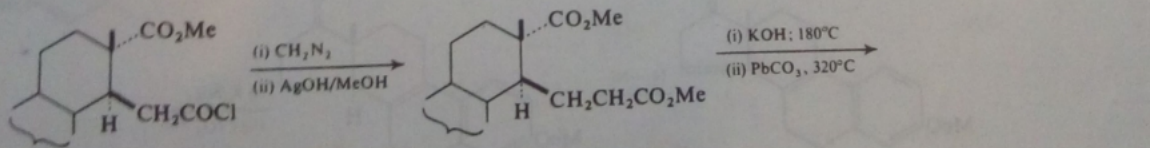
The stereochemical problems involved in the synthesis of oestrone are not so complicated as in cholesterol, since only four chiral centres are present in the hormone (*cf.* §5). (VI) contains 3 chiral centres and so four racemates are possible. Three have been isolated by Anner and Miescher, and one of these was converted into (±)-oestrone (C/D *trans*) and the stereoisomer (C/D *cis*), (±)-iso-oestrone. These were separated and the (±)-oestrone resolved with (−)-menthoxyacetic acid. The (+)-enantiomer that was obtained was shown to be identical with the natural compound. The *trans*-B/C fusion of the racemate used (for the oestrone synthesis) was deduced from other synthetic work, and the

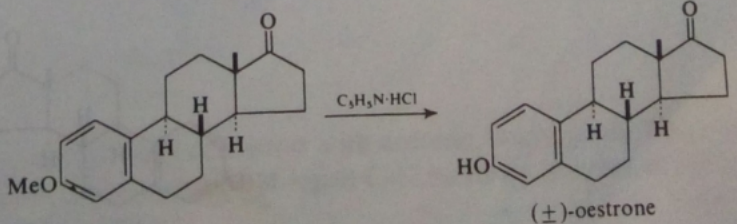


(VI)



(VII)



