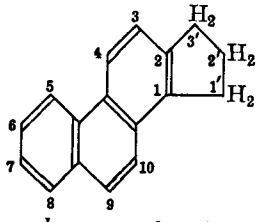
CHAPTER XI

STEROIDS

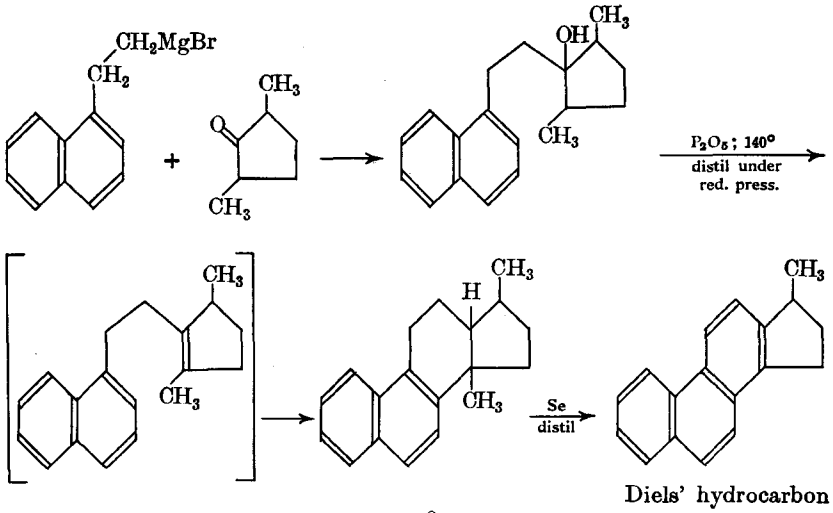
§1. Introduction. The steroids form a group of structurally related compounds which are widely distributed in animals and plants. Included in the steroids are the sterols (from which the name *steroid* is derived), vitamin D, the bile acids, a number of sex hormones, the adrenal cortex hormones, some carcinogenic hydrocarbons, certain sapogenins, etc. The structures of the steroids are based on the 1 : 2-*cyclo*pentenophenanthrene skeleton (Rosenheim and King, 1932; Wieland and Dane, 1932). All the steroids



1:2-*cyclo*pentenophenanthrene

give, among other products, Diels' hydrocarbon on dehydrogenation with selenium at 360° (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°, the steroids give mainly chrysene (§4. X) and a small amount of picene (§4a. X).

Diels' hydrocarbon is a solid, m.p. 126–127°. 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-*cyclo*pentenophenanthrene. This structure for the compound was definitely established by synthesis, *e.g.*, that of Harper, Kon and Ruzicka (1934) who used the Bogert–Cook method [§2 (vi) *e*. X], starting from 2-(1-naphthyl)-ethylmagnesium bromide and 2 : 5-dimethyl*cyclo*pentanone.



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, cholestanol and coprostanol (coprosterol) 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°. 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.

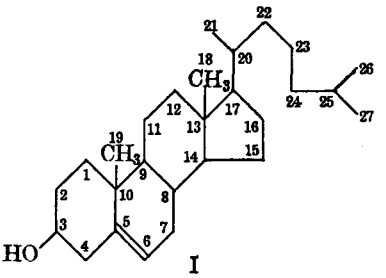
Cholesterol is a white crystalline solid which is optically active (lævorotatory). Cholesterol (and other sterols) gives many colour reactions, *e.g.*,

(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 §19. iii), 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). Digitonide formation is used for the estimation of cholesterol.

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



I

have the established structure of cholesterol at the beginning of our discussion. I is the structure of 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.

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

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.

When cholesterol is distilled with selenium at 360°, 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 *cyclo*pentenophenanthrene 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 §4a) 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 *cyclo*pentenophenanthrene 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. This structure has now been confirmed by the synthesis of cholesterol (see later in this section).

Although an account of the oxidative degradation of the steroids cannot be discussed here, the following points in this connection are of some interest.

(i) The nature of the *nucleus* in sterols and bile acids was shown to be the same, since cholanic acid or *allo*cholanic acid is one of the oxidation products (see §4a for the significance of the prefix *allo*).

(ii) 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, and further work showed that the hydroxyl groups behaved differently towards a given reagent, *e.g.*,

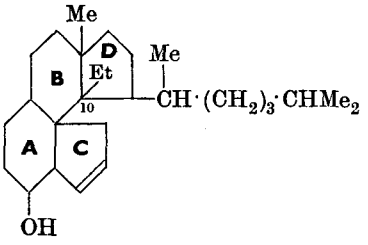
(*a*) The ease of oxidation of hydroxyl groups to keto groups by means of chromic acid is $C_7 > C_{12} > C_3$. More recently, Fonken *et al.* (1955) have shown that *tert.*-butyl hypochlorite apparently oxidises the 3-OH group selectively to the keto group; this reaction, however, failed with cholesterol. Sneedon *et al.* (1955) have also shown that the 3-OH group in steroids is oxidised by oxygen-platinum, but not those at 6, 7 or 12.

(*b*) The three keto groups are not equally readily reduced to a methylene group (by the Clemmensen reduction) or to an alcoholic group (by H_2—platinum). The ease of reduction is $C_3 > C_7 > C_{12}$. This is also the order

for the ease of hydrolysis or acetylation when these positions are occupied by hydroxyl groups (see also testosterone, §13). More recently, it has been shown that the modified Wolff-Kishner reduction of Huang-Minlon (see Vol. I) on steroid ketones reduces keto groups at 3, 7, 12, 17 and 20, but not at 11. Another interesting point in this connection is that lithium aluminium hydride, in the presence of aluminium chloride, does not reduce unsaturated ketones to alcohols, *e.g.*, cholest-4-en-3-one, under these conditions, is reduced to cholest-3-ene (Broome *et al.*, 1956).

Thus a knowledge of (*a*) and (*b*) enabled workers to open the molecule at *different* points by oxidation under the appropriate conditions. This led to a large variety of degradation products, the examination of which enabled the nature of the nucleus to be ascertained.

(*c*) Blanc's rule was also used to determine the sizes of the various rings, but the failure of the rule in certain cases led to an erroneous formula; *e.g.*, ring C was originally believed to be five-membered. Thus Windaus and Wieland (1928) proposed 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 *cyclo*pentenophenanthrene 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.

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

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

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_2·CO·CH_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:

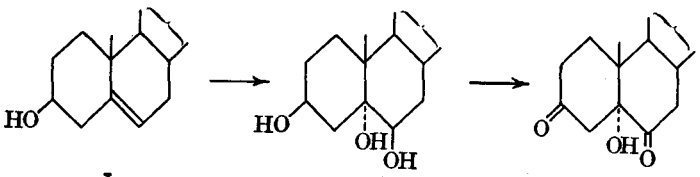
$$\text{Cholesterol } C_{27}H_{46}O \xrightarrow[CH_3\cdot CO_3H]{H_2O_2} \text{Cholestanetriol } C_{27}H_{48}O_3 \xrightarrow{CrO_3} \text{Hydroxycholestanedione } C_{27}H_{44}O_3$$

I VII VIII

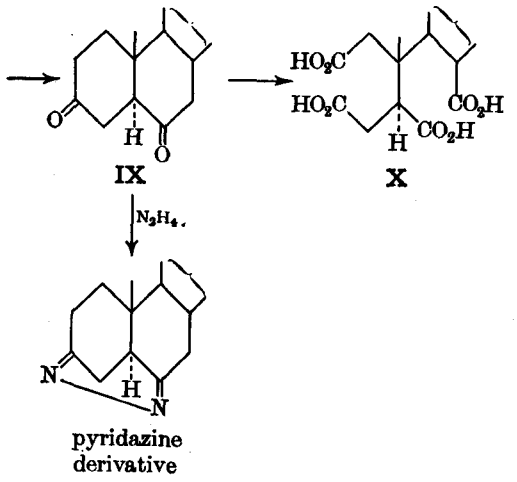
$$\xrightarrow[\text{(ii) } Zn—CH_3\cdot CO_2H]{\text{(i) } -H_2O} \text{Cholestanedione } C_{27}H_{44}O_2 \xrightarrow{CrO_3} \text{Tetracarboxylic acid } C_{27}H_{44}O_8$$

IX X

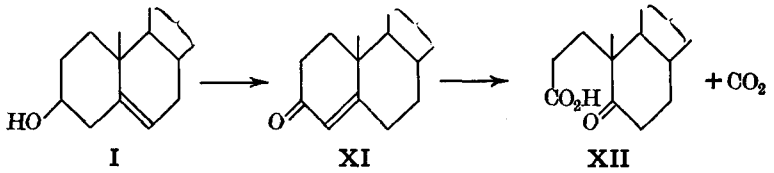
In the conversion of I into VII, the double bond in I is hydroxylated. Since only two of the three hydroxyl groups in VII are oxidised to produce VIII, these two groups are secondary alcoholic groups (one of these being the secondary alcoholic group in cholesterol), and the third, being resistant to 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 §§4, 4a, 4b). 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°, 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 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



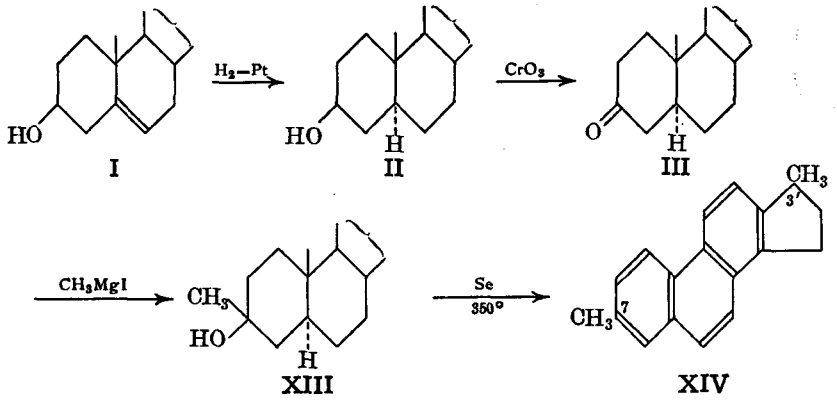
I VII VIII



is in position 3 and the double bond between 5 and 6, *position 5 being common to both rings A and B.* Thus:

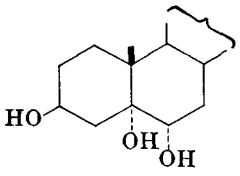


The position of the hydroxyl group at position 3 is definitely proved by the experiments of Kon *et al.* (1937, 1939). These authors reduced cholesterol, I, to cholestanol, II, oxidised this to cholestanone, III, treated this with methylmagnesium iodide and dehydrogenated the product, a tertiary alcohol, XIII, to 3′ : 7-dimethyl*cyclo*pentenophenanthrene, 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.

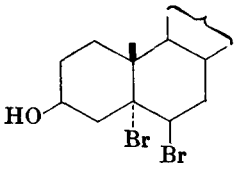


It might be noted here that the orientation of the two hydroxyl groups (introduced across the double bond in cholesterol) depends on the nature of the reagent used. With hydrogen peroxide, or *via* the oxide, the cholestanetriol is *trans*-5 : 6 (VII); with potassium permanganate or osmium

tetroxide, the product is *cis*-5 : 6 (VII*a*; *cf.* §5a. IV). These orientations may be explained as follows. When the addition of the two hydroxyl groups occurs *via* the oxide (the 5 : 6-oxide), the oxide ring will be formed *behind* the plane of the molecule due to the steric effect of the methyl group. Since opening of the epoxide ring occurs by attack on the conjugate acid (§5a. IV), the water molecule will attack from the back of the ring (*i.e.*, from the *front* of the molecule), and also preferably at position 6 due to the steric effect of the methyl group. Thus the orientation of the two hydroxyl groups (*trans*) will be as shown in VII. With permanganate (and osmium tetroxide),



VII *a*



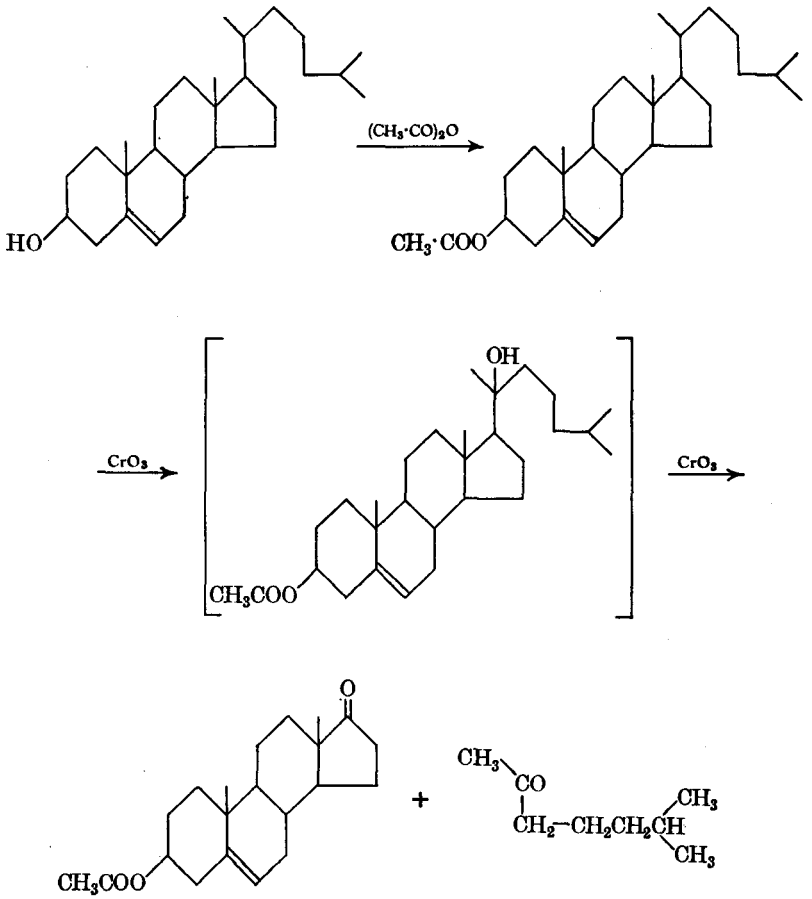
VII *b*

the plane of the cyclic compound will lie at the back of the molecule, again due to the steric effect of the methyl group. Moreover, since in the formation of the dihydroxy compound, *both* glycol oxygen atoms come from the permanganate ion (§5a. IV), it follows that *both* hydroxyl groups will be at the back of the molecule (VII*a*).

The addition of bromine, occurring *via* a brominium ion (§5a. IV), will produce the dibromide VII*b*, the reasons for the orientation being the same as those for the formation of VII (*via* the epoxide).

Since secondary alcoholic groups in steroids are readily oxidised to keto groups, and the latter may be located by mass spectra measurements (see §4b), this offers a very good way of locating secondary hydroxyl groups in the steroid molecule.

(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 *iso*hexyl methyl ketone, $CH_3{\cdot}CO{\cdot}(CH_2)_3{\cdot}CH(CH_3)_2$. 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:

$$R{\cdot}CH_2{\cdot}CO_2H \xrightarrow[HCl]{CH_3OH} R{\cdot}CH_2{\cdot}CO_2CH_3 \xrightarrow{2C_6H_5MgBr}$$

$$R{\cdot}CH_2{\cdot}C(OH)(C_6H_5)_2 \xrightarrow{-H_2O} R{\cdot}CH{=}C(C_6H_5)_2 \xrightarrow{CrO_3}$$

$$R{\cdot}CO_2H + (C_6H_5)_2CO$$

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

$$R{\cdot}CH_2{\cdot}C(OH)(CH_3)_2 \xrightarrow{CrO_3} R{\cdot}CO_2H + (CH_3)_2CO$$

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 coprostane (a stereoisomer of cholestane; see §§4, 4a). If we represent the nucleus of coprostane as Ar, and

the side-chain as C_n, then we may formulate the degradation of coprostane as follows (B–W represents a Barbier–Wieland degradation):

$$\underset{Ar—C_n}{\text{Coprostane}} \xrightarrow{CrO_3} CH_3{\cdot}CO{\cdot}CH_3 + \underset{Ar—C_{n-3}}{\text{Cholanic acid}} \xrightarrow{B\text{–}W}$$

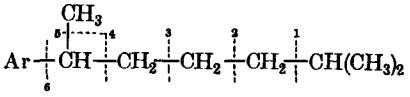
$$(C_6H_5)_2CO + \underset{Ar—C_{n-4}}{\text{Norcholanic acid}} \xrightarrow{B\text{–}W}$$

$$(C_6H_5)_2CO + \underset{Ar—C_{n-5}}{\text{Bisnorcholanic acid}} \xrightarrow{B\text{–}W}$$

$$(C_6H_5)_2CO + \underset{Ar—C_{n-6}}{\text{Ætiocholyl methyl ketone}} \xrightarrow{CrO_3} \underset{Ar—C_{n-7}}{\text{Etianic acid}}$$

The formation of acetone from coprostane indicates that the side-chain terminates in an *iso*propyl group. The conversion of bisnorcholanic 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 etianic acid (formerly known as ætiocholanic 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 bisnorcholanic 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, ætiocholanone, is obtained and this, on oxidation with nitric acid, gives a dicarboxylic acid, ætiobilianic acid, *without loss of any carbon atoms*. Thus ætiocholanone 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 more recent methods for degrading the side-chain:

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

$$Ar{\cdot}CHMe{\cdot}CH_2{\cdot}CH_2{\cdot}CO_2H \xrightarrow[(ii)\ CH_2N_2]{(i)\ SOCl_2} Ar{\cdot}CHMe{\cdot}CH_2{\cdot}CH_2{\cdot}CO{\cdot}CHN_2 \xrightarrow{HCl}$$

$$Ar{\cdot}CHMe{\cdot}CH_2{\cdot}CH_2{\cdot}CO{\cdot}CH_2Cl \xrightarrow[AcOH]{Zn} Ar{\cdot}CHMe{\cdot}CH_2{\cdot}CH_2{\cdot}CO{\cdot}CH_3 \xrightarrow[(ii)\ -HBr]{(i)\ Br_2}$$

$$Ar{\cdot}CHMe{\cdot}CH{=}CH{\cdot}CO{\cdot}CH_3 \xrightarrow{CrO_3} Ar{\cdot}CHMe{\cdot}CO_2H$$

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

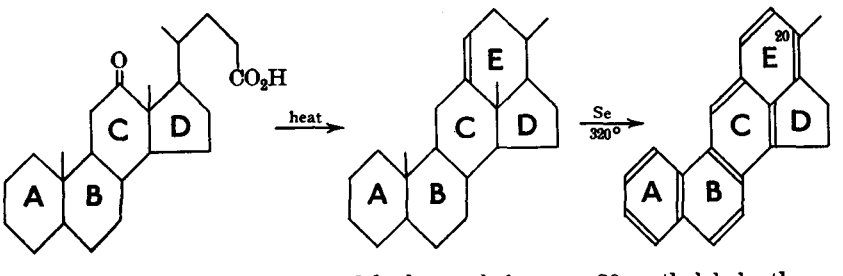
$$Ar{\cdot}CHMe{\cdot}CH_2{\cdot}CH_2{\cdot}CO_2Me \xrightarrow{2PhMgBr} Ar{\cdot}CHMe{\cdot}CH_2{\cdot}CH_2{\cdot}C(OH)Ph_2 \xrightarrow{-H_2O}$$

$$Ar{\cdot}CHMe{\cdot}CH_2{\cdot}CH{=}CPh_2 \xrightarrow[succinimide]{N\text{-bromo-}} Ar{\cdot}CHMe{\cdot}CHBr{\cdot}CH{=}CPh_2 \xrightarrow{-HBr}$$

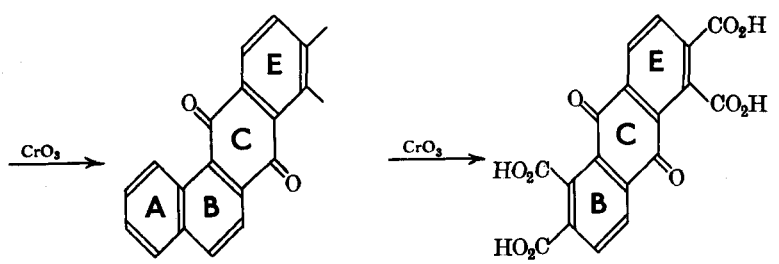
$$Ar{\cdot}CMe{=}CH{\cdot}CH{=}CPh_2 \xrightarrow{CrO_3} Ar{\cdot}COMe$$

(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, ætiobilianic acid, forms an anhydride when heated with acetic anhydride. Thus the ketone (ætiocholanone) 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 §2 vii. X). 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, cholanic acid may be obtained by the oxidation of coprostane. Cholanic acid may also be obtained by the oxidation of deoxycholic acid (a bile acid; see §8) 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-ketocholanic acid which, on heating to 320°, 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 (Fieser *et al.*, 1935; see §3f. X). The foregoing facts can be explained only if the side-chain in cholesterol is in position 17; thus:



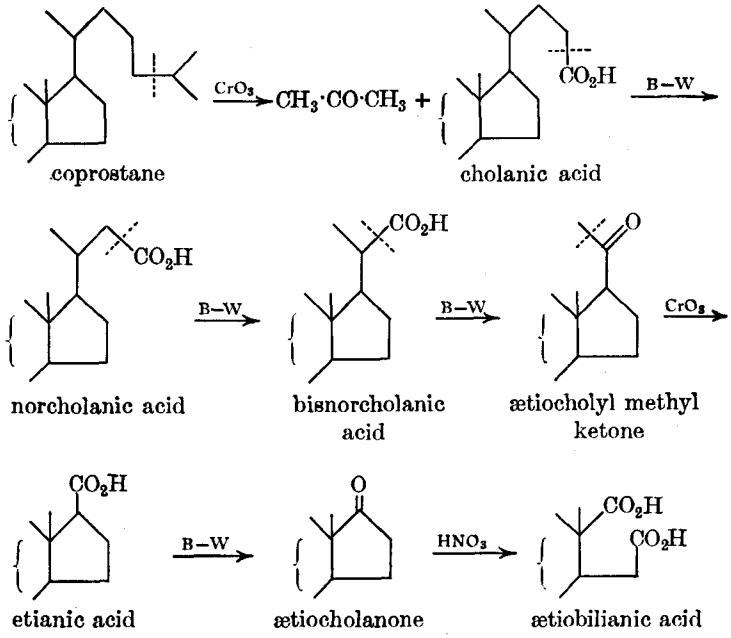
12-ketocholanic acid dehydronorcholene 20-methylcholanthrene



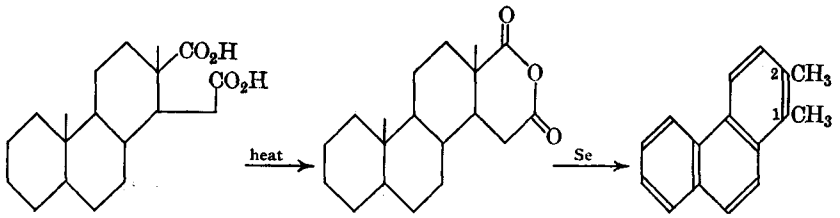
5:6-dimethyl-1:2-benzanthraquinone anthraquinone-1:2:5:6-tetracarboxylic acid

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

Thus, now that we know the nature and position of the side-chain, we can formulate the conversion of coprostane into ætiobilianic acid as follows:

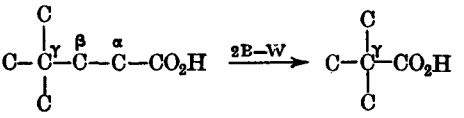


A point of interest in this connection is that when the anhydride of ætiobilianic acid is distilled with selenium, 1 : 2-dimethylphenanthrene is obtained (Butenandt *et al.*, 1933). This also provides proof for the presence of the phenanthrene nucleus in cholesterol, and also evidence for the position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

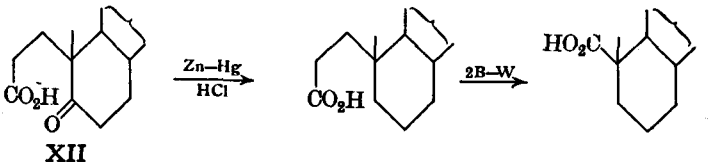


(iv) **Positions of the two angular methyl groups.** The *cyclo*pentenophenanthrene nucleus of cholesterol accounts for seventeen carbon atoms, and the side-chain for eight. Thus twenty-five carbon atoms in all have been accounted for, but since the molecular formula of cholesterol is $C_{27}H_{46}O$, 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, §31. VIII), the side-chain in XII must be of the type

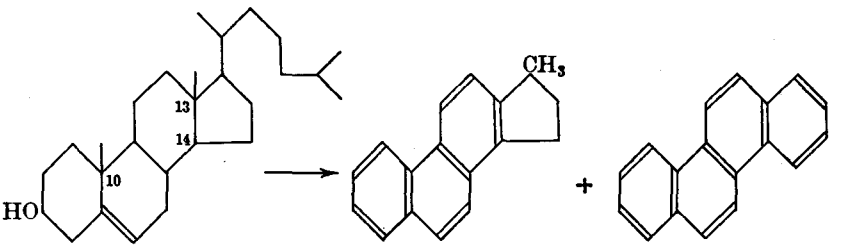


Thus there must be an alkyl group at position 10 in XII. This could be an ethyl group (as originally believed by Windaus and Wieland) or a methyl group, provided that in the latter case the second " missing " carbon atom can be accounted for. As we shall see later, there is also a methyl group at position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



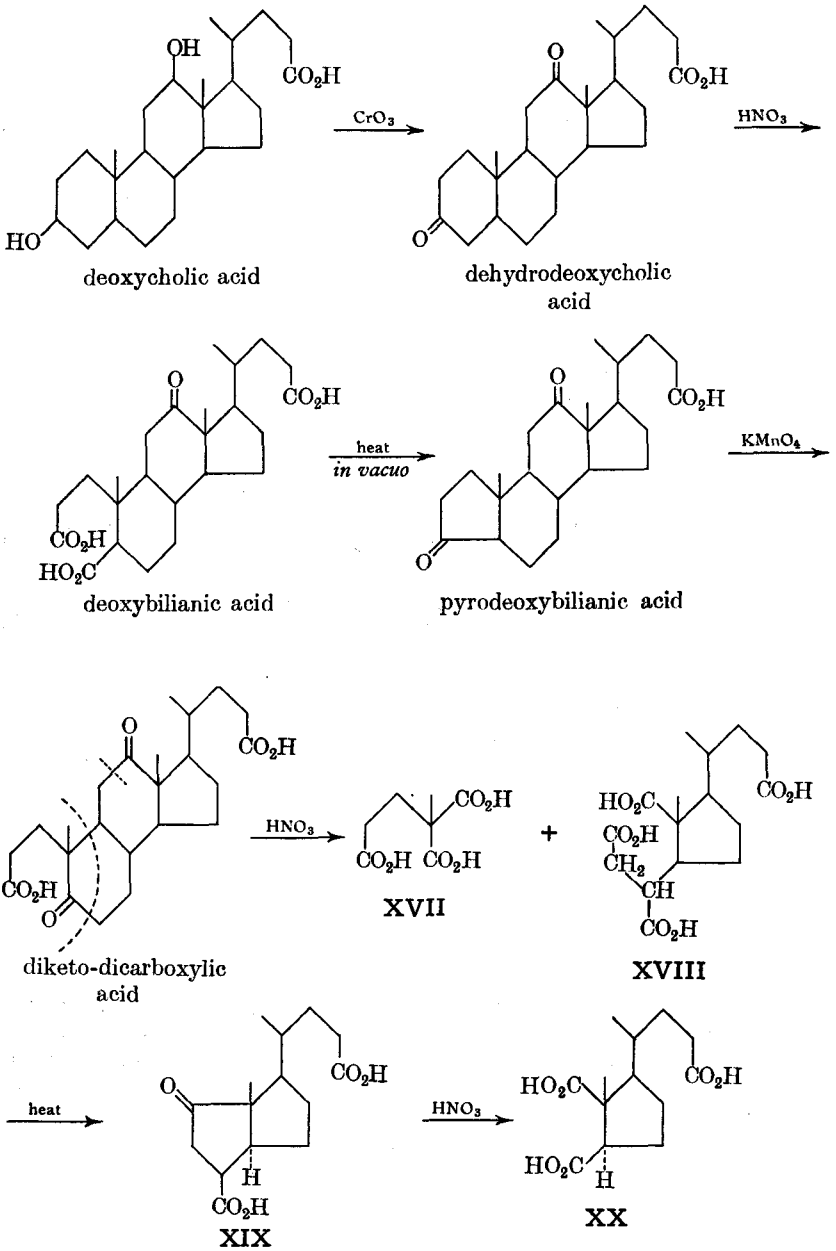
XII

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) (see overleaf).



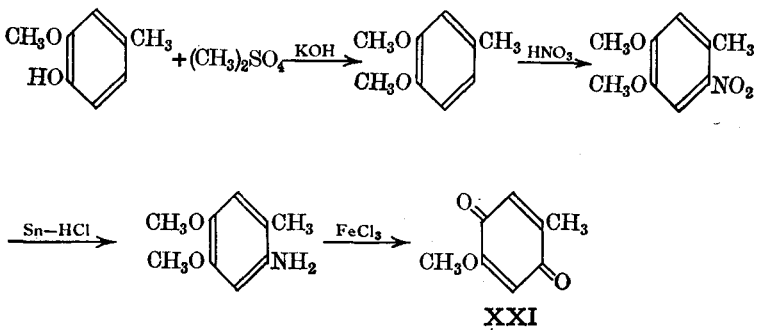
XVII was shown to be butane-2 : 2 : 4-tricarboxylic acid; thus there is a methyl group at position 10. XVIII was shown to be a tetracarboxylic acid containing a *cyclo*pentane ring with a side-chain

$$—CH(CH_3)\cdot CH_2\cdot CH_2\cdot CO_2H.$$

Thus this compound is derived from ring D. XX was also shown to be a tricarboxylic acid containing a *cyclo*pentane ring. Furthermore, one carb-

oxyl 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) to be at 13, and this is supported by the fact that ætiobilianic acid (XV, section iii) gives 1 : 2-dimethylphenanthrene (XVI) on dehydrogenation with selenium. Had the angular methyl group been at position 14, 1-methylphenanthrene would most likely have been obtained.

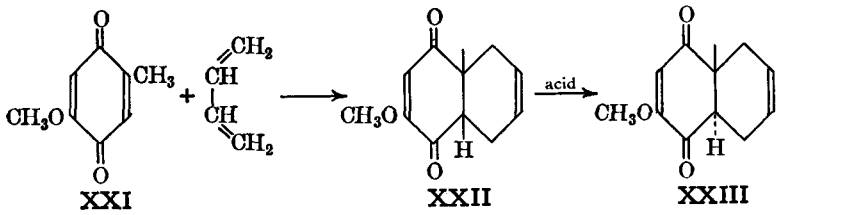
(v) **Synthesis of cholesterol.** Two groups of workers, *viz.*, Sir R. Robinson *et al.* (1951) and Woodward *et al.* (1951), have synthesised cholesterol. One of the outstanding difficulties in the synthesis of steroids is the stereochemical problem. The cholesterol nucleus contains eight asymmetric carbon atoms and so 256 optical isomers are possible (see also §4 for further details). Thus every step in the synthesis which produced a new asymmetric carbon atom 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 other parts. This problem led to the development of specific reagents. The following is an outline of the Woodward synthesis. 4-Methoxy-2 : 5-toluquinone, XXI, was prepared from 2-methoxy-*p*-cresol as follows:

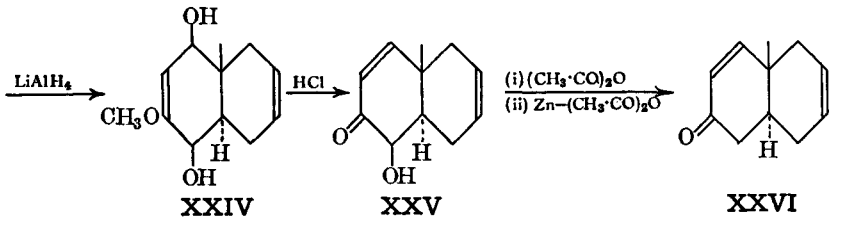


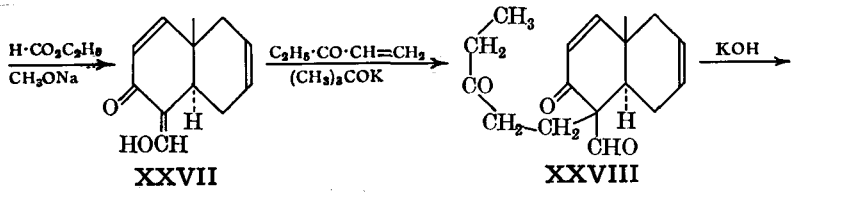
XXI

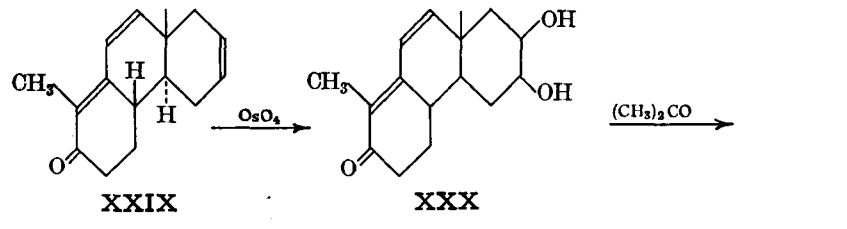
XXI was condensed with butadiene (Diels–Alder reaction) to give XXII. This had the *cis* configuration and was isomerised (quantitatively) to the *trans*-isomer XXIII by dissolving in aqueous alkali, adding a seed crystal of the *trans*-form, and then acidifying. XXIII, on reduction with lithium aluminium hydride, gave the glycol XXIV, and this, on treatment with aqueous acid, gave XXV. Conversion of XXV to XXVI by removal of the hydroxyl group was carried out by a new technique: XXV was acetylated and the product, the ketol acetate, was heated with zinc in acetic anhydride to give XXVI (reduction with metal and acid usually reduces $\alpha : \beta$-unsaturated bonds in ketones). XXVI, on treatment with ethyl formate in the presence of sodium methoxide, gave the hydroxymethylene ketone XXVII (Claisen condensation). When this was treated with ethyl vinyl ketone in the presence of potassium *tert.*-butoxide, XXVIII was formed (Michael condensation). The object of the double bond in the ketone ring in XXVI 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

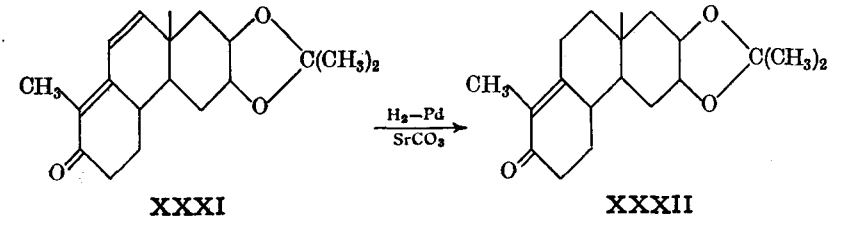
group (this is now flanked on *both* sides by carbonyl groups). The necessity for this " activation " lies in the fact that ethyl vinyl ketone tends to self-condense, and consequently decrease the yield of XXVIII. XXVIII was now cyclised (quantitatively) by means of potassium hydroxide in aqueous dioxan to the single product XXIX. This is the desired compound; the other possible isomer (XXIX with the two hydrogens *cis* instead of *trans* as shown) is not formed since the *cis*-isomer is less stable than the *trans*-. XXIX was then treated with osmium tetroxide to give two *cis*-glycols of structure XXX. These 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 *iso*propylidene derivative XXXI. This, on catalytic reduction (H_2—$Pd/SrCO_3$) gave XXXII which was condensed with ethyl formate in the presence of sodium methoxide to give XXXIII, and this was then converted into XXXIV 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). When XXXIV 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 XXXV (methyl group in front and propionic acid group behind the plane of the rings) was converted into the enol lactone XXXVI which, on treatment with methylmagnesium bromide, gave XXXVII, and this, on ring closure by means of alkali, gave XXXVIII. When this was oxidised with periodic acid in aqueous dioxan, the dialdehyde XXXIX was obtained, and this, when heated in benzene solution in the presence of a small amount of piperidine acetate, gave XL (and a small amount of an isomer). This ketoaldehyde was oxidised to the corresponding acid which was then converted into the methyl ester XLI with diazomethane. XLI, 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 (+)-XLI. This was catalytically reduced (H_2—Pt) to XLII, which was then oxidised to XLIII which was a mixture of stereoisomers (from the mixture of XLII; H at 17 behind and in front). These were separated, reduced (sodium borohydride), and hydrolysed. The β-isomer, XLIV, was converted into the methyl ketone by first acetylating, then treating with thionyl chloride and finally with dimethylcadmium. This acetylated hydroxyketone, XLV, on treatment with *iso*hexylmagnesium bromide, gave XLVI. This was a mixture of isomers (a new asymmetric carbon has been introduced at position 20). XLVI, on dehydration, gave one product, XLVII, and this, on catalytic hydrogenation (H_2—Pt), gave a mixture of cholestanyl acetates (the asymmetric C_{20} has been re-introduced). These acetates were separated and the desired isomer, on hydrolysis, gave cholestanol, XLVIII, which was identical with natural cholestanol. The conversion of cholestanol into cholesterol, LIII, is then carried out by a series of reactions introduced by various workers: XLVIII to XLIX (Bruce, 1943); XLIX to L (Butenandt *et al.*, 1935); L to LI (Ruzicka, 1938); LI to LII (Westphal, 1937); LII to LIII (Dauben *et al.*, 1950).