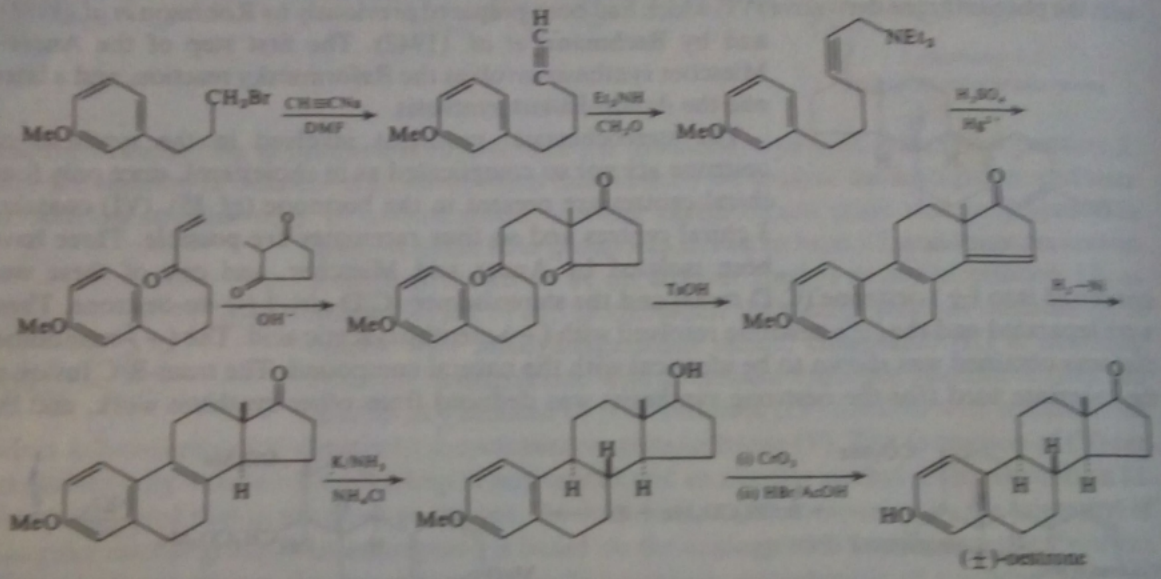


(±)-oestrone

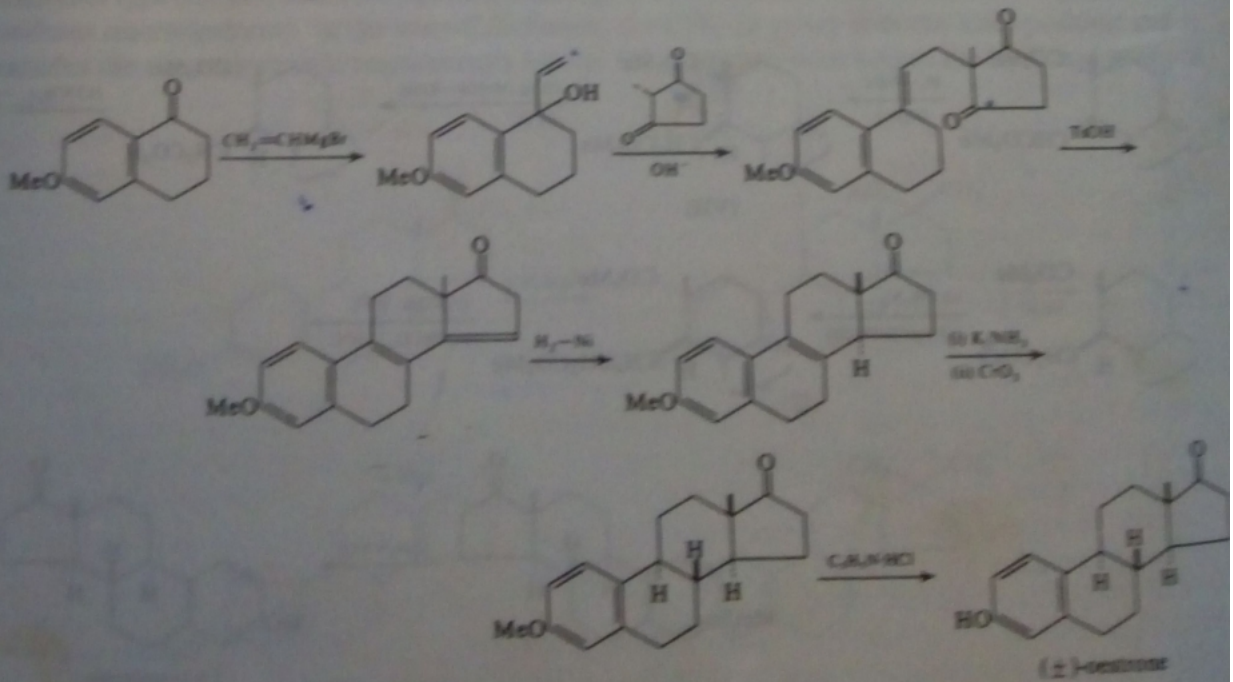
β-configuration of the CH_3-18 had already been established (see above). The catalytic reduction step produced a mixture of stereoisomers (dimethyl esters). These were separated by fractional crystallisation and the one chosen for the oestrone synthesis, (VII), was that which was identical with the methyl ether dimethyl ester of 'natural' (+)-*trans*-marrianolic acid (see formula II, §21).

Miescher and Anner have also prepared various isomers of oestrone by using other stereoisomers of (VI) and (VII), *e.g.*, (±)-iso-oestrone (C/D *cis*).

Johnson *et al.* (1958, 1962) have also carried out a total synthesis of oestrone; each step in their synthesis was stereospecific, but Hughes *et al.* (1960) have described total syntheses of oestrone which appear to be simpler than any previous method and just as efficient. The better method is as follows and involves a Mannich reaction and a Michael condensation (see Vol. I).

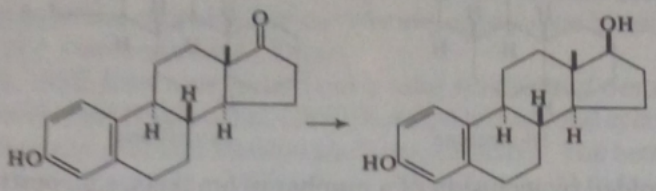


On the other hand, Torgov *et al.* (1960–1962) have synthesised oestrone as follows:



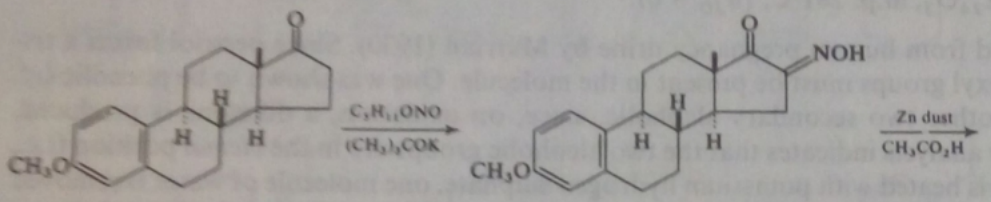
The chemical relationship between oestrone, oestriol and oestradiol (§22) is shown by the following reactions.

(i) Oestrone may be reduced to oestradiol by catalytic hydrogenation, by aluminium isopropoxide (the Meerwein–Ponndorf–Verley reduction), or by lithium aluminium hydride.

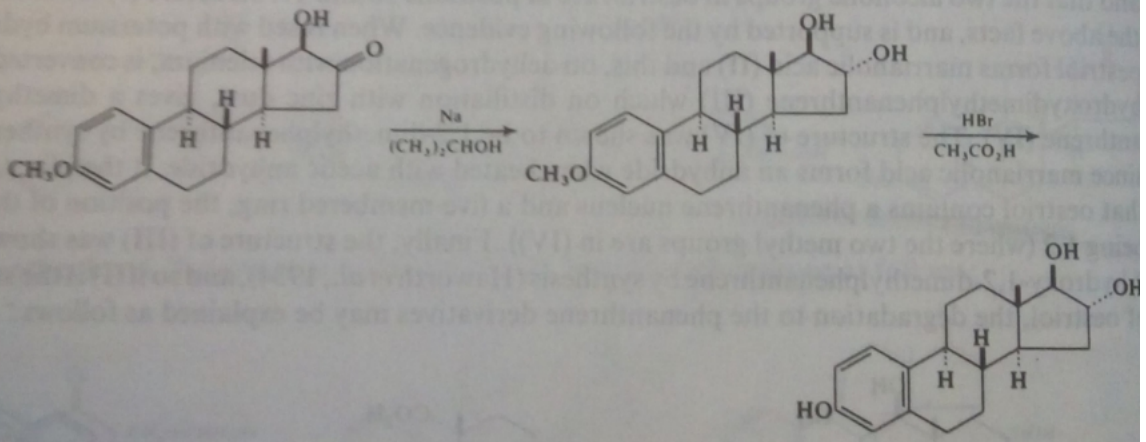


oestrone oestradiol

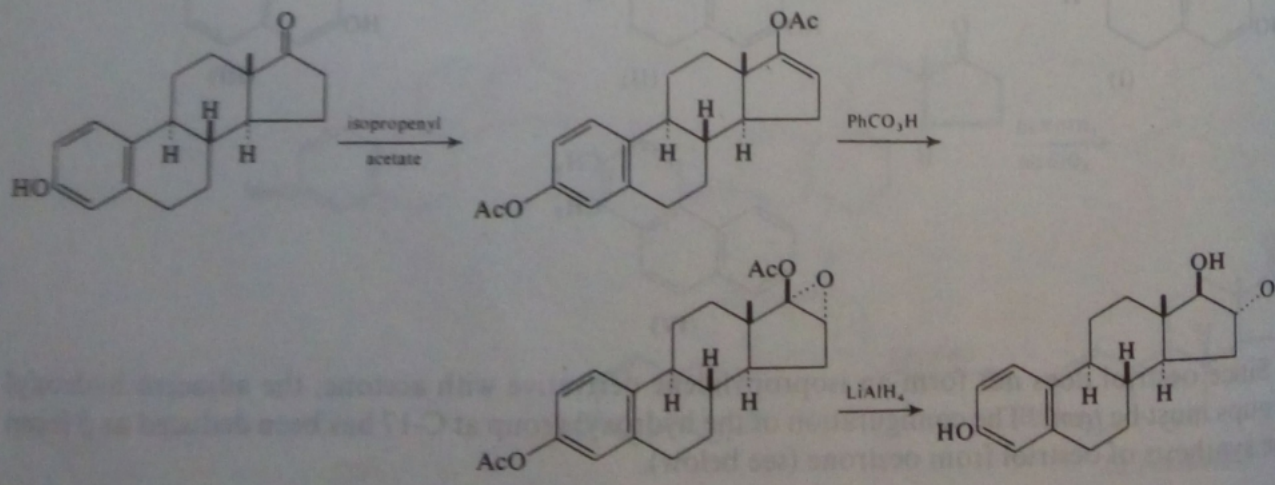
(ii) Oestriol may be converted into oestrone by the action of potassium hydrogen sulphate (see above), and oestrone may be converted into oestriol as follows (Huffman *et al.*, 1947, 1948).