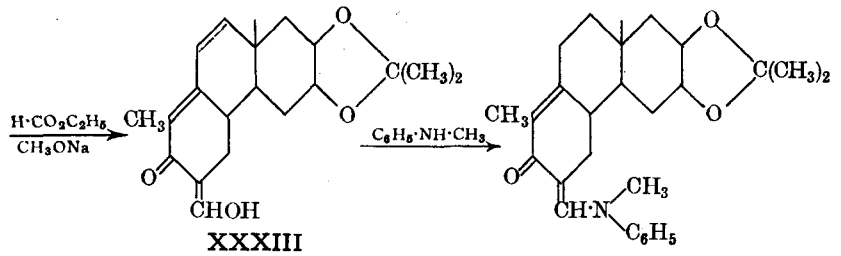

CH3O
CH3
CH2
CH
CH
CH2
acid
XXI
XXII
XXIII


LiAlH4
HCl
(i) (CH3·CO)2O
(ii) Zn–(CH3·CO)2O
XXIV
XXV
XXVI

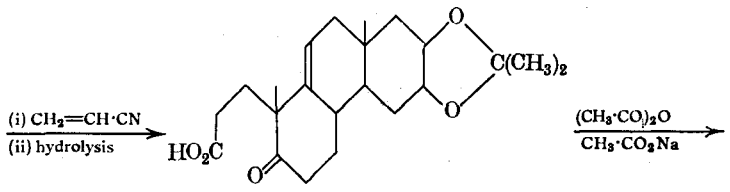

H·CO2C2H5
CH3ONa
C2H5·CO·CH=CH2
(CH3)3COK
KOH
HOCH
XXVII
XXVIII


OsO4
(CH3)2CO
XXIX
XXX

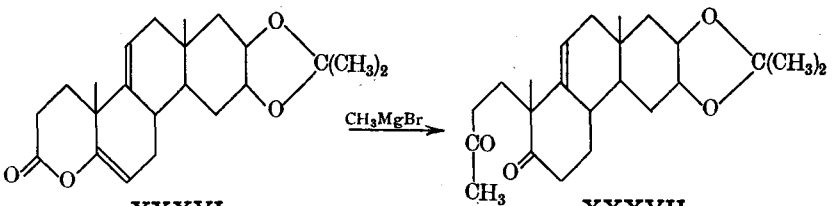

C(CH3)2
H2–Pd
SrCO3
XXXI
XXXII



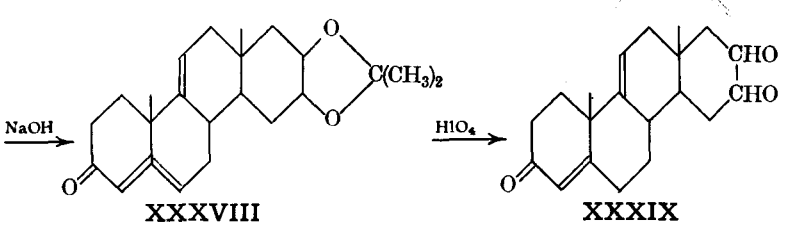
XXXIII XXXIV



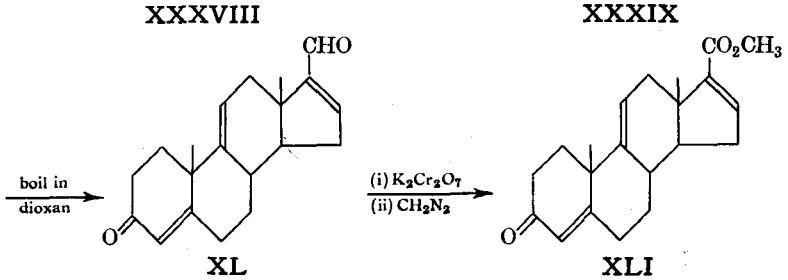
XXXV



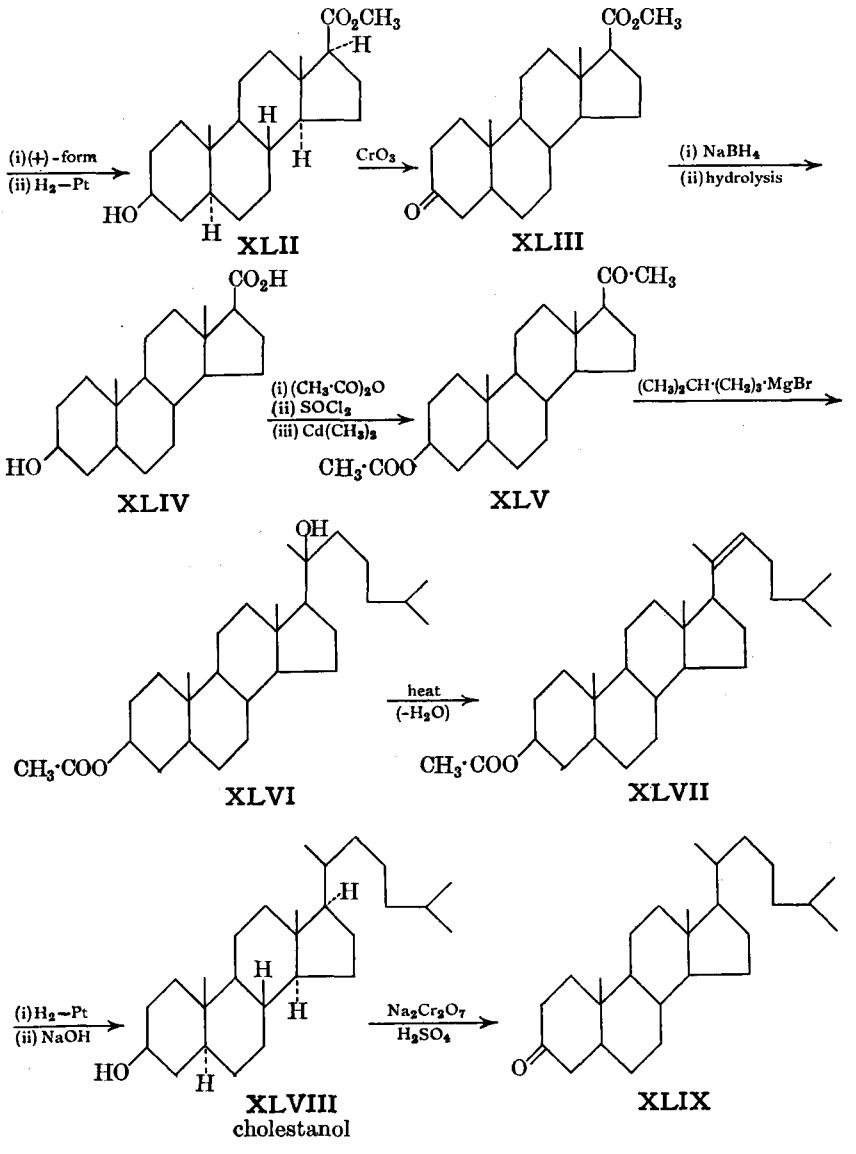
XXXVI XXXVII



XXXVIII XXXIX



XL XLI



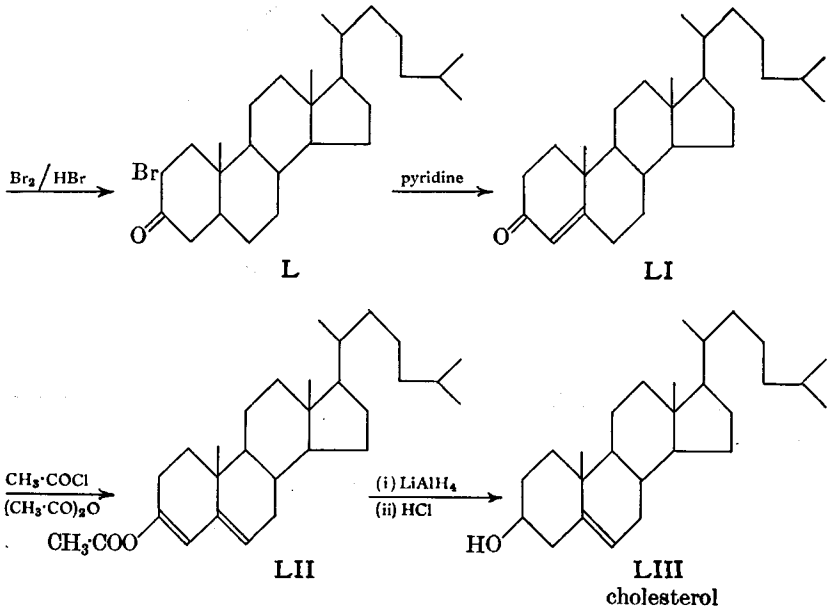
XLII XLIII

XLIV XLV

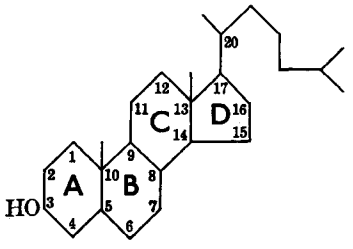
XLVI XLVII

XLVIII
cholestanol

XLIX



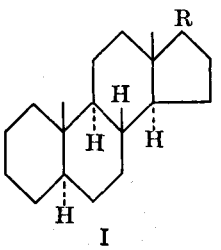
§4. Stereochemistry of the steroids. If we examine the fully saturated sterol, we find that there are eight dissimilar asymmetric carbon atoms in the nucleus (3, 5, 8, 9, 10, 13, 14 and 17). Thus there are $2^8 = 256$ optical isomers possible. If we also include the asymmetric carbon atom in the side-chain (20), then there are 512 optical isomers possible.



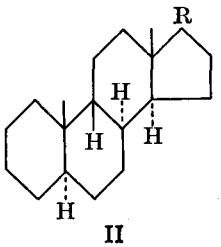
The stereoisomerism of the steroids is conveniently classified into two types, one dealing with the way in which the rings are fused together, and the other with the configurations of substituent groups, particularly those at C_3 and C_{17}.

§4a. Configuration of the nucleus. There are six asymmetric carbon atoms in the nucleus (5, 8, 9, 10, 13 and 14), and therefore there are $2^6 = 64$ optically active forms possible. X-ray analysis has shown that the steroid molecule is long and *thin, i.e.*, the molecule is essentially *flat* (Bernal, 1932). This is only possible if rings B and C are fused together in a *trans* manner (*cf. trans*-decalin, §11 vii. IV); rings A/B and C/D could be *cis* or *trans*. It has been found that all naturally occurring saturated steroids, except those of the heart poisons, belong either to the *cholestane series* or to the *coprostane series*; in the former the rings A/B are *trans*, and in the latter *cis*, the rings B/C and C/D being *trans* in *both* series. By convention a full line represents groups above the plane of the molecule, and a dotted (or broken) line represents groups below the plane (see also §11 vii. IV for

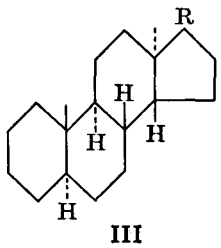
conventions). Furthermore, *by convention*, the methyl group at C_{10} in *cholestane* has been placed *above* the plane of the molecule, and therefore this leads to *four* possible stereoisomers for cholestane (I–IV). X-ray



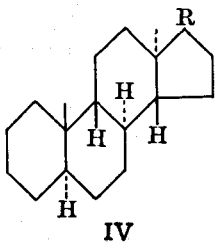
I



II

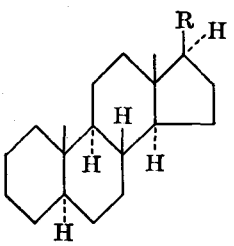


III



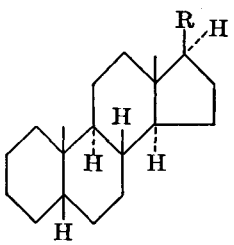
IV

analysis has shown that the hydrogen atom at C_9 is *trans* to the methyl group at C_{10} (Bernal *et al.*, 1940), and this conclusion is supported by chemical evidence. Thus cholestane must be I or III. Further chemical work has shown that the methyl groups at C_{10} and C_{13} are *cis*, and so cholestane is I, and consequently coprostane is also I, except that in this case the hydrogen atom at C_5 is *above* the plane (rings A/B are *cis* in coprostane). The final point to be settled in connection with this problem of the configuration of cholestane is the orientation of the side-chain R at C_{17}. Chemical evidence and X-ray analysis studies have shown that this side-chain is *above* the plane of the molecule (*i.e.*, *cis* with respect to the two angular methyl groups). Thus cholestane and coprostane are:



Cholestane

A/B *trans*
B/C *trans*
C/D *trans*
allo series



Coprostane

A/B *cis*
B/C *trans*
C/D *trans*
normal series

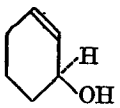
Compounds derived from cholestane are known as the ***allo*-compounds,** the prefix *allo* being reserved to indicate this configuration at C_5. Compounds derived from coprostane are known as the ***normal*-compounds,** but it should be noted that it is not customary to prefix compounds of this series by the word *normal*, *e.g.*, *allo*cholanic acid can be derived from cholestane, whereas cholanic acid can be derived from coprostane.

§4b. **Configurations of substituent groups.** The configuration of the side-chain at C_{17} has already been mentioned above. The only other configuration that we shall discuss here is that of the hydroxyl group at C_3. *By convention,* the hydroxyl at C_3 in **cholestanol** (and cholesterol) is taken as being **above** the plane of the ring, *i.e.,* the hydroxyl group is taken as being in the *cis* position with respect to the methyl group at C_{10}. This configuration occurs in all *natural* sterols, and gives rise to the **β-series,** the prefix β always indicating that the substituent group lies *above* the plane of the molecule. When the hydroxyl group lies **below** the plane, the compounds are said to belong to the **α-** or **epi series;** the prefix epi indicates the *epimer* due to the inversion of the configuration of C_3.

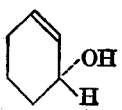
X-ray analysis studies have shown that the hydroxyl group in cholesterol is above the plane of the molecule, *i.e.,* it is *cis* to the methyl group at C_{10}. This has been confirmed by chemical evidence (Shoppee, 1947).

The assignment of the configurations of C_7 and C_{17} in steroid alcohols has been determined by Prelog *et al.* (1953) by arguments based on asymmetric syntheses (see §7. III). It has been shown that the configuration of the hydroxyl group in, *e.g.,* cholestan-7α-ol and androsten-17β-ol is in agreement with the accepted conventional steroid formula.

Mills (1952) has also correlated the configurations of steroids with glyceraldehyde. This author collected the molecular optical rotations of a number of pairs of epimeric *cyclo*hex-2-enols and their esters, and on the assumption that the configurations given (in the literature) were correct, Mills showed that the alcohol represented as I is more lævorotatory than its epimer II, irrespective of the positions of alkyl groups in these allylic terpene alcohols (these compounds had already been correlated with glyceraldehyde by the work of Fredga; §23e. VIII). The differences in rotation are large, and are increased on esterification. Mills then applied this rule to seven known



I



II

pairs of epimeric, allylic steroid alcohols, and found that the differences were those which may be predicted on the basis that the conventional steroid formulæ represent the absolute configurations. Thus the configuration of the 3β-hydroxyl group in cholesterol corresponds to that of D(+)-glyceraldehyde.

These stereochemical relationships of steroids to D(+)-glyceraldehyde have now been proved by the degradation of cholesterol to derivatives of (+)-citronellal (§23e. VIII), in which the only asymmetric carbon atom is the C_{20} of the steroid (Cornforth *et al.,* 1954; Riniker *et al.,* 1954). Thus the arbitrary choice of placing the angular methyl groups above the plane in the cholesterol nucleus (*i.e.,* the β-configuration) has proved to be the absolute configuration. Furthermore, since the configuration of the 3-hydroxyl group in cholesterol is β, this configuration is also the absolute one.

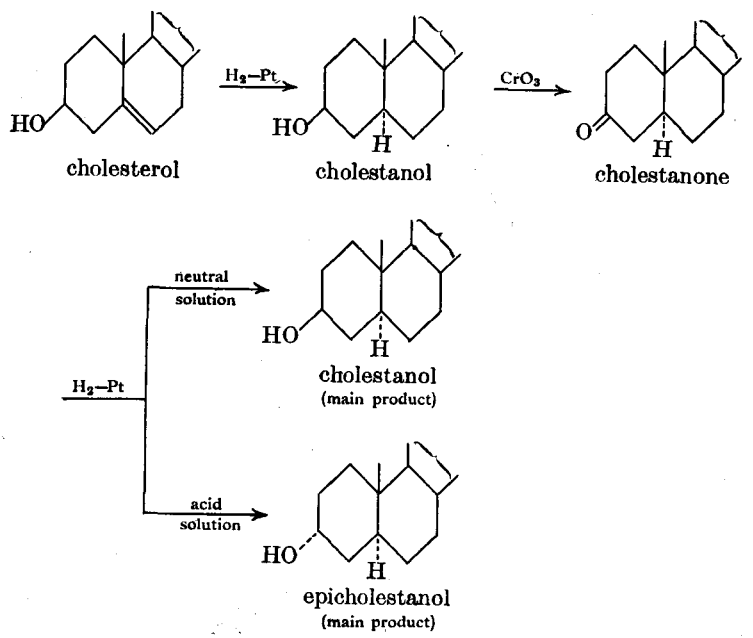
Barton (1944–) has also applied the method of optical rotations to steroid chemistry, and has called his treatment the Method of Molecular Rotation Differences (this is a modification of the Rule of Shift, §12. I). The basis of this method is that the molecular rotation of any steroid is considered as the sum of the rotation of the fundamental structure (which is the parent hydrocarbon cholestane, androstane, or pregnane) and the rotations contributed by the functional groups (these are called the Δ values). The Δ

value of a given group is a characteristic of its position and orientation, and the Δ values of different groups are independent of one another provided that unsaturated groups are not present, *i.e.*, conjugation is absent, or that the groups are not too close together, *i.e.*, are separated by 3 or 4 saturated carbon atoms. In this way it has been possible to assign configurations and also the positions of double bonds.

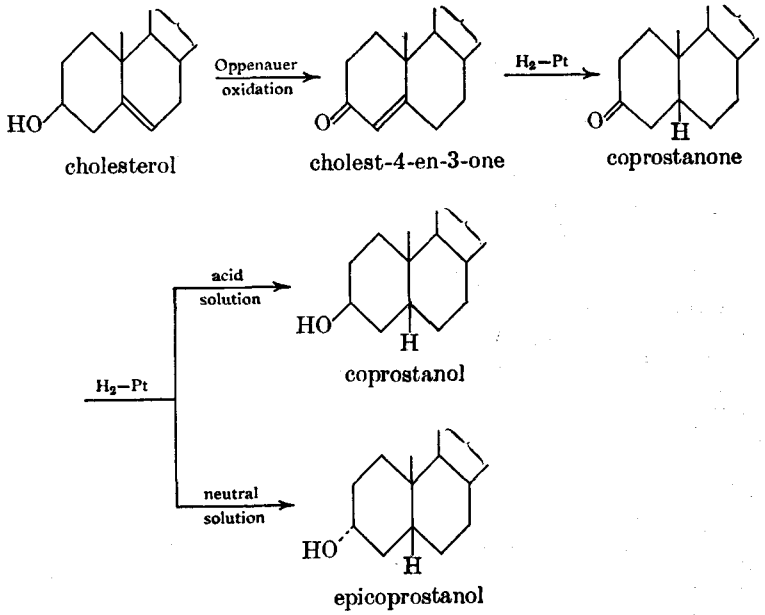
Correlation of configurations in steroids has also been carried out by the method of rotatory dispersion (§12a. I). Saturated steroids have been examined (Djerassi *et al.*, 1956) and the results show that as the position of the carbonyl group changes in A/B *trans*-steroids, the curves change in sign, shape and/or amplitude. Thus this method may be used to locate the unknown position of a carbonyl group in a steroid. The authors also showed that for a *given* position of the carbonyl group, the shape of the curve depends on the conformation of the molecule. Thus, by comparing the curve of the compound under investigation with that of a compound of *known* absolute configuration and containing the carbonyl group in the *same* position, it is then possible to deduce the absolute configuration of a group in the unknown compound.

On the other hand, Djerassi *et al.* (1962) have shown that mass spectra measurements of keto steroids offer a means of *locating* the carbonyl group in a steroid molecule. However, when mass spectrometry is combined with optical rotatory dispersion measurements, it is then possible to locate in an unambiguous manner the carbonyl group.

§4c. The preparation of the " stanols ". The catalytic hydrogenation (platinum) of cholesterol (cholest-5-en-3β-ol) produces only cholestanol (cholestan-3β-ol). On the other hand, oxidation of cholestanol with chromium trioxide in acetic acid gives cholestanone and this, on catalytic reduction in *neutral* solution, gives mainly cholestanol, whereas catalytic reduction in *acid* solution gives mainly epicholestanol (cholestan-3α-ol).



The corresponding C_5 epimers, coprostanol (coprostan-3β-ol) and epicoprostanol (coprostan-3α-ol), may be prepared from cholesterol as follows, the first step being the conversion of cholesterol into cholest-4-en-3-one by means of the Oppenauer oxidation (aluminium *tert.*- butoxide in acetone; see also Vol. I).



A detailed study of the catalytic reduction of the decalones has shown that in an acid medium the product is usually the *cis*-compound, whereas in a neutral or alkaline medium the product is usually the *trans*-compound (von Auwers, 1920; Skita, 1920). This principle, which is known as the *Auwers–Skita rule of catalytic hydrogenation*, was used by Ruzicka (1934) to determine the configurations of the above " stanols ". The configurations assigned have been supported by measurement of the rates of hydrolysis of the acetates of the various " stanols " (Ruzicka *et al.*, 1938). The acetates of cholestanol and epicoprostanol are hydrolysed much faster than those of epicholestanol and coprostanol (see §4d).

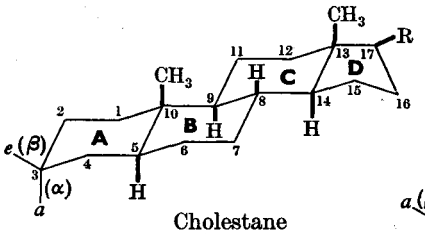
A point of interest in connection with the Auwers–Skita rule is that this generalisation does not allow for the possibility of isomerisation. Schuetz *et al.* (1962) have shown that in the hydrogenation of the three xylenes, the yield of the *trans*-isomer increased with temperature.

Now let us consider the configuration at C_5. The results of experiments on the catalytic hydrogenation of substituted *cyclo*hexanones and substituted phenols have led to the generalisation that the initial addition is *cis*, and occurs on the more accessible side of the double bond (Peppiatt *et al.*, 1955; Wicker, 1956). In accordance with this generalisation, it has been found that when saturated steroids of the A/B-*cis*- and the A/B-*trans*- series are produced by catalytic hydrogenation of 3α-substituted Δ^5-steroids, then the larger the size of the 3α-substituent, the larger is the proportion of the A/B-*cis*-steroid; in some cases, this *cis*-steroid is apparently formed exclusively (Shoppee *et al.*, 1955).