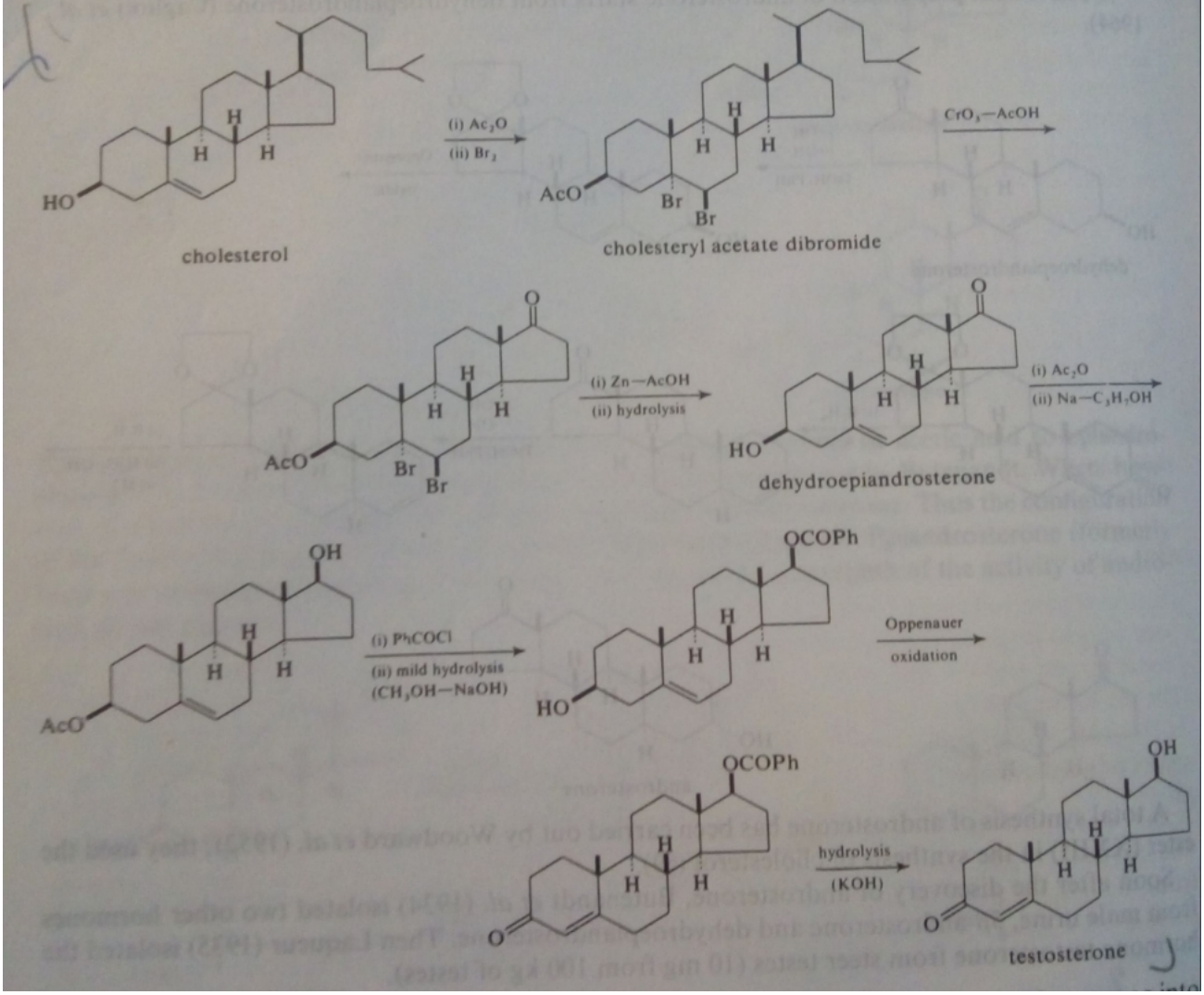
methyl ether of oestrone



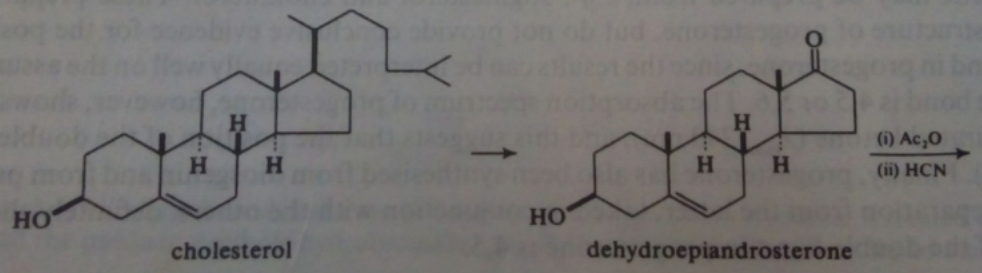
Leeds *et al.* (1954) have converted oestrone into oestriol by a simpler method:



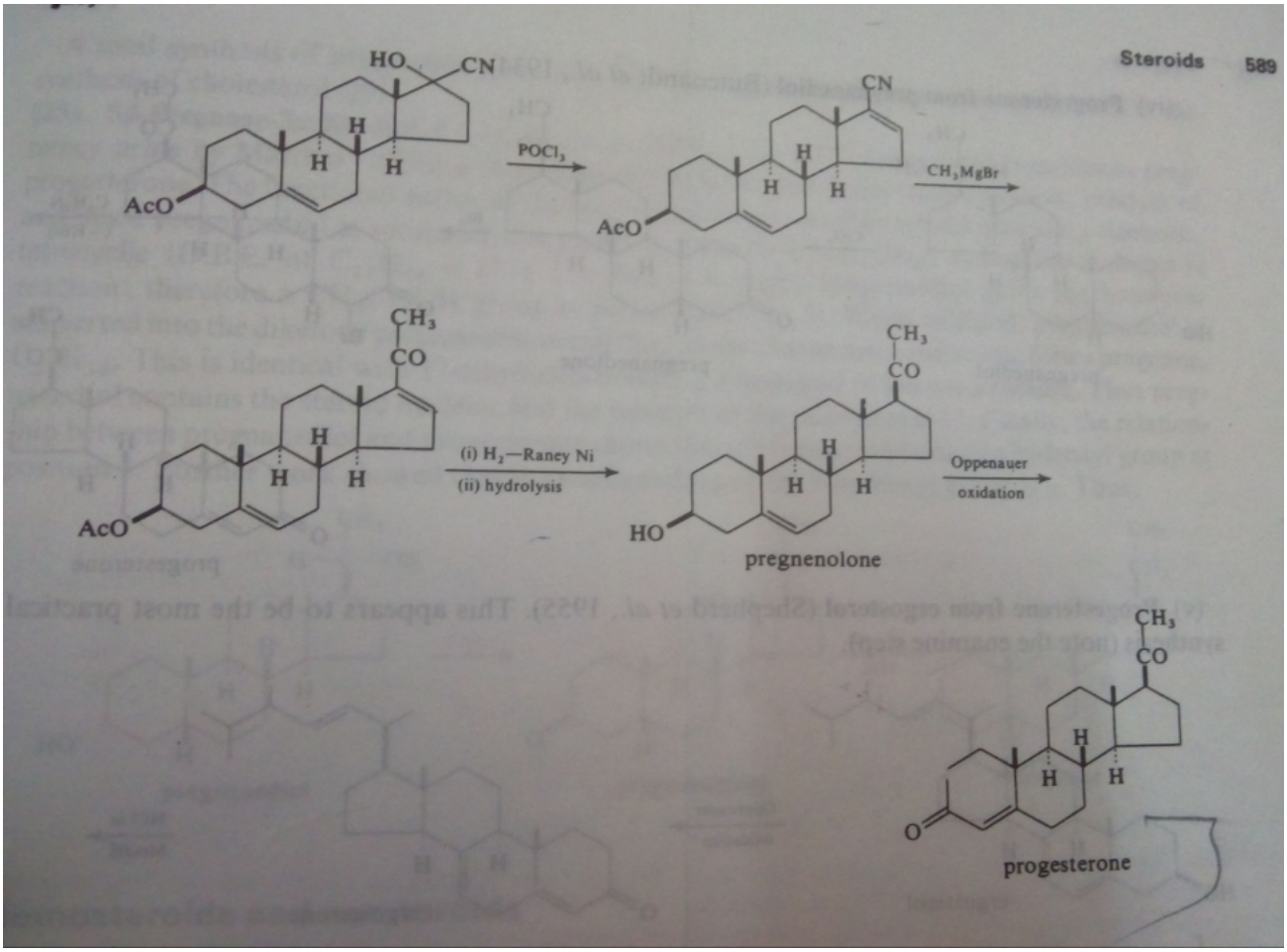
This preparation of testosterone establishes the structure of this hormone which had been shown to contain one hydroxyl group and an α,β-unsaturated ketone group.