§4d. Conformational analysis of steroids. The Auwers–Skita rule of catalytic hydrogenation (§4c) cannot be used with certainty since, as pointed

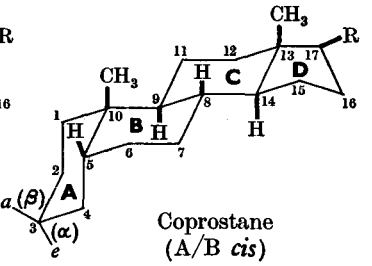
out, the product is *usually* mainly *cis* or *trans* according to the conditions, and hence the exceptions can only be ascertained as such by other evidence. Barton (1953) has restated this Auwers–Skita rule of catalytic hydrogenation as follows: Catalytic hydrogenation of ketones in strongly acid media (rapid hydrogenation) produces the axial hydroxyl compound, whereas hydrogenation in neutral media (slow hydrogenation) produces the equatorial alcohol if the ketone is unhindered or the axial alcohol if the ketone is very much hindered.

All the evidence obtained has shown that all the *cyclo*hexane rings in the steroid nucleus are chair forms; thus I is cholestane, and II is coprostane.



Cholestane
(A/B *trans*)

I

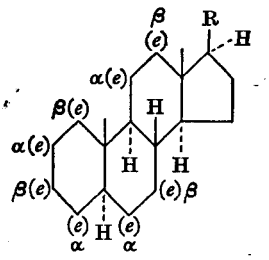


Coprostane
(A/B *cis*)

II

The effect of conformation on the course and rate of reactions has been discussed in §12. IV. The following is a summary of the generalisations that have been formulated:

(i) Equatorial groups are normally more stable than axial. Thus, when a (polycyclic) secondary alcohol is equilibrated with alkali, it is the equatorial isomer that predominates in the product. Similarly, when a (polycyclic) ketone is reduced with sodium and ethanol, the predominant isomer in the product is the equatorial alcohol (the more stable form). Furthermore, because of the rigidity of the system (which prevents interconversion of chair forms), the stable configurations of hydroxyl groups at different positions in the cholestane series will be as shown in III (compare this with I).



III

The following are examples of equilibration (using sodium ethoxide at 180°) (see also §8. II):

Cholestanol [3β(*e*)] ⇄ Epicholestanol [3α(*a*)] (10% ↓, 90% ↑)

Coprostanol [3β(*a*)] ⇄ Epicoprostanol [3α(*e*)] (10% ↑, 90% ↓)

(ii) Equatorial hydroxyl and carboxyl groups are esterified more rapidly

than the corresponding axial groups. Similarly, hydrolysis of equatorial esters and acyloxy groups is more rapid than for the corresponding axial isomers. These principles explain Ruzicka's results on the "stanols" (§4c); in the acetates of cholestanol and epicoprostanol, the acetoxy groups are equatorial, whereas in the acetates of epicholestanol and coprostanol these groups are axial and therefore subject to 1 : 3-interactions. Hence the former pair are hydrolysed more rapidly than the latter pair.

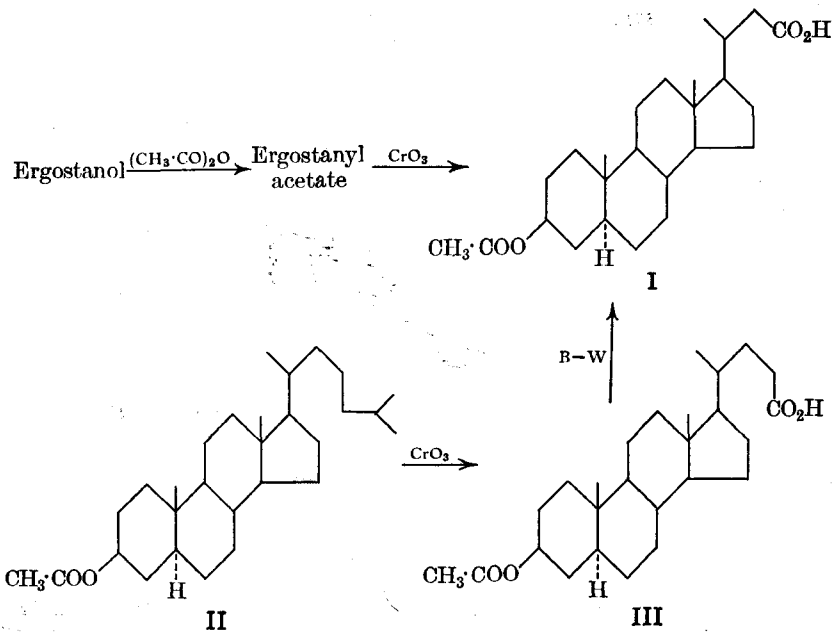
Empirical methods, using infra-red spectra, have been developed by Jones *et al.* (1951, 1952) for determining the conformation of 3-hydroxy (and 3-acetoxy) steroids; characteristic bands are given by the axial and equatorial groups.

(iii) Secondary axial alcohols are more rapidly oxidised by chromic acid (or hypobromous acid) than secondary equatorial alcohols. Schreiber *et al.* (1955) have shown that the more hindered the alcohol, the faster is the oxidation (with chromic acid).

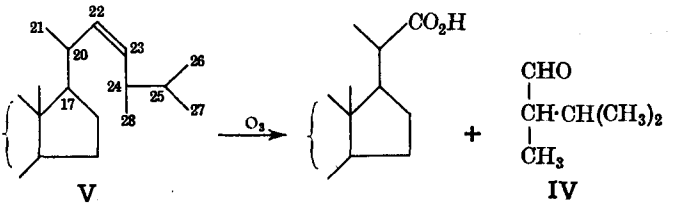
(iv) Bimolecular ionic elimination reactions occur readily when the two groups (which are eliminated) are *trans*-diaxial, and less readily when *trans*-diequatorial or *cis*-axial : equatorial.

(v) Epoxides are attacked by, *e.g.*, hydrogen bromide, to give the *trans*-diaxial compound. Reduction with lithium aluminium hydride or catalytic hydrogenation converts epoxides into the axial hydroxy compound.

§5. **Ergosterol,** $C_{28}H_{44}O$, m.p. 163°, occurs in yeast. Ergosterol forms esters, *e.g.*, an acetate with acetic anhydride; thus there is a hydroxyl group present in ergosterol. Catalytic hydrogenation (platinum) of ergosterol produces ergostanol, $C_{28}H_{50}O$; thus there are three double bonds in ergosterol. When ergostanol is acetylated and the product then oxidised, the acetate of 3β-hydroxynor*allo*cholanic acid, I, is obtained (Fernholz *et al.*, 1934). The identity of I is established by the fact that cholestanyl acetate, II (a compound of known structure), gives, on oxidation, the acetate of 3β-hydroxy*allo*cholanic acid, III, and this, after one Barbier–Wieland degradation (§3 iii), gives I; thus:

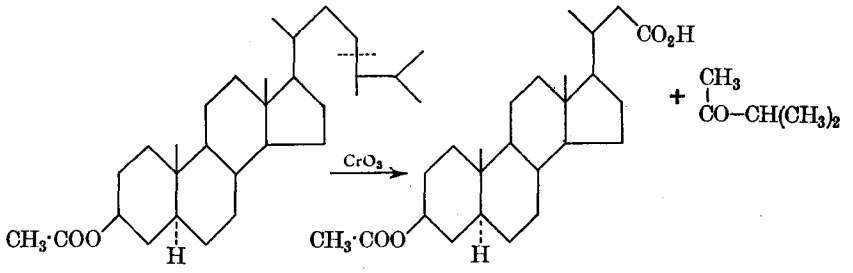


Thus ergostanol and cholestanol have identical nuclei, the same position of the hydroxyl group and the same position of the side-chain. The only difference must be the *nature* of the side-chain, and hence it follows that ergosterol contains one more carbon atom in its side-chain than cholesterol (the former compound is $C_{28}H_{44}O$ and the latter $C_{27}H_{46}O$). Ozonolysis of ergosterol gives, among other products, methyl*iso*propylacetaldehyde, IV. This can be accounted for if the side-chain of ergosterol is as shown in V (Windaus *et al.*, 1932).



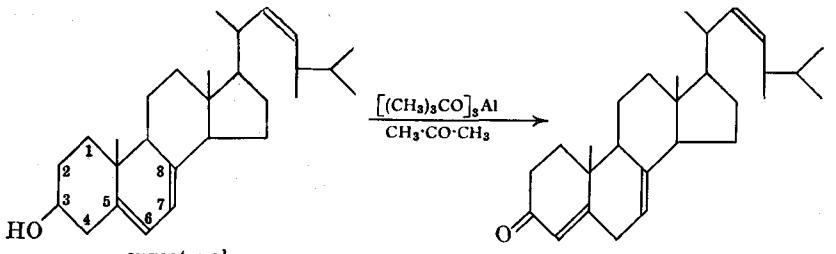
V IV

On this basis, the oxidation of ergostanyl acetate to the acetate of 3β-hydroxynor*allo*cholanic acid, I, is readily explained.



ergostanyl acetate I

We have now accounted for all the structural features of ergosterol except the positions of the three double bonds. The position of one of these is actually shown in the above account; it is C_{22}—C_{23}. The side-chain must contain only *one* double bond, since if more than one were present, more than one fragment (IV) would have been removed on ozonolysis. Thus the other two double bonds must be in the nucleus. When heated with maleic anhydride at 135°, ergosterol forms an adduct, and so it follows that the two double bonds (in the nucleus) are conjugated (Windaus *et al.*, 1931). Now ergosterol has an absorption maximum at 2810 A. Conjugated *acyclic* dienes absorb in the region of 2200–2500 A, but if the diene is in a *ring system*, then the absorption is shifted to the region 2600–2900 A. Thus the two double bonds in the nucleus of ergosterol are in *one* of the rings (Dimroth *et al.*, 1936). When ergosterol is subjected to the Oppenauer oxidation (aluminium *tert.*-butoxide and acetone), the product is an $\alpha:\beta$-unsaturated ketone (as shown from its absorption spectrum). This can only be explained by assuming that one of the double bonds is in the 5 : 6-position, and moves to the 4 : 5-position during the oxidation (*cf.* cholesterol, §3 ii). The other double bond is therefore 7 : 8 in order to be conjugated with the one that is 5 : 6. Thus the conjugated system is in ring B and the oxidation is explained as follows:

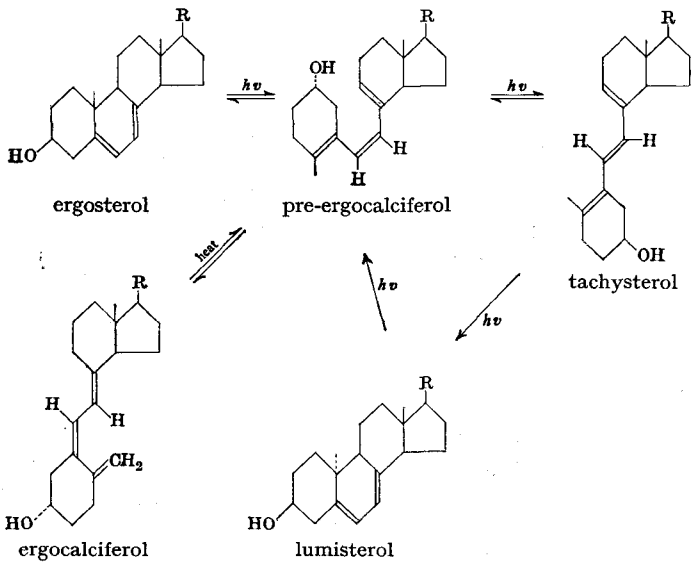


ergosterol

§6. Vitamin D. This vitamin is the antirachitic vitamin; it is essential for bone formation, its function being the control of calcium and phosphorus metabolism.

Steenbock *et al.* (1924) showed that when various foods were irradiated with ultraviolet light, they acquired antirachitic properties. This was then followed by the discovery that the active compound was in the unsaponifiable fraction (the sterol fraction). At first, it was believed that the precursor of the active compound was cholesterol, but subsequently the precursor was shown to be some " impurity " that was in the cholesterol fraction (*e.g.*, by Heilbron *et al.*, 1926). The ultraviolet absorption spectrum of this " impure cholesterol " indicated the presence of a small amount of some substance that was more unsaturated than cholesterol. This led to the suggestion that ergosterol was the provitamin D in the " impure cholesterol ", and the investigation of the effect of ultraviolet light on ergosterol resulted in the isolation from the irradiated product of a compound which had very strong antirachitic properties. This compound was named **calciferol** by the Medical Research Council (1931), and **vitamin D_1** by Windaus (1931). This potent crystalline compound, however, was subsequently shown to be a molecular compound of calciferol and lumisterol (one molecule of each). Windaus (1932) therefore renamed the pure potent compound as **vitamin D_2**, but the M.R.C. retained the original name calciferol. The *Chemical Society* (1951) has proposed the name **ergocalciferol** for this pure compound.

A detailed study of the irradiation of ergosterol with ultraviolet light has led to the proposal that the series of changes is as follows (R = C_9H_{17}):



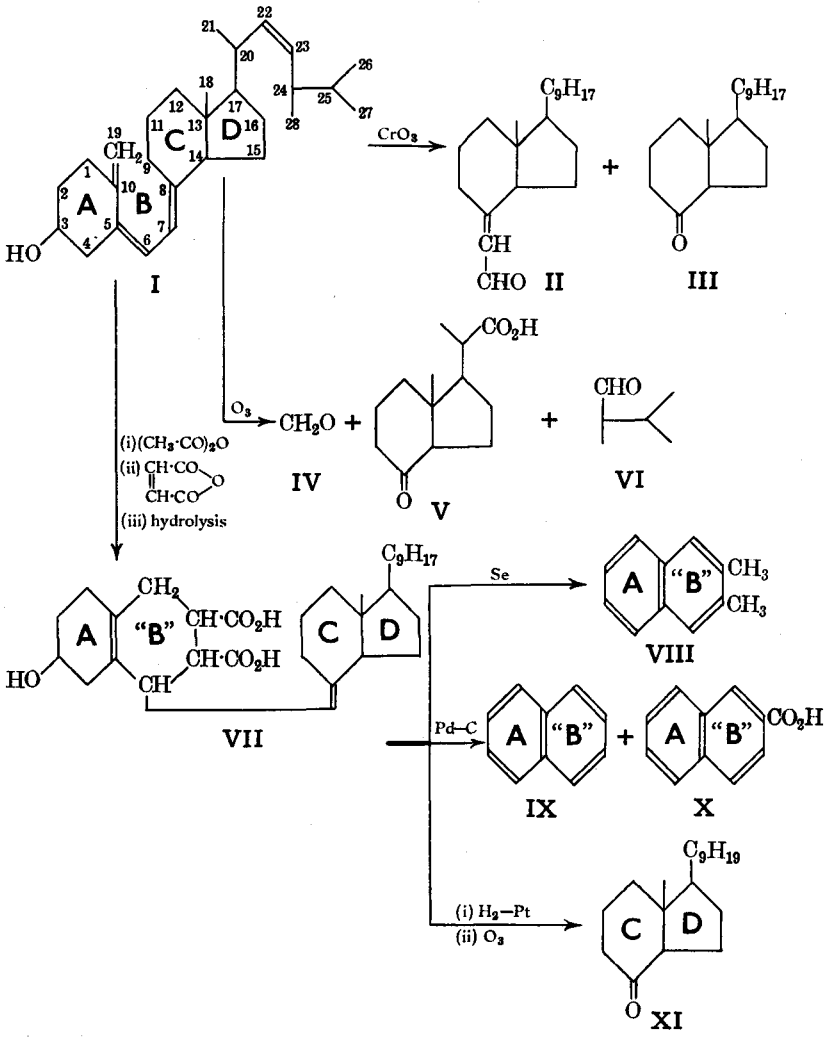
Velluz *et al.* (1949) isolated the pre-ergocalciferol (P) by irradiation of ergosterol at 20°, and showed that it formed ergocalciferol (E) on heating (see also below). Velluz *et al.* (1955) and Havinga *et al.* (1955) showed that pre-ergocalciferol is the 6 : 7-*cis*-isomer of tachysterol (T), and the interconversion of these two compounds has been studied by Inhoffen *et al.* (1959) and Havinga *et al.* (1959). Lumisterol (L) is converted directly into pre-ergocalciferol (Rappoldt, 1960). It should be noted that tachysterol and lumisterol are formed in a side reaction from pre-ergocalciferol and are *not* directly involved in the formation of ergocalciferol as postulated in the original scheme of Windaus *et al.*, who carried out the irradiation in solution and allowed the temperature to rise to 50°:

$$\text{Ergosterol} \xrightarrow{h\nu} \text{L} \xrightarrow{h\nu} \text{T} \xrightarrow{h\nu} \text{E}$$

§6a. Ergocalciferol (calciferol, vitamin D_2) is an optically active crystalline solid, m.p. 115–117°. Its molecular formula is $C_{28}H_{44}O$, and since it forms esters, the oxygen is present as a hydroxyl group. Furthermore, since ergocalciferol gives a ketone on oxidation, this hydroxyl group is a secondary alcoholic group. Ozonolysis of ergocalciferol produces, among other products, methyl*iso*propylacetaldehyde. Thus the side-chain in ergocalciferol is the same as that in ergosterol. Catalytic hydrogenation converts ergocalciferol into the fully saturated compound octahydroergocalciferol, $C_{28}H_{52}O$. This shows that there are four double bonds present, and since one is in the side-chain, three are therefore in the nucleus. The parent hydrocarbon of ergocalciferol is $C_{28}H_{52}$, and since this corresponds to the general formula C_nH_{2n-4}, the molecule therefore is *tricyclic*. Furthermore, ergocalciferol does not give Diels' hydrocarbon when distilled with selenium. These facts indicate that ergocalciferol does not contain the four-ring system of ergosterol. The problem is thus to ascertain which of the rings in ergosterol has been opened in the formation of ergocalciferol. The following reactions of ergocalciferol are readily explained on the assumption that its structure is I. The absorption spectrum of the semicarbazone of II ($C_{21}H_{34}O$) was shown to be characteristic of $\alpha : \beta$-unsaturated aldehydes. The absence of the hydroxyl group and the carbon content of II indicate the *absence* of ring A. These facts suggest that in ergocalciferol " ring B " is open between C_9 and C_{10}, and that II arises by scission of the molecule at a double bond in position 5 : 6, and can be an $\alpha : \beta$-unsaturated aldehyde only if there is a double bond at 7 : 8 (these double bonds are also present in ergosterol). The isolation of the ketone III ($C_{19}H_{32}O$) confirms the presence of the double bond at 7 : 8 (Heilbron *et al.*, 1935).

The isolation of formaldehyde (IV) shows the presence of an exocyclic methylene group, and the presence of this group at C_{10} is in keeping with the opening of ring B at 9 : 10. The formation of V ($C_{13}H_{20}O_3$), a keto-acid, suggests that ring B is open at 9 : 10, and that there are two double bonds at 7 : 8 and 22 : 23. The position of the latter double bond is confirmed by the isolation of methyl*iso*propylacetaldehyde, VI (Heilbron *et al.*, 1936).

Structure I for ergocalciferol is also supported by the formation of VII, the structure of which is shown by the products VIII, IX, X and XI (Windaus *et al.*, 1936). The production of 2 : 3-dimethylnaphthalene (VIII) is in keeping with the fact that carboxyl groups sometimes give rise to methyl groups on selenium dehydrogenation (*cf.* §2 vii. X). Similarly, the formation of naphthalene, IX, and naphthalene-2-carboxylic acid, X, shows the presence of rings A and " B " in VII. Catalytic reduction of VII (to reduce the double bond in the *side-chain* only), followed by ozonolysis, gives XI. Thus the formation of these compounds VIII–XI establishes the structure of VII, and shows that the double bonds are at 5 : 6, 10 : 19 and 7 : 8.