(ii) **Progesterone from cholesterol** (Butenandt *et al.*, 1939). Cholesterol is first converted into dehydroepiandrosterone (see §19), and then as follows:

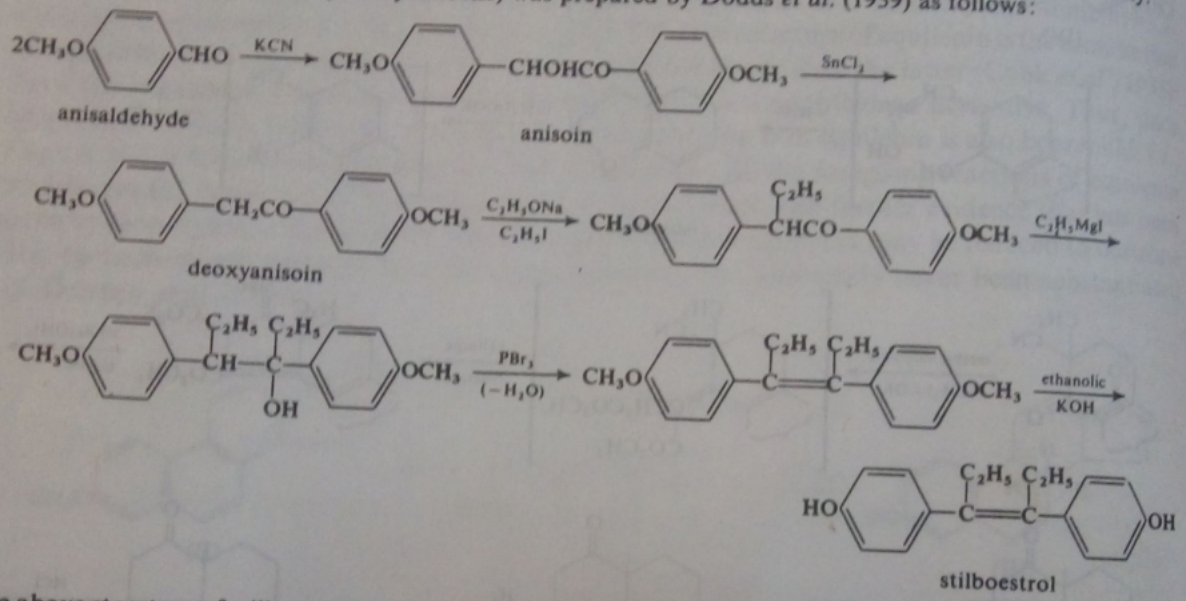


cholesterol dehydroepiandrosterone

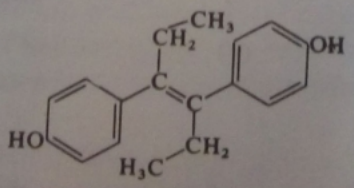

HO
CN
AcO
H
H
H
POCl3
CN
AcO
CH3MgBr
CH3
CO
(i) H2—Raney Ni
(ii) hydrolysis
HO
Oppenauer
oxidation
pregnenolone
O
progesterone