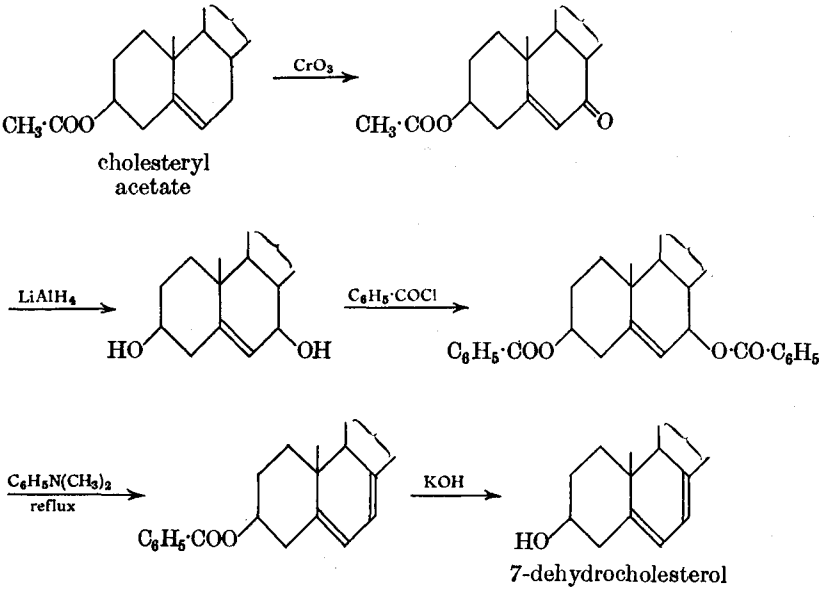
X-ray analysis studies of the 4-iodo-3-nitrobenzoate of ergocalciferol confirm structure I for ergocalciferol (Crowfoot *et al.*, 1948).

The presence of the two double bonds 5 : 6 and 7 : 8 gives rise to the possibility of various geometrical isomeric forms for ergocalciferol. Ultraviolet spectroscopic studies (Braude *et al.*, 1955) and other work (§6) have led to the conclusion that ergocalciferol has the configuration shown in the chart in §6. This is further supported by the work of Crowfoot *et al.* (1957) who, from calculations of electron densities in the ester crystal (the 4-iodo-3-nitrobenzoate), have shown that their results agree with the configuration given in the chart.

Lythgoe *et al.* (1957) have carried out a partial synthesis of ergocalciferol from the aldehyde II.

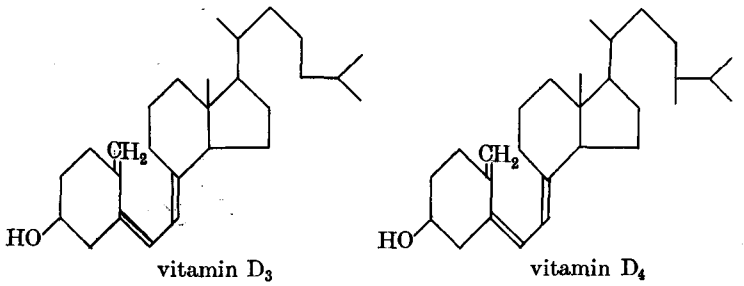
§6b. Vitamins D_3 and D_4. A detailed biological investigation has shown that the vitamin D in cod-liver oil is not identical with ergocalciferol, and that vitamin D activity could be conferred on cholesterol, or on some

impurity in cholesterol other than ergosterol. Windaus (1935) therefore suggested that natural vitamin D (in cod-liver oil) is derived from 7-dehydrocholesterol. The following chart shows the method of preparing 7-dehydrocholesterol (originated by Windaus, 1935, and improved by Buser, 1947, and by Fieser *et al.*, 1950).



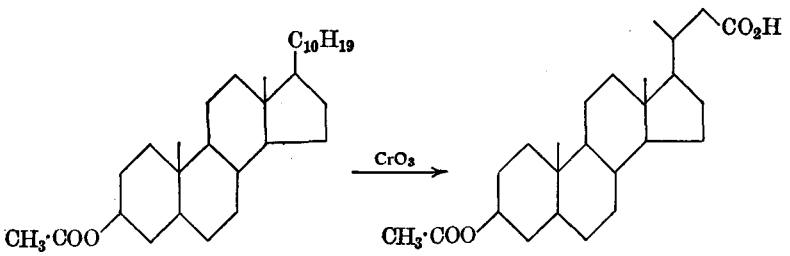
7-Dehydrocholesterol, on irradiation with ultraviolet light, gives a product that is about as active as ergocalciferol (vitamin D_2). This product was shown to be impure, and the pure active constituent was isolated as the 3 : 5-dinitrobenzoate (Windaus *et al.*, 1936). This vitamin D with a cholesterol side-chain is named **vitamin D_3**, and has been shown to be identical with the natural vitamin that is isolated from tunny-liver oil (Brockman, 1937). Vitamin D_3 has also been isolated from other fish-liver oils, *e.g.*, halibut. The *Chemical Society* (1951) has proposed the name **cholecalciferol** for vitamin D_3. It has now been shown that the irradiation of 7-dehydrocholesterol (at low temperature) first produces the previtamin D_3, and this, on gentle heating, is converted into the vitamin itself (*cf.* ergocalciferol, §6).

Irradiation of 22 : 23-dihydroergosterol gives a compound with antirachitic properties (Windaus *et al.*, 1937); this is known as **vitamin D_4**.



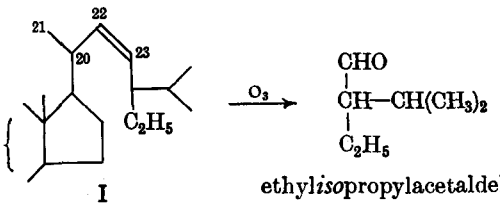
§7. **Stigmasterol,** $C_{29}H_{48}O$, m.p. 170°, is best obtained from soya bean oil. Since stigmasterol forms an acetate, etc., a hydroxyl group is therefore

present. Stigmasterol also forms a tetrabromide; thus it contains two double bonds. Hydrogenation of stigmasterol produces stigmastanol, $C_{29}H_{52}O$, and since the acetate of this gives the acetate of 3β-hydroxynor-*allo*cholanic acid on oxidation with chromium trioxide, it follows that stigmastanol differs from cholestanol only in the *nature* of the side-chain (Fernholz



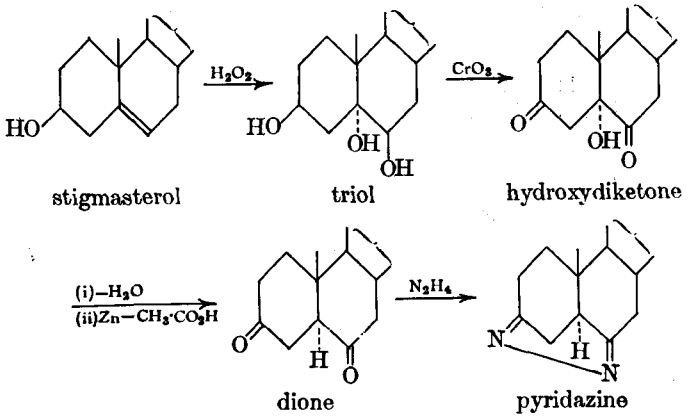
stigmastanyl acetate — acetate of 3β-hydroxynor-*allo*cholanic acid

et al., 1934; *cf.* ergosterol, §5). Ozonolysis of stigmasterol gives, among other products, ethyl*iso*propylacetaldehyde (Guiteras, 1933). This suggests that the side-chain is as shown in I, with a double bond at 22 : 23.



I — ethyl*iso*propylacetaldehyde

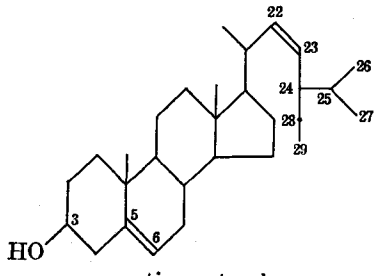
Thus the final problem is to ascertain the position of the second double bond in stigmasterol. This has been shown to be 5 : 6 by the method used for cholesterol (Fernholz, 1934). Stigmasterol, on hydroxylation with hydrogen peroxide in acetic acid, gives a triol which, on oxidation with chromium trioxide, forms a hydroxydiketone. This, on dehydration followed by reduction, forms a dione which combines with hydrazine to form a pyridazine derivative. These reactions can be explained as follows (*cf.* cholesterol, §3 ii):



stigmasterol — triol — hydroxydiketone

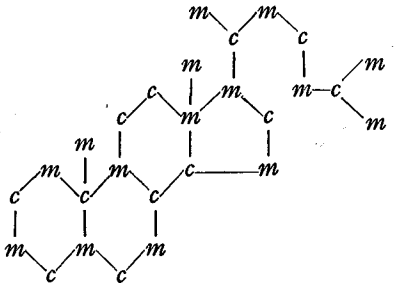
dione — pyridazine

This position for the nuclear double bond is supported by other evidence; thus stigmasterol is:



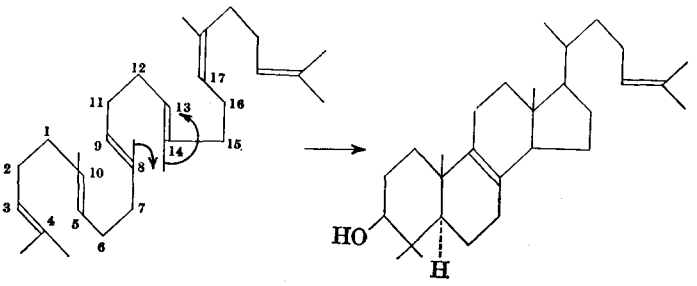
stigmasterol

§7a. Biosynthesis of sterols. It has long been known that animals can synthesise cholesterol, but the possible pathways were unknown until biosynthetic cholesterol was prepared from acetic acid labelled isotopically (with ^{14}C) in either the methyl (*m*) or the carboxyl (*c*) group, or labelled in both groups ($^{13}CH_3 \cdot ^{14}CO_2H$). These tracer studies were carried out mainly by Bloch *et al.* (1942–) and by Cornforth *et al.* (1953–), and the results established that the distribution of the carbon atoms is as shown in I. Thus

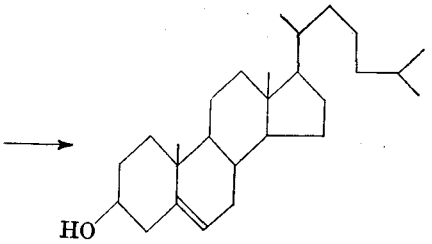


I

acetic acid can be regarded as the fundamental unit. Evidence was also obtained that *iso*valeric acid can serve as a precursor for cholesterol, and then Tavormina *et al.* (1956), using labelled mevalonic acid (MVA), showed that this is converted almost completely into cholesterol by rat liver; the route from acetic acid to MVA has been described in §32a. VIII. The problem now is to discover the route whereby MVA is converted into cholesterol. As far back as 1926 Heilbron *et al.* suggested that squalene (§32. VIII) is a precursor of cholesterol, and Robinson (1934) proposed a scheme for the cyclisation of the squalene molecule with the loss of three methyl groups. Woodward *et al.* (1953), however, suggested that squalene is first cyclised to lanosterol, and then this loses three methyl groups to give cholesterol. Bloch *et al.* (1952) showed that squalene is a precursor of cholesterol in the intact animal. Furthermore, Bloch *et al.* (1955) showed that lanosterol is converted into cholesterol in rats, and in 1956 carried out the biosynthesis of lanosterol from labelled acetate. Thus we have evidence for the suggested route from squalene to cholesterol. As mentioned above, Woodward *et al.* (1953) suggested that squalene ring-closes to form lanosterol, and proposed a 1,3-shift of the methyl group at C_8 to C_{13} (the squalene molecule is numbered to give the numbering in the closed-ring system in the steroid). On the other hand, Ruzicka *et al.* (1955) and Bloch *et al.* (1957) proposed a 1,2-shift of the methyl group from C_{14} to C_{13} and another 1,2-shift from C_8 to C_{14}. Further work by Bloch *et al.* (1958) showed that the 1,2-shifts were correct; this is also supported by the work of Cornforth *et al.* (1958).



squalene lanosterol



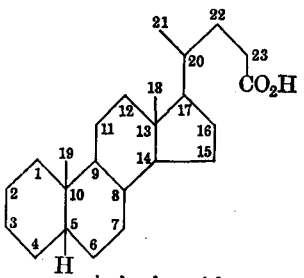
cholesterol

Bloch *et al.* (1957) also found that the three methyl groups of lanosterol are eliminated as carbon dioxide (*via* oxidation to carboxyl groups). Several intermediates and new precursors which function between lanosterol and cholesterol have now been identified (Cornforth, 1959; Crabbé, 1959). Finally, studies with yeast extracts have shown the mevalonic acid 5-pyrophosphate, *iso*pentenyl pyrophosphate, geranyl pyrophosphate and farnesyl pyrophosphate are successive intermediates in the biosynthesis of squalene (see §32a. VIII).

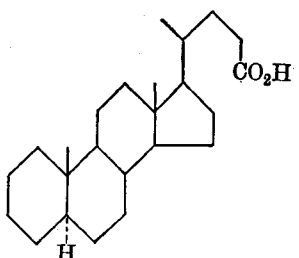
The biosynthesis of ergosterol from acetate has been carried out by Bloch *et al.* (1951), and the distribution pattern corresponds to that of cholesterol. Bloch *et al.* (1957) also showed that formate is an efficient source for the methyl group at C_{28}.

BILE ACIDS

§8. Introduction. The bile acids occur in bile (a secretion of the liver which is stored in the gall-bladder) of most animals combined as amides with either glycine ($NH_2{\cdot}CH_2{\cdot}CO_2H$) or taurine ($NH_2{\cdot}CH_2{\cdot}CH_2{\cdot}SO_3H$), *e.g.*, glycocholic acid (= glycine + cholic acid), taurocholic acid (= taurine + cholic



cholanic acid



*allo*cholanic acid

acid). The bile acids are present as sodium salts, and they function as emulsifying agents in the intestinal tract, *e.g.*, fats, which are insoluble in water, are rendered " soluble ", and so may be absorbed in the intestine.

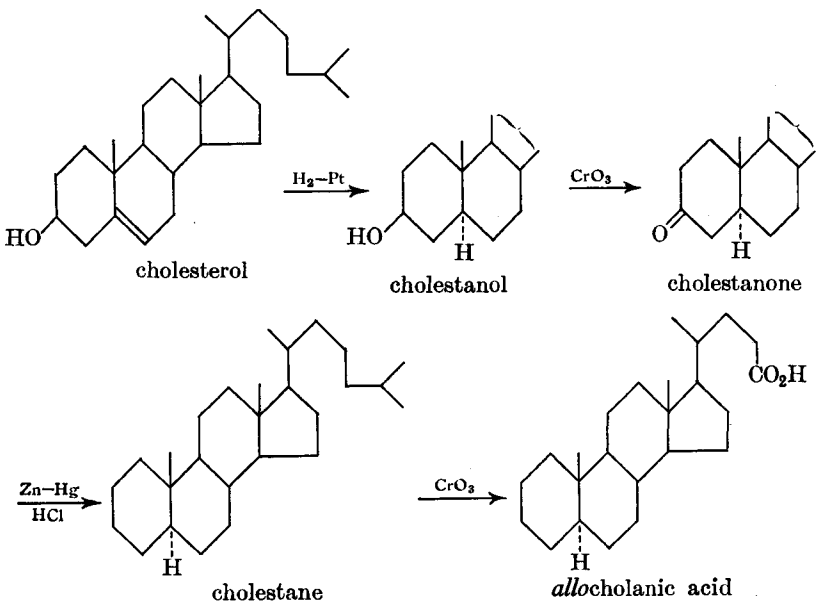
The bile acids are hydroxy derivatives of either cholanic acid or *allo*-cholanic acid (but see §10). Dehydration of a bile acid by heating in a vacuum, followed by catalytic reduction, gives either cholanic or *allo*cholanic acid.

About twelve natural bile acids have been characterised, and a number of others are synthetic. The positions of the hydroxyl groups are any of the following: **3, 6, 7, 11, 12** and **23**, and in almost all of the natural bile acids the configurations of the hydroxyl groups are α (see §4b). Some of the more important natural bile acids are:

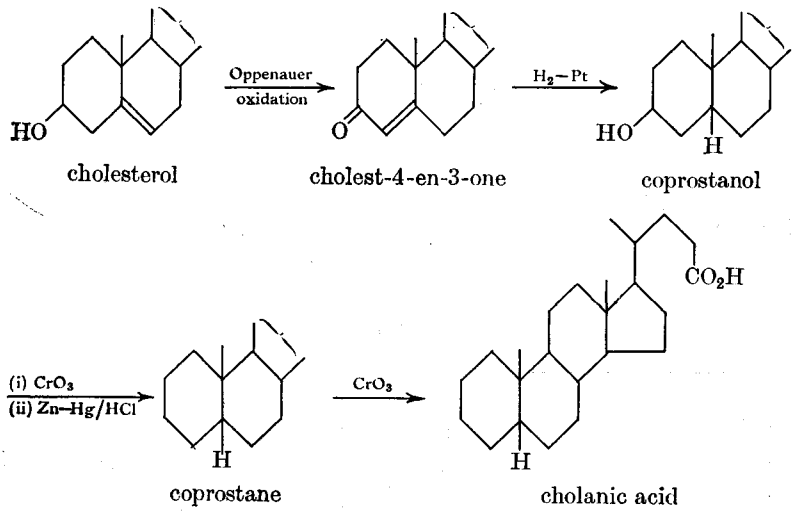
Name	*M.p.*	*Hydroxyl groups*	*Source*
Cholic acid	195°	3α : 7α : 12α	Man, ox
Deoxycholic acid	172°	3α : 12α	Man, ox
Lithocholic acid	186°	3α	Man, ox
Chenodeoxycholic acid . .	140°	3α : 7α	Man, ox, hen
Hyodeoxycholic acid . . .	197°	3α : 6α	Pig

§9. The structures of cholanic acid and *Allo*cholanic acid. These acids may be derived from coprostane and cholestane, respectively, as follows (*cf.* §4c). At the same time, these reactions show the relationship between the bile acids and the sterols (Windaus, 1919).

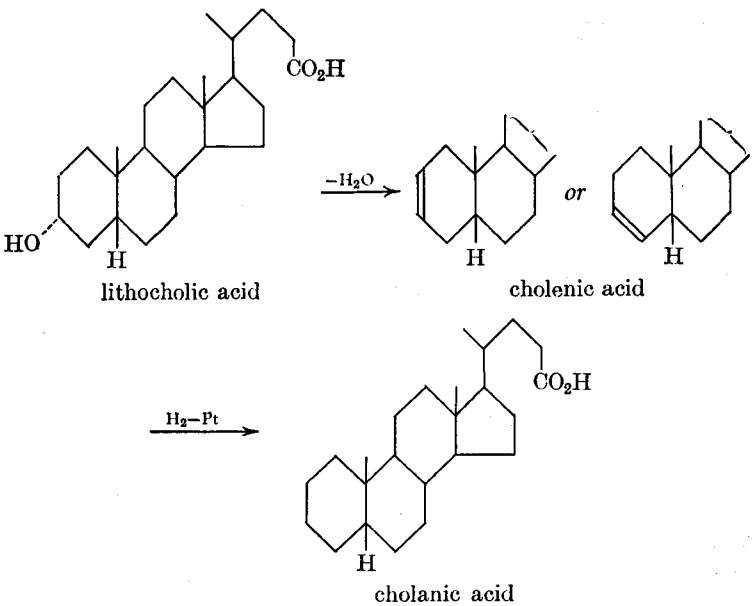
***Allo*cholanic acid.**



Cholanic acid.



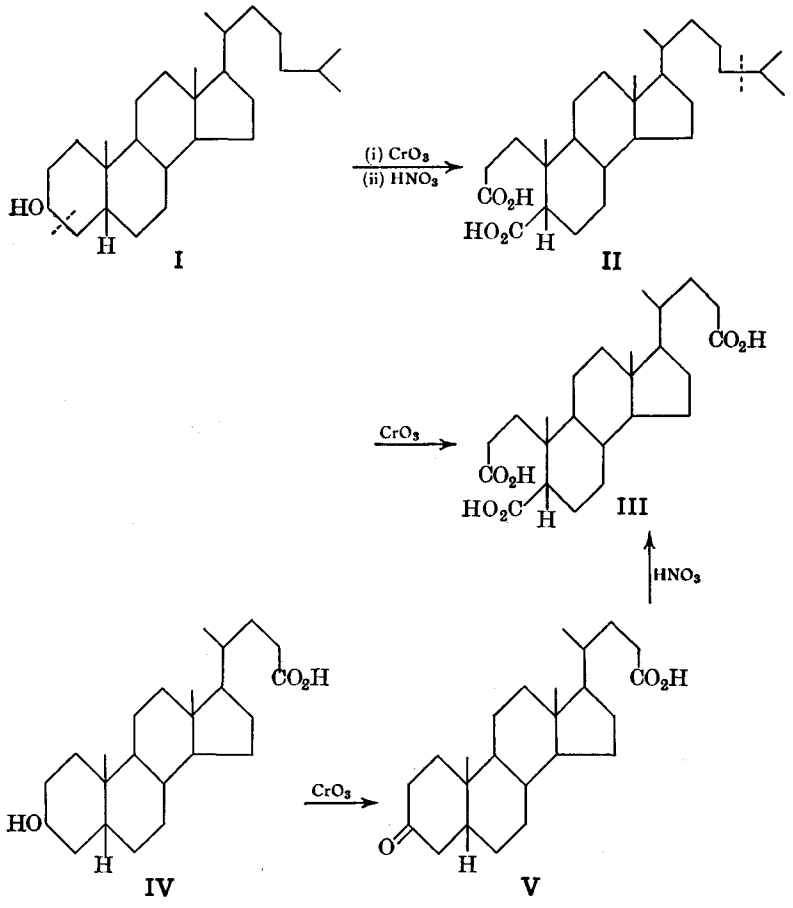
§10. Structure of the bile acids. Since all the bile acids can be converted into either of the cholanic acids, the former are therefore hydroxy derivatives of the latter, *e.g.*, lithocholic acid can be converted into cholanic acid as follows:



According to Fieser *et al.* (1955), cholenic acid is a mixture of the two compounds shown, the chol-3-enic acid being the main constituent.

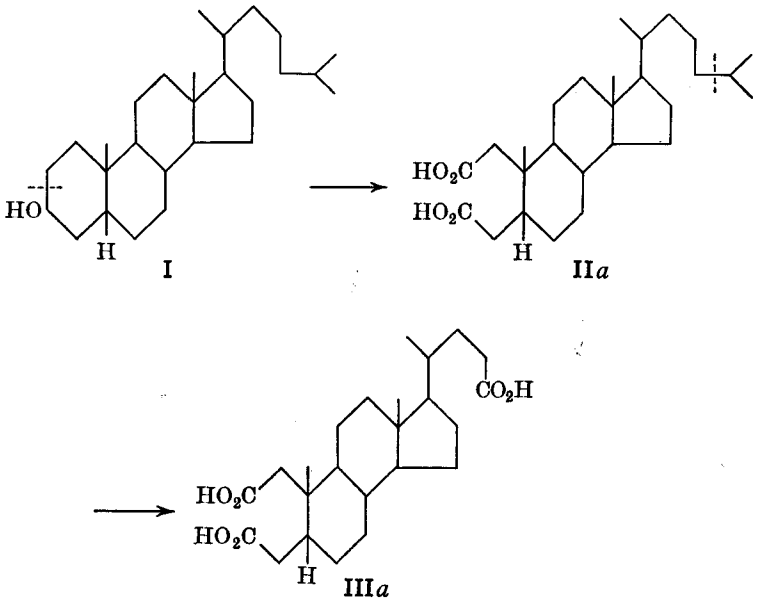
The positions of the hydroxyl groups in the bile acids have been determined by means of oxidative degradation, *e.g.*, the position of the hydroxyl group in lithocholic acid is shown to be at 3 as follows. Cholesterol can be

converted into coprostanol I (see, *e.g.*, §9) which, on oxidation with chromium trioxide, forms a ketone and this, when oxidised with nitric acid, gives a dicarboxylic acid, II. II, on further oxidation with nitric acid, produces the tricarboxylic acid, lithobilianic acid, III. Lithocholic acid, IV, on oxidation with chromium trioxide, forms dehydrolithocholic acid, V, and this, when oxidised with nitric acid, forms III. It therefore follows that the hydroxyl group in lithocholic acid is probably in the same position as in coprostanol, *viz.*, position 3. Thus:

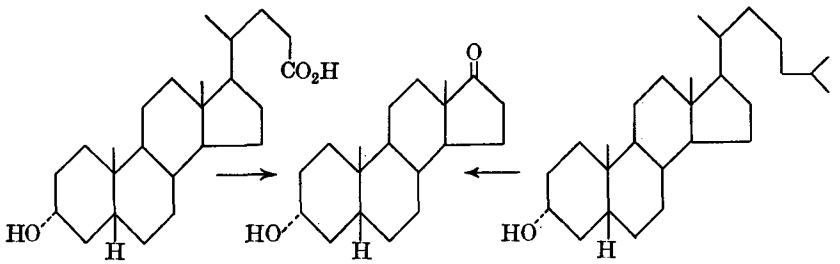


The above evidence is not conclusive, since had the hydroxyl group in lithocholic acid been at position 4, III could still have been obtained. In practice, however, the oxidation of I produces two isomeric acids for II, one being II as shown, and the other II*a*, in which the ring A is opened between C_2 and C_3; this acid, on further oxidation, gives *iso*lithobilianic acid, III*a*. Since the oxidation of lithocholic acid, IV, also produces a mixture of the *same* two acids, III and III*a*, there can be no doubt that the hydroxyl group is at position 3.

The configuration of the hydroxyl group in lithocholic acid has been shown to be α by, *e.g.*, the oxidative degradation of the acetates of lithocholic acid and epicoprostanol to 5-*iso*androsterone (formerly known as 3α-hydroxy-ætiocholan-17-one). Since all of the natural bile acids except one ("β"



hyodeoxycholic acid) can be converted into lithocholic acid, all have therefore the α-configuration for the hydroxyl group at C_3.



lithocholic acid 5-*iso*androsterone epicoprostanol

The bile acids form molecular compounds with various substances. Cholic acid, in particular, forms these molecular compounds with such compounds as fatty acids, esters, alcohols, etc.; these are known as the **choleic acids.** These choleic acids are of the channel complex type (like urea complexes; see Vol. I).

The bile acids discussed in the foregoing account are all derivatives of cholanic or *allo*cholanic acid. There are, however, some bile acids which are not derivatives of the cholanic acids, *e.g.*, in the bile of crocodiles there is the bile acid 3α : 7α : 12α-trihydroxycoprostanic acid, $C_{27}H_{46}O_5$.

SEX HORMONES

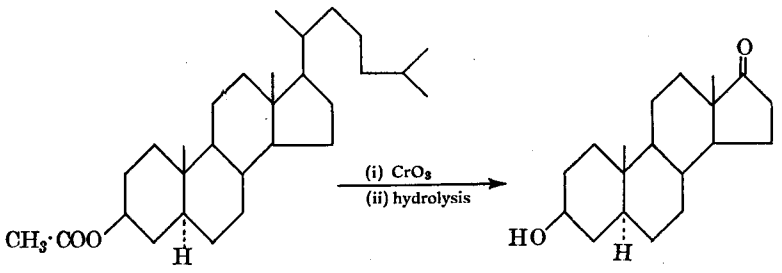
§11. Introduction. Hormones are substances which are secreted by the ductless glands, and only minute amounts are necessary to produce the various physiological reactions in the body. As a group, hormones do not resemble one another chemically, and their classification is based on their physiological activity. There appear to be about 60 different hormones recognised so far, and more than half of these are steroids. The sex hormones

belong to the steroid class of compounds, and are produced in the gonads (testes in the male, and ovaries in the female). Their activity appears to be controlled by the hormones that are produced in the anterior lobe of the pituitary gland. Because of this, the sex hormones are sometimes called the secondary sex hormones, and the hormones of the anterior lobe of the pituitary (which are protein in nature) are called the primary sex hormones.

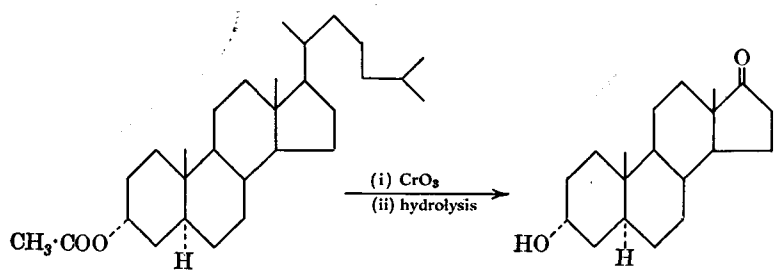
The sex hormones are of three types: the **androgens** (male hormones), the **œstrogens** (female or follicular homones) and **progesterone** (the corpus luteum hormone). The sex hormones are responsible for the sexual processes, and for the secondary characteristics which differentiate males from females.

ANDROGENS

§12. **Androsterone,** $C_{19}H_{30}O_2$, m.p. 184–185°, is dextrorotatory. It was first isolated by Butenandt *et al.* (1931) from male urine (about 15 mg. from 15,000 litres of urine). Androsterone behaves as a saturated compound, and since it forms mono-esters, *one* oxygen atom is present as a hydroxyl group. The functional nature of the other oxygen atom was shown to be oxo, since androsterone forms an oxime, etc. The parent hydrocarbon of androsterone, $C_{19}H_{30}O_2$, is therefore $C_{19}H_{32}$, and since this corresponds to the general formula C_nH_{2n-6}, the molecule is tetracyclic. This led to the suggestion that androsterone probably contains the steroid nucleus, and since it is a hydroxyketone, it was thought that it is possibly related to



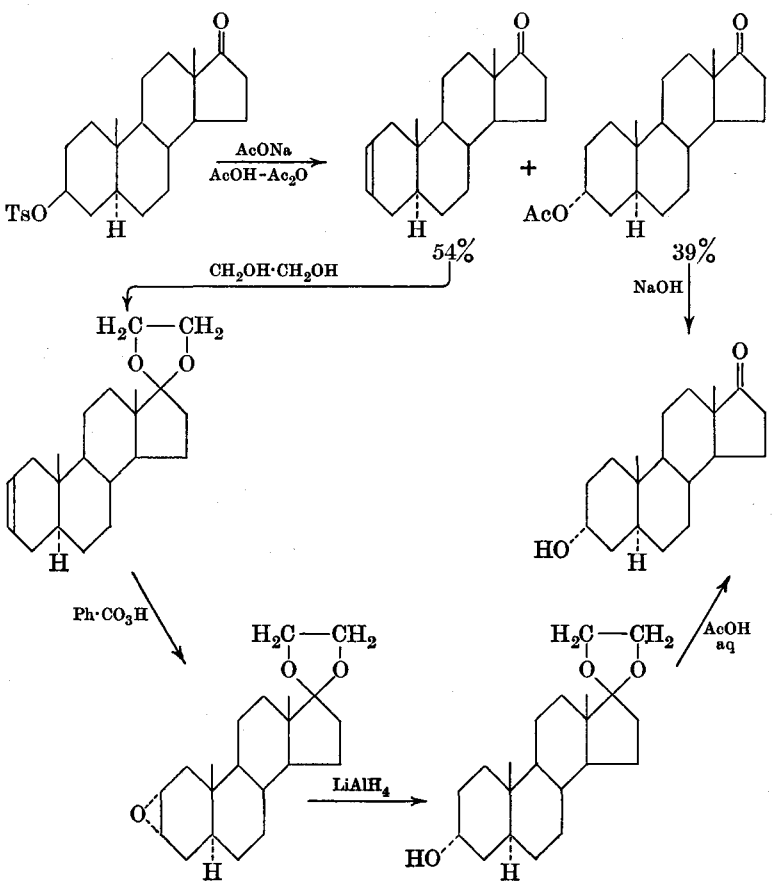
cholestanyl acetate — epiandrosterone



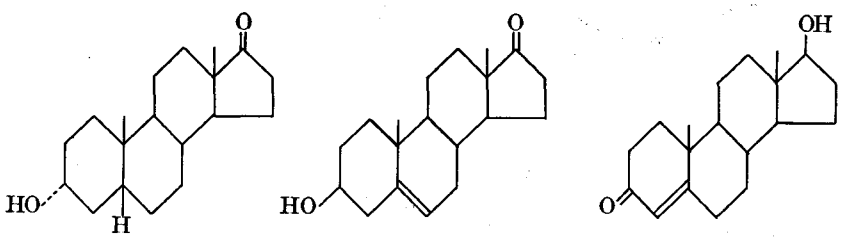
epicholestanyl acetate — androsterone

œstrone (§14). Butenandt (1932) therefore proposed a structure which was proved correct by Ruzicka (1934) as follows. Ruzicka oxidised cholestanyl acetate with chromium trioxide in acetic acid to **epiandrosterone,** a hydroxyketone with the structure proposed for androsterone by Butenandt. When, however, epicholestanyl acetate was oxidised, the product was androsterone. Thus the configuration of the hydroxyl group at C_3 is α and not β as Butenandt suggested. Epiandrosterone (formerly known as *iso*androsterone) has about one-eighth of the activity of androsterone.

Sondheimer *et al.* (1955) have converted epiandrosterone into androsterone, starting with epiandrosterone *p*-toluenesulphonate (*cf.* tosyl esters of sugars, §9. VII).



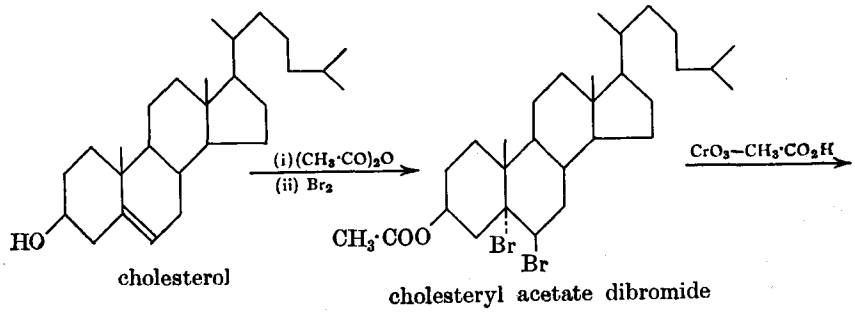
Soon after the discovery of androsterone, Butenandt *et al.* (1934) isolated two other hormones from male urine, 5-*iso*androsterone and dehydroepiandrosterone. Then Laqueur (1935) isolated the hormone testosterone from steer testes (10 mg. from 100 kg. of testes).



5-*iso*androsterone dehydroepiandrosterone testosterone

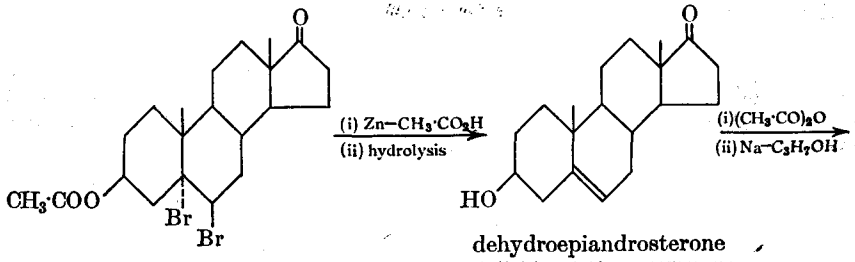
§13. **Testosterone,** $C_{19}H_{28}O_2$, m.p. 155°, is dextrorotatory. Testosterone has been produced commercially by the following method of Butenandt

(1935) and Ruzicka (1935); the Oppenauer oxidation step in this method was introduced by Oppenauer (1937). This preparation of testosterone establishes the structure of this hormone. This method has been improved

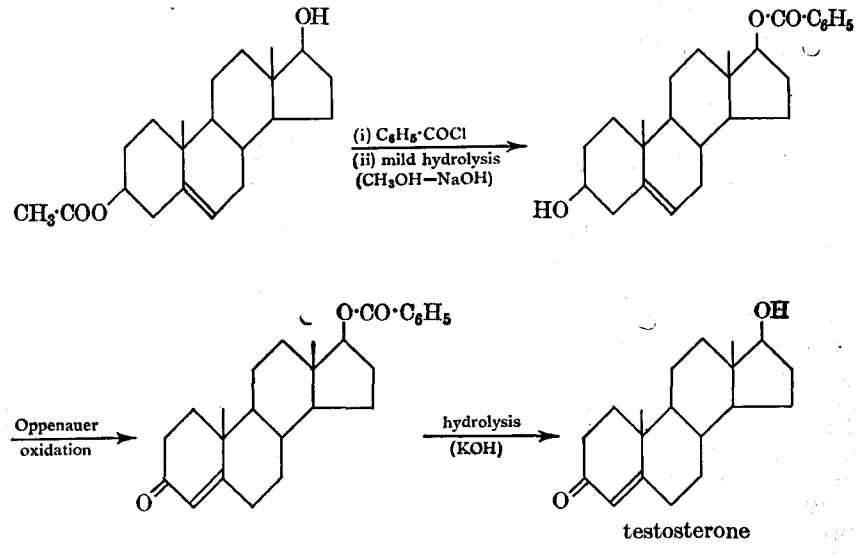


cholesterol

cholesteryl acetate dibromide



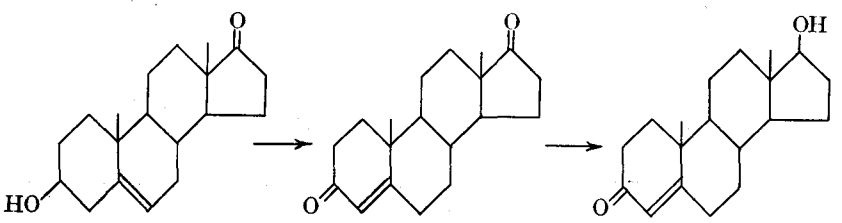
dehydroepiandrosterone



testosterone

by Mamoli (1938), who converted dehydroepiandrosterone into testosterone by means of micro-organisms; the first stage uses an oxidising yeast in the presence of oxygen, and the second stage a fermenting yeast.

Elisberg *et al.* (1952) have shown that sodium borohydride selectively reduces the 3-keto group in the presence of others at 11, 12, 17 or 20. On



dehydroepiandrosterone androst-4-ene-3:17-dione testosterone

the other hand, Norymberski *et al.* (1954) have shown that if there is a double bond in position 4 : 5, then the keto group at 17 or 20 is preferentially reduced to that at 3. Thus androst-4-ene-3 : 17-dione is reduced to testosterone by sodium borohydride (*cf.* §3 i). Johnson *et al.* (1960) have adapted Johnson's synthesis of equilenin (§17) to provide an improved synthesis of testosterone.

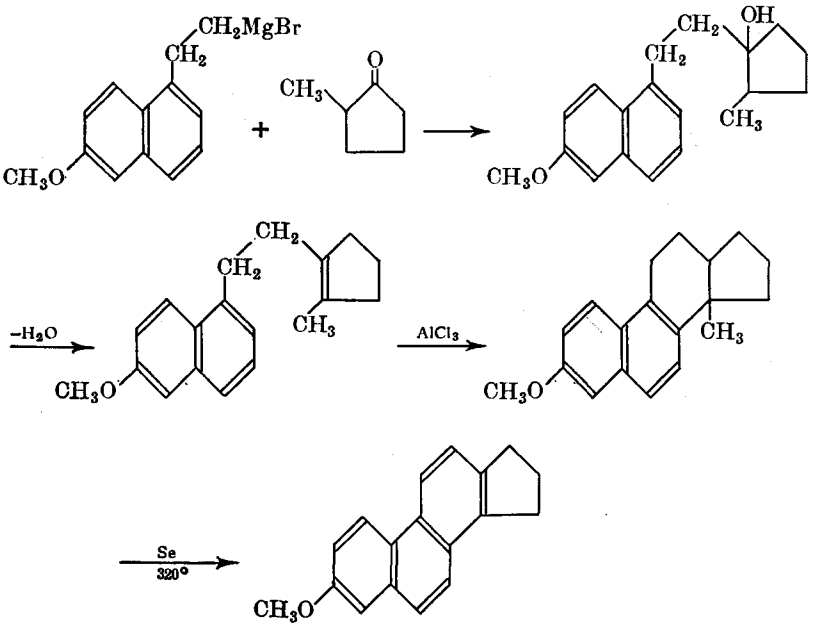
It appears that testosterone is the real male sex hormone in the body; the others are metabolic products of testosterone. The ketonic steroids are separated from the non-ketonic steroids (all from urine) by means of Girard's reagents (P and T); the ketonic compounds form *soluble* derivatives, and may be regenerated by hydrolysis (see also Vol. I). Many other hormones have also been isolated from urine.