§24. Artificial hormones

Many compounds with oestrogenic activity but not of steroid structure have been prepared synthetically. **Stilboestrol** (4,4′-dihydroxydiethylstilbene) was prepared by Dodds *et al.* (1939) as follows:



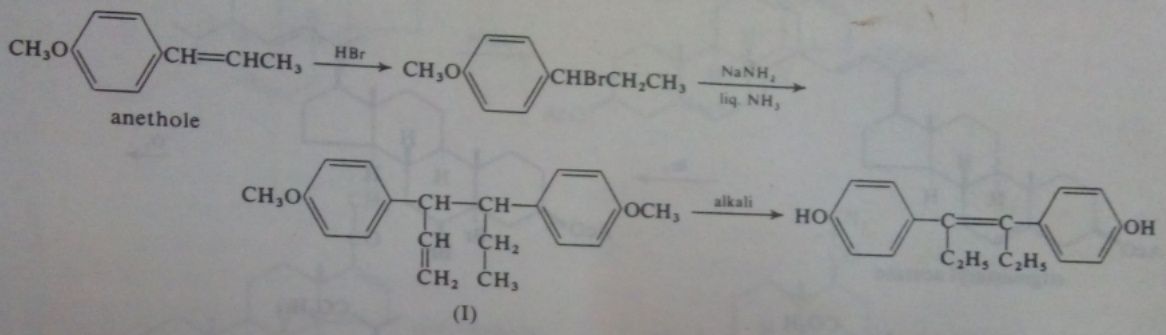
The above structure of stilboestrol can exist in two geometrical isomeric forms; it is the *trans*-form which is the active substance, and this configuration has been confirmed by X-ray analysis (Crowfoot *et al.*, 1941).



trans-stilboestrol

Kharasch *et al.* (1943) have introduced a simpler synthesis of stilboestrol. Anethole is treated with hydrobromic acid and the product, anethole hydrobromide, is then treated with sodamide in liquid ammonia. The

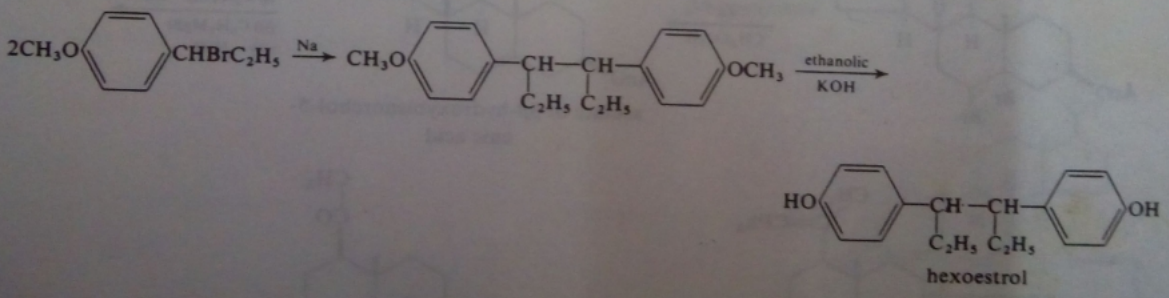
resulting compound (I) gives stilboestrol on demethylation and isomerisation in the presence of alkali. The structure of (I) is uncertain, but it is believed to be the one given.



(I)

Stilboestrol is more active than oestrone when administered subcutaneously, and it can also be given orally.

Hexoestrol (dihydrostilboestrol) may be prepared from anethole hydrobromide as follows:



hexoestrol

The active form is the *meso*-isomer (as shown by X-ray crystallography by Crowfoot *et al.*, 1941).