ŒSTROGENS

§14. Œstrone. 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 **œstrone** from the urine of pregnant women. Œstrone is the first known member of the sex hormones, and soon after its discovery two other hormones were isolated, œstriol and œstradiol.

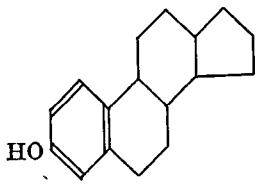
(+)-Œstrone, m.p. 259°, has the molecular formula $C_{18}H_{22}O_2$. It behaves as a ketone (forms an oxime, etc.), and contains one hydroxyl group (it forms a monoacetate and a monomethyl ether). Furthermore, this hydroxyl group is *phenolic*, since œstrone couples with diazonium salts in alkaline solution (this reaction is typical of phenols). When distilled with zinc dust, œstrone forms chrysene; this led to the suggestion that œstrone is related to the steroids (*cf.* §1). The X-ray analysis of œstrone 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, œstrone forms octahydro-œstrone, $C_{18}H_{30}O_2$. 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 œstrone is supported by measurements of the molecular refractivity and the ultraviolet absorption spectrum.

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

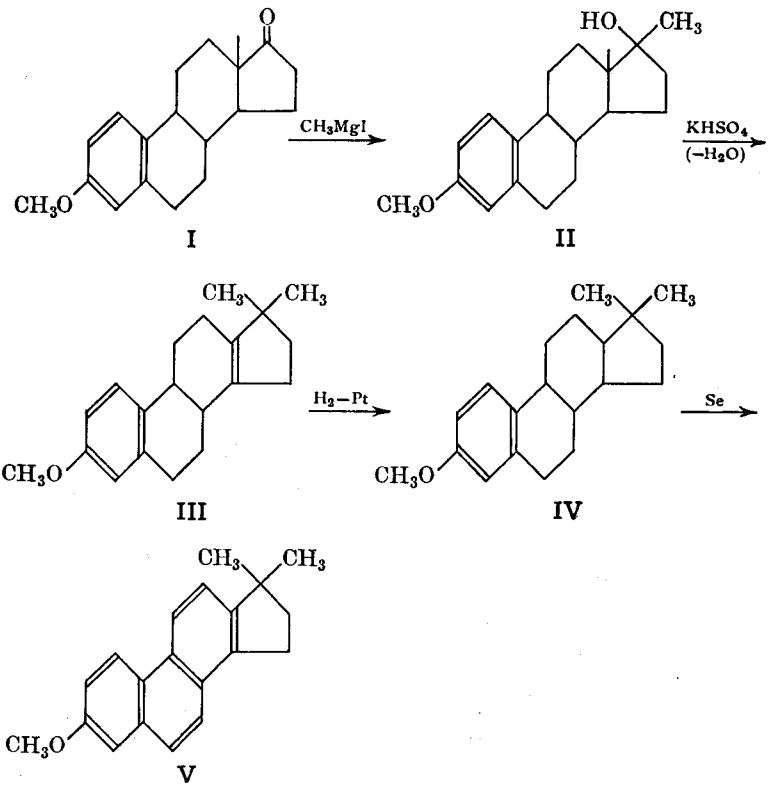


7-methoxy-1:2-*cyclo*pentenophenanthrene

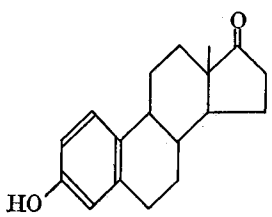
Thus the benzene ring in œstrone is ring A, and the (phenolic) hydroxyl group is at position 3; hence the skeleton of œstrone is:



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 œstrone, 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-*cyclo*pentenophenanthrene, 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 (see overleaf).



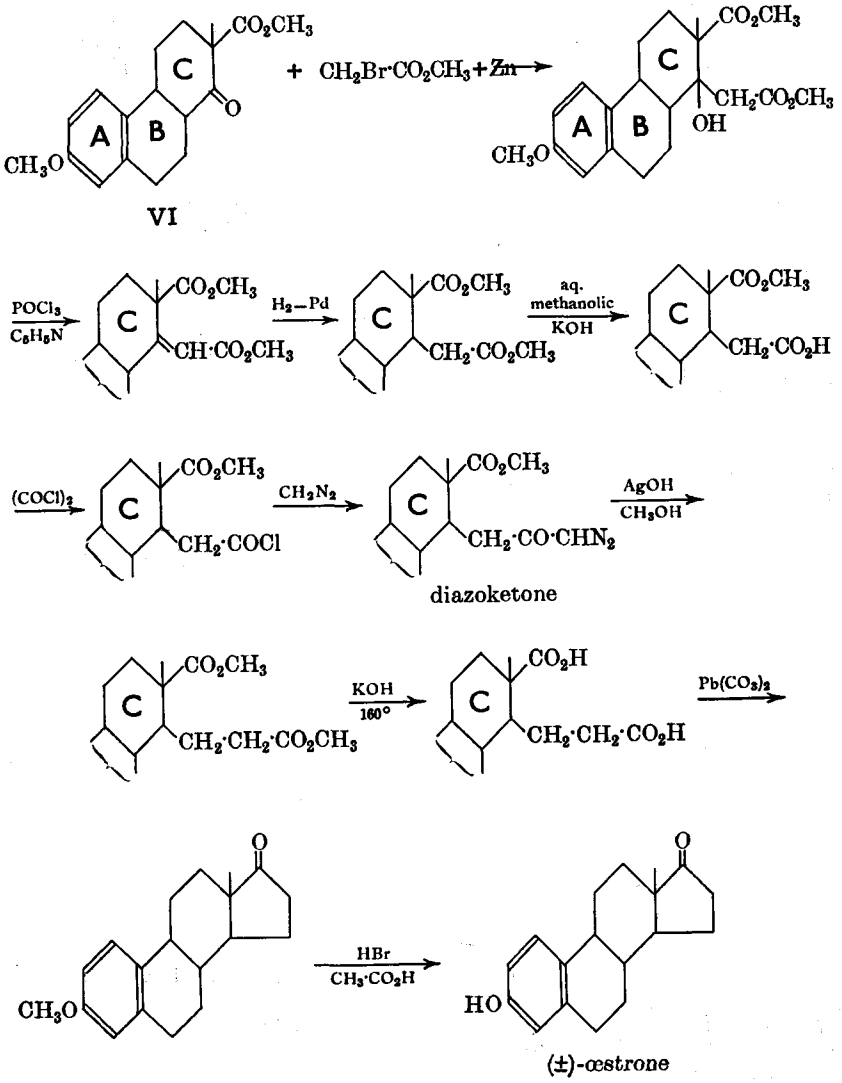
The structure of V has been confirmed by synthesis (Cook *et al.*, 1935). Thus the structure of œstrone is:



œstrone

This has been confirmed by the 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.

The stereochemical problems involved in the synthesis of œstrone are not so complicated as in cholesterol, since only four asymmetric carbon atoms are present in the hormone (*cf.* §3). VI contains 3 asymmetric carbon atoms, and so four racemates are possible. Three have been isolated by Anner and Miescher, and one of these was converted into (±)-œstrone (C/D *trans*) and the stereoisomer (C/D *cis*) as shown above. These were separated and the

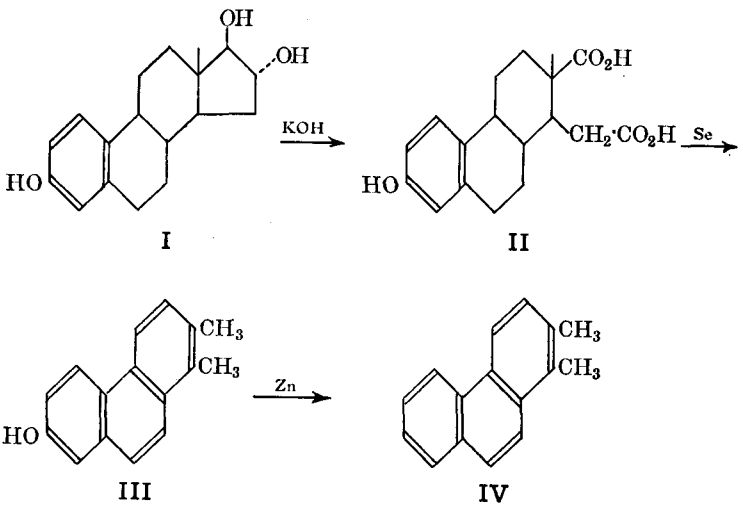


(±)-œstrone resolved with (−)-menthoxyacetic acid. The (+)-enantiomorph that was obtained was shown to be identical with the natural compound.

Johnson *et al.* (1958, 1962) have carried out a total synthesis of œstrone; each step in their synthesis was stereoselective. Hughes *et al.* (1960) have reported total syntheses of œstrone which appear to be simpler than any previous method and just as efficient.

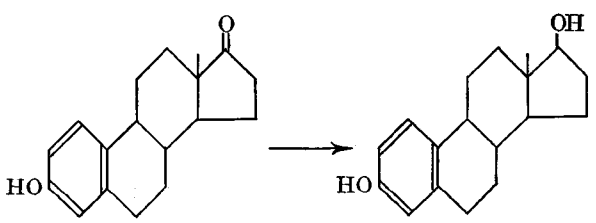
§15. Œstriol, $C_{18}H_{24}O_3$, m.p. 281°, was isolated from human pregnancy urine by Marrian (1930). Since œstriol forms a triacetate, three hydroxyl groups must be present in the molecule. One was shown to be phenolic (*cf.* œstrone), and the other two secondary alcoholic, since, on oxidation, a diketone is produced. Furthermore, X-ray analysis indicates that the two alcoholic groups are in the *vicinal* position (*i.e.*, 1 : 2-). When œstriol is heated with potassium hydrogen sulphate, one molecule of water is removed

and œstrone is produced. It therefore follows that œstriol has the same carbon skeleton as œstrone, and that the two alcoholic groups in œstriol are at positions 16 and 17. Structure I for œstriol fits the above facts, and is supported by the following evidence. When fused with potassium hydroxide, œstriol forms marrianolic acid, II, and this, on dehydrogenation with selenium, is converted into a hydroxydimethylphenanthrene, III, which, on distillation with zinc dust, gives a dimethylphenanthrene, IV. The structure of IV was shown to be 1 : 2-dimethylphenanthrene by synthesis, and since marrianolic acid forms an anhydride when heated with acetic anhydride, it therefore follows that œstriol contains a phenanthrene nucleus and a five-membered ring, the position of the latter being 1 : 2 (where the two methyl groups are in IV). Finally, the structure of III was shown to be 7-hydroxy-1 : 2-dimethylphenanthrene by synthesis (Haworth *et al.*, 1934), and so if I is the structure of œstriol, the degradation to the phenanthrene derivatives may be explained as follows:

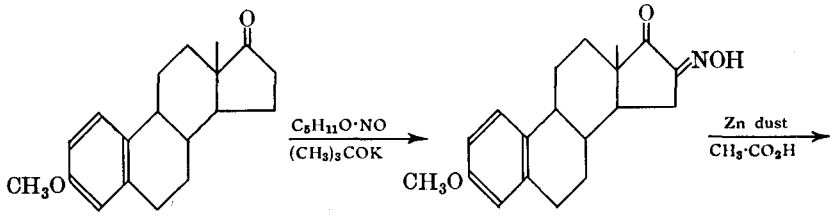


The chemical relationship between œstrone, œstriol and œstradiol (§16) is shown by the following reactions.

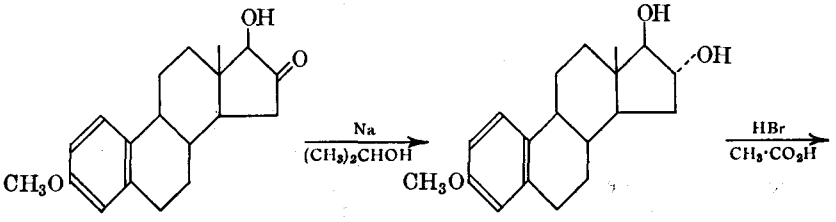
(i) Œstrone may be reduced to œstradiol by catalytic hydrogenation, by aluminium *iso*propoxide (the Meerwein–Ponndorf–Verley reduction), or by lithium aluminium hydride.

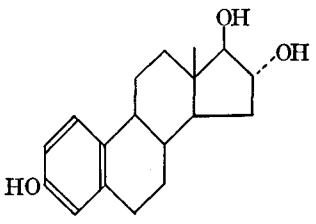


(ii) Œstriol may be converted into œstrone by the action of potassium hydrogen sulphate (see above), and œstrone may be converted into œstriol as follows (Huffman *et al.*, 1947, 1948).

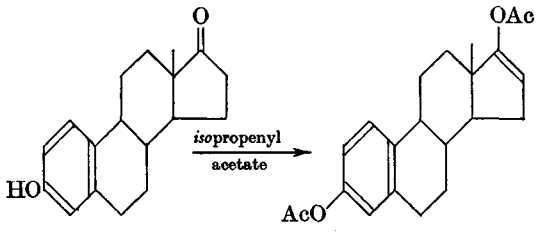


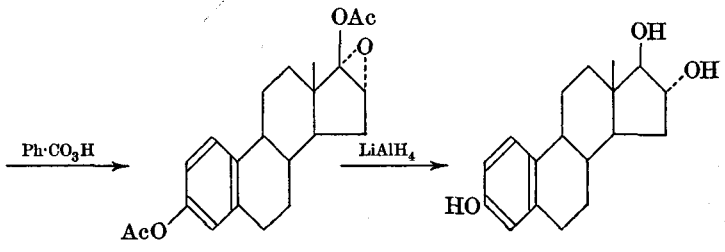
methyl ether of œstrone





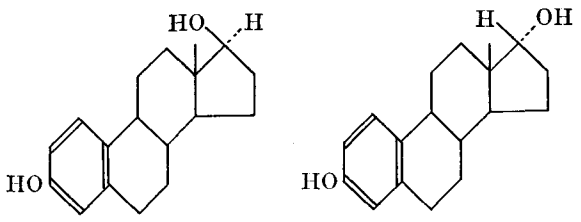
Leeds *et al.* (1954) have converted œstrone into œstriol by a simpler method:





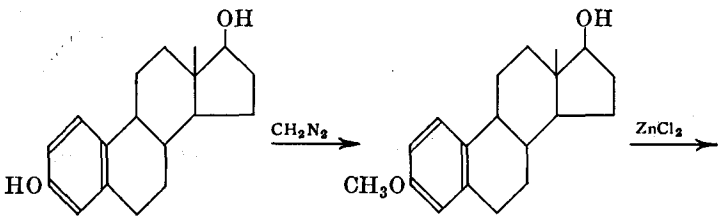
Œstriol is more soluble than œstrone in water, and is more potent than either œstrone or œstradiol when taken orally.

§16. Œstradiol, $C_{18}H_{24}O_2$. There are two stereoisomeric œstradiols, α and β; the α-isomer is much more potent than the β-.

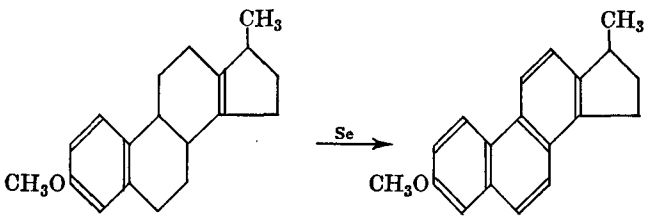


α-œstradiol (œstradiol-17β) β-œstradiol (œstradiol-17α)

α-Œstradiol was first obtained by the reduction of œstrone (see §15), but later it was isolated from the ovaries of sows (Doisy *et al.*, 1935). When the phenolic methyl ether of œstradiol is heated with zinc chloride, a molecular rearrangement occurs, the angular methyl group migrating to the *cyclo*pentane ring D (*cf.* §2 viii. X). This compound, when dehydrogenated with selenium, produces 7-methoxy-3′-methyl-1 : 2-*cyclo*pentenophenanthrene, the structure of which has been ascertained by synthesis (Cook *et al.*, 1934). Thus the structure of œstradiol is established.



α-œstradiol



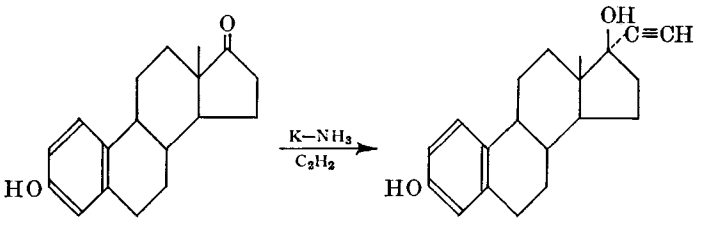
7-methoxy-3′-methyl-1:2-*cyclo*pentenophenanthrene

Velluz *et al.* (1960) have synthesised œstradiol starting from 6-methoxy-1-tetralone; this is therefore a total synthesis of the hormone.

β-Œstradiol has been isolated from the pregnancy urine of mares (Wintersteiner *et al.*, 1938). α-Œstradiol is much more active than œstrone, whereas β-œstradiol is much less active. It appears that œstradiol is the real hormone, and that œstrone and œstriol are metabolic products. It might be noted here that when the second œstradiol was discovered, the earlier one was arbitrarily designated as the "α"-isomer. Subsequently, this

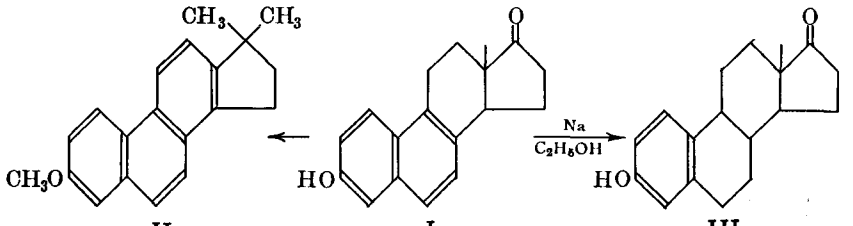
isomer was shown to have the 17β configuration, and the "β"-isomer the 17α configuration.

A very active synthetic œstrogen is **17α-ethinylœstradiol,** and has the advantage that it is very active when taken orally. This synthetic compound has been prepared by the action of acetylene on œstrone in a solution of liquid ammonia containing potassium.



œstrone 17α-ethinylœstradiol

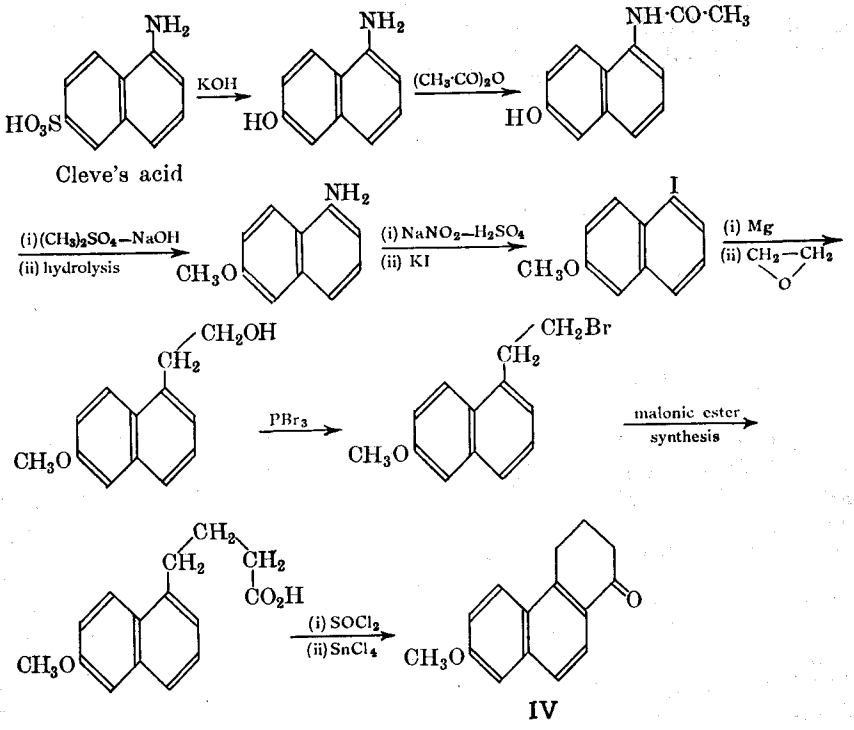
§17. **(+)-Equilenin,** $C_{18}H_{18}O_2$, m.p. 258–259°, has been isolated from the urine of pregnant mares by Girard *et al.* (1932); it is not a very potent œstrogen. The reactions of equilenin show that a phenolic hydroxyl group and a ketonic group are present, and also that the molecule contains five double bonds (*cf.* œstrone, §14). When the methyl ether of equilenin is treated with methylmagnesium iodide, then the alcohol dehydrated, catalytically reduced and then dehydrogenated with selenium, the product is 7-methoxy-3′ : 3′-dimethyl-1 : 2-*cyclo*pentenophenanthrene, II (*cf.* œstrone, §14). Thus the structure of equilenin is the same as that of œstrone, except that the former has two more double bonds than the latter (Cook *et al.*, 1935). Now the absorption spectrum of equilenin shows that it is a naphthalene derivative. Thus, since ring A in œstrone is benzenoid, it appears probable that ring B in equilenin is also benzenoid, *i.e.*, rings A and B form the naphthalene nucleus in equilenin. All the foregoing reactions of equilenin may be readily explained by assuming that I is its structure, and further evidence that has been given to support this is the claim by Marker *et al.* (1938) that equilenin may be reduced to œstrone, III, by sodium and ethanol. This reduction, however, has apparently never been substantiated (*cf.* Dauben *et al.*, 1956).



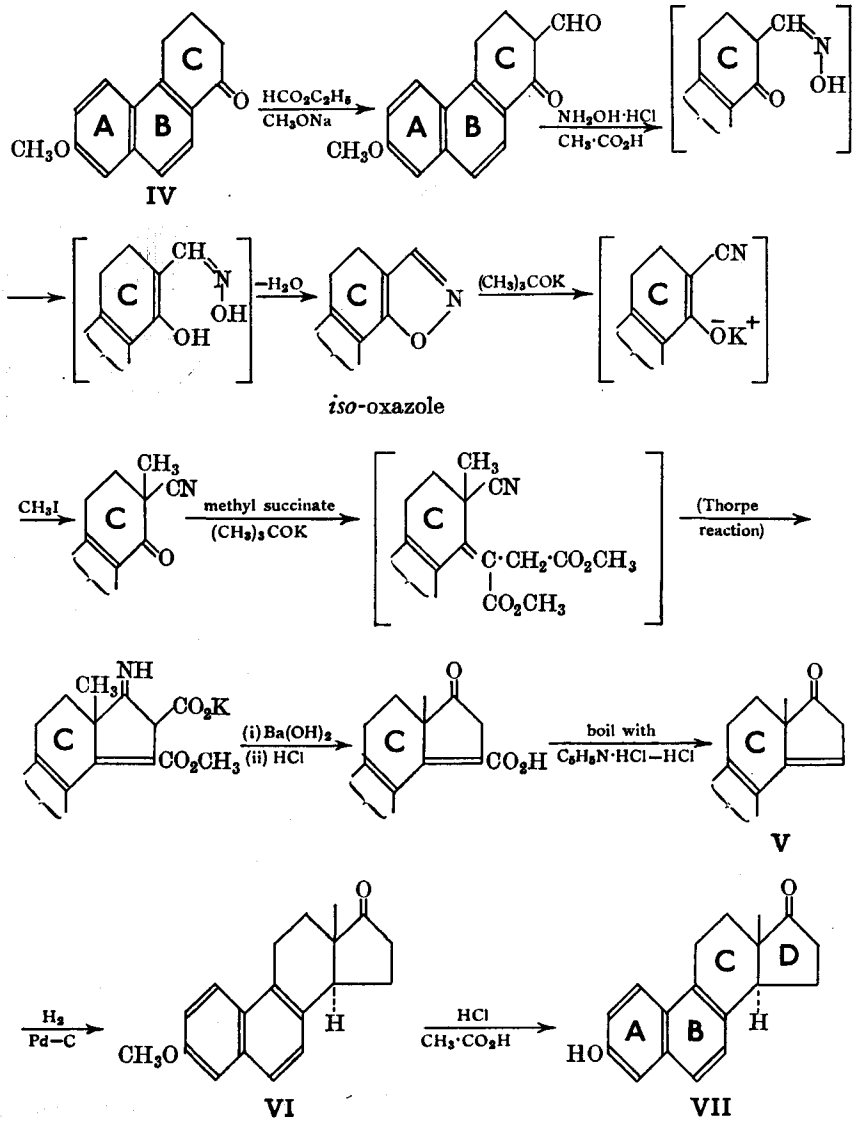
II I equilenin III œstrone

This structure of equilenin has been confirmed by synthesis. The first synthesis was by Bachmann *et al.* (1940), but was somewhat improved by Johnson *et al.* (1947). In the following chart, compound IV is synthesised by the method of Bachmann, and the rest of the synthesis is that of Johnson,

who started with compound IV (Johnson's synthesis involves fewer steps than Bachmann's).

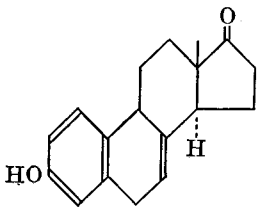


Johnson's synthesis starting from IV.



Reduction of V gives a mixture of (±)-equilenin methyl ether, VI (rings C/D *trans*), and *iso*equilenin methyl ether (rings C/D *cis*); these are separated by fractional crystallisation from acetone-methanol, the equilenin derivative being the less soluble isomer. Product VII is (±)-equilenin, and is resolved *via* the menthoxyacetic ester. The (+)-equilenin so obtained is identical with the natural product. It should be noted here that equilenin contains only two asymmetric carbon atoms, and so the stereochemical problems involved are far simpler than those for cholesterol and œstrone.

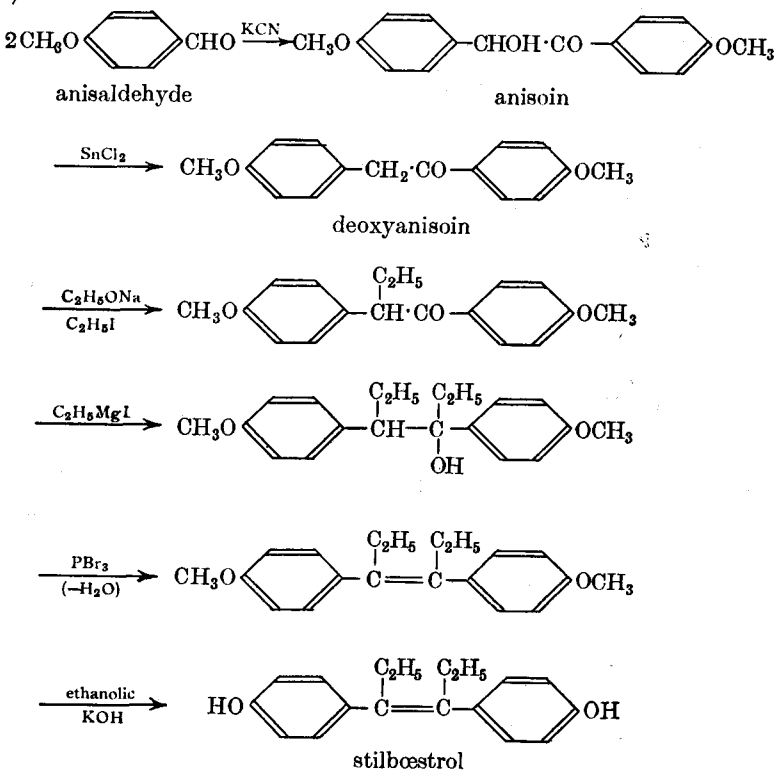
§17a. (+)-Equilin, $C_{18}H_{20}O_2$, m.p. 238–240°, has also been isolated from the urine of pregnant mares (Girard *et al.*, 1932), and its structure has been shown to be:



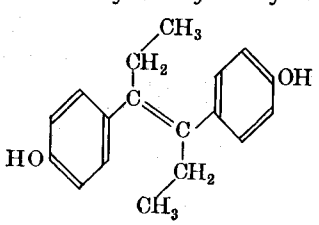
equilin

§18. Artificial hormones. Many compounds with œstrogenic activity but not of steroid structure have been prepared synthetically.

Stilbœstrol (4 : 4′-dihydroxydiethylstilbene) was prepared by Dodds *et al.* (1939) as follows:

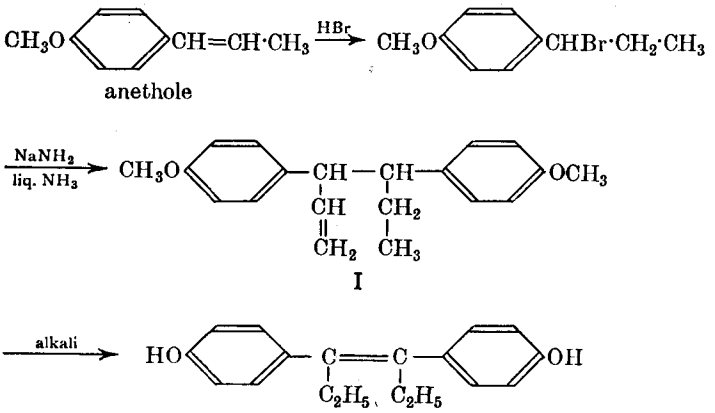


The above structure of stilbœstrol can exist in two geometrical isomeric forms; it is the *trans*-form which is the active substance, and this configuration has been confirmed by X-ray analysis (Crowfoot *et al.*, 1941).



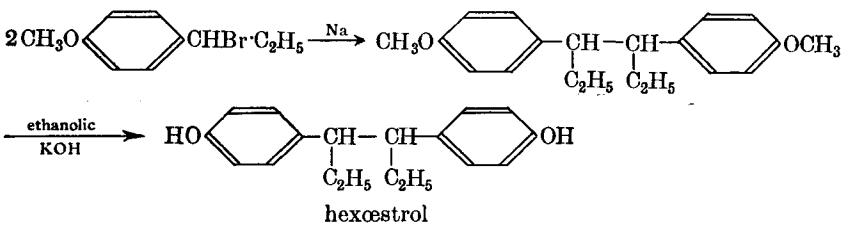
trans-stilbœstrol

Kharasch *et al.* (1943) have introduced a simpler synthesis of stilbœstrol. Anethole is treated with hydrobromic acid and the product, anethole hydrobromide, is then treated with sodamide in liquid ammonia. The resulting compound, I, gives stilbœstrol on demethylation and isomerisation in the presence of alkali. The structure of I is uncertain, but it is believed to be the one given.



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

Hexœstrol (dihydrostilbœstrol) may be prepared from anethole hydrobromide as follows:



hexœstrol

The active form is the *meso*-isomer (as shown by X-ray crystallography by Crowfoot *et al.*, 1941), and this compound appears to be the most potent of the œstrogens.

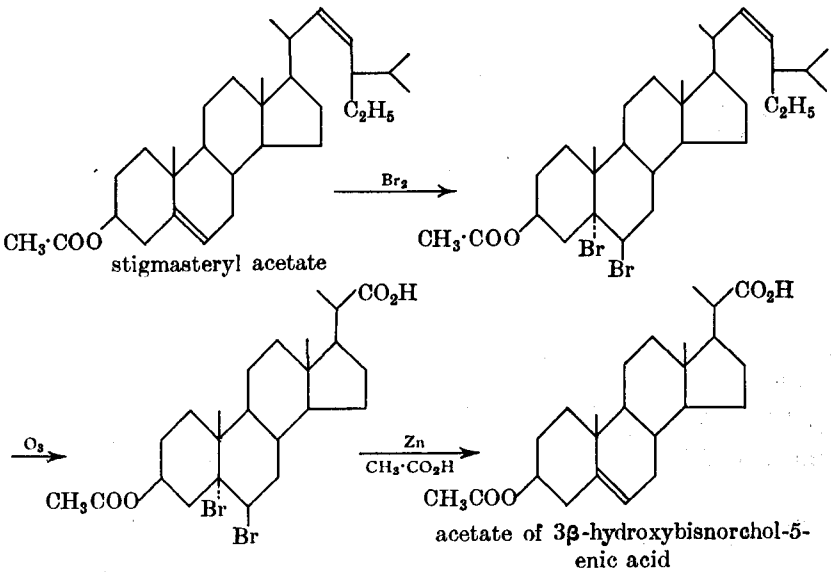
GESTOGENS

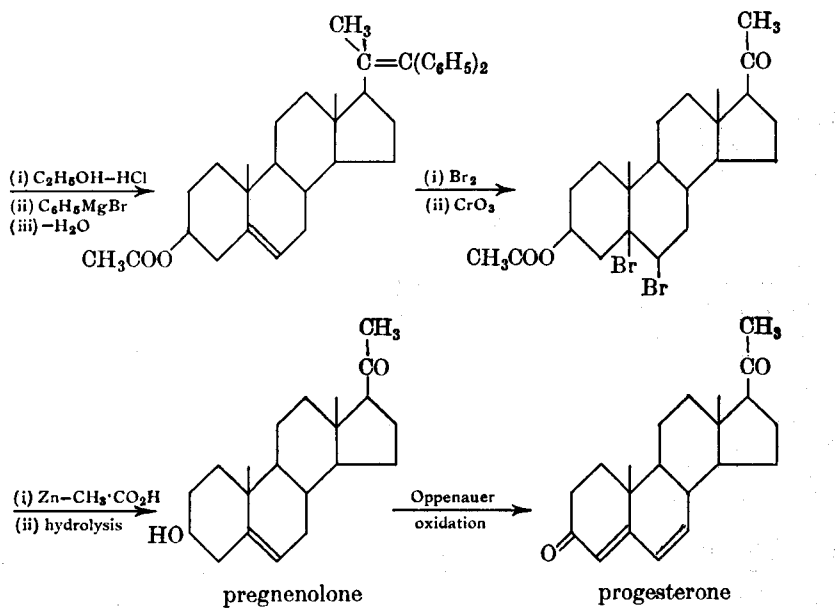
§19. Progesterone, $C_{21}H_{30}O_2$, m.p. 128°, was first isolated in a pure form by Butenandt *et al.* (1934) from the *corpora lutea* of pregnant sows.

The chemical reactions of progesterone show that there are two keto groups present, and since on catalytic reduction three molecules of hydrogen are added to form the dialcohol $C_{21}H_{36}O_2$, it therefore follows that progesterone contains one double bond (four hydrogen atoms are used to convert the two keto groups to alcoholic groups). Thus the parent hydrocarbon of progesterone is $C_{21}H_{36}$, and since this corresponds to the general formula C_nH_{2n-6}, progesterone is therefore tetracyclic. Furthermore, X-ray studies have shown that progesterone contains the steroid nucleus, and this is further supported by the fact that progesterone may be prepared from, *e.g.*, stigmasterol and cholesterol. These preparations also show the structure of progesterone, but do not provide conclusive evidence for the position of the double bond in progesterone, since the results can be interpreted equally well on the assumption that the double bond is 4 : 5 or 5 : 6. The

absorption spectrum of progesterone, however, shows that it is an $\alpha:\beta$-unsaturated ketone, and this suggests that the position of the double bond is 4 : 5 (see below). Finally, progesterone has also been synthesised from diosgenin and from pregnanediol, and the preparation from the latter, taken in conjunction with the others, definitely shows that the position of the double bond in progesterone is 4 : 5.

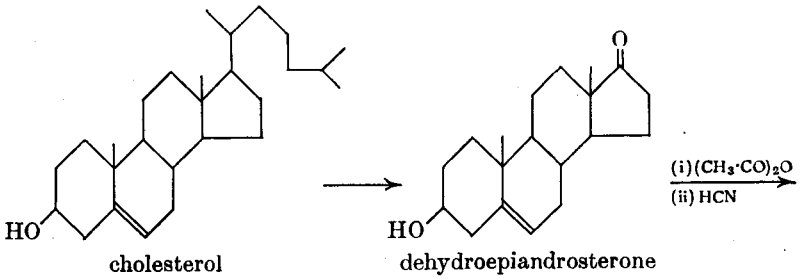
(i) *Progesterone from stigmasterol* (Butenandt *et al.*, 1934, with improvements by other workers).



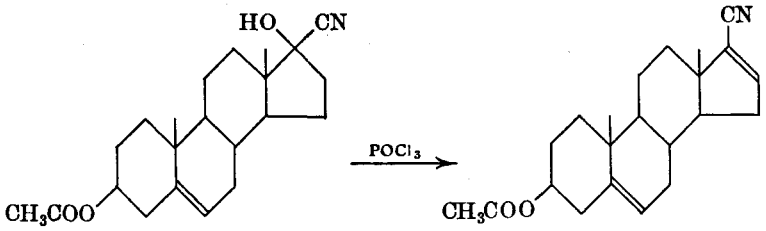


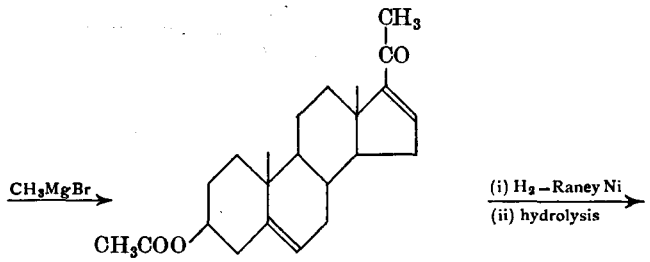
Pregnenolone has also been isolated from the *corpus luteum*.

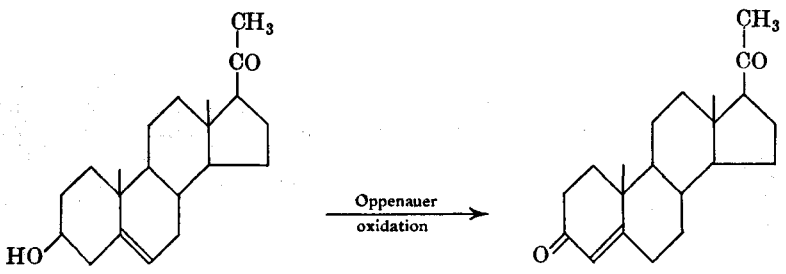
(ii) *Progesterone from cholesterol* (Butenandt *et al.*, 1939). Cholesterol is first converted into dehydroepiandrosterone (see §13), and then as follows:



cholesterol dehydroepiandrosterone

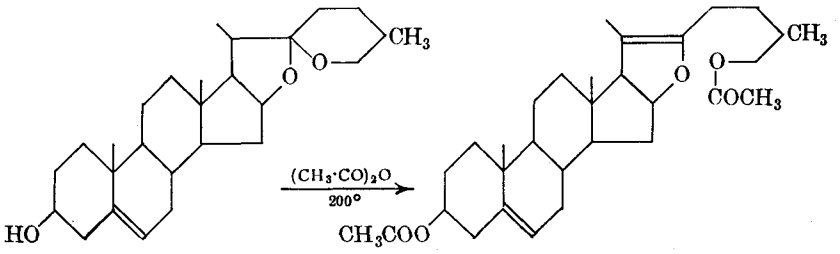




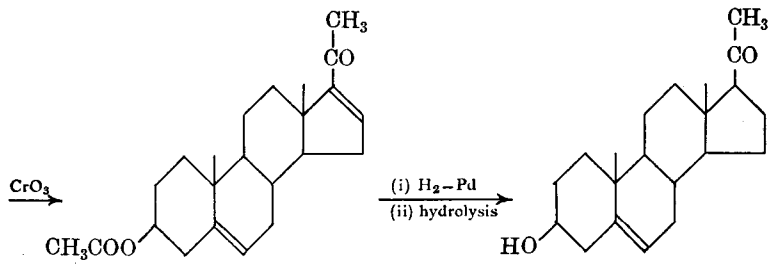


pregnenolone progesterone

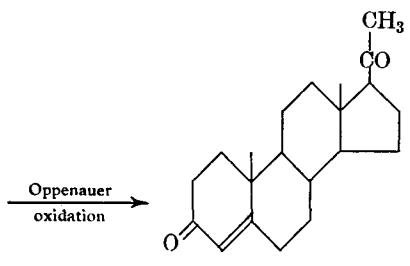
(iii) *Progesterone from diosgenin* (Marker *et al.*, 1940, 1941). Diosgenin (a sapogenin) occurs as a glycoside in the root of *Trillium erectum*.



diosgenin



pregnènolone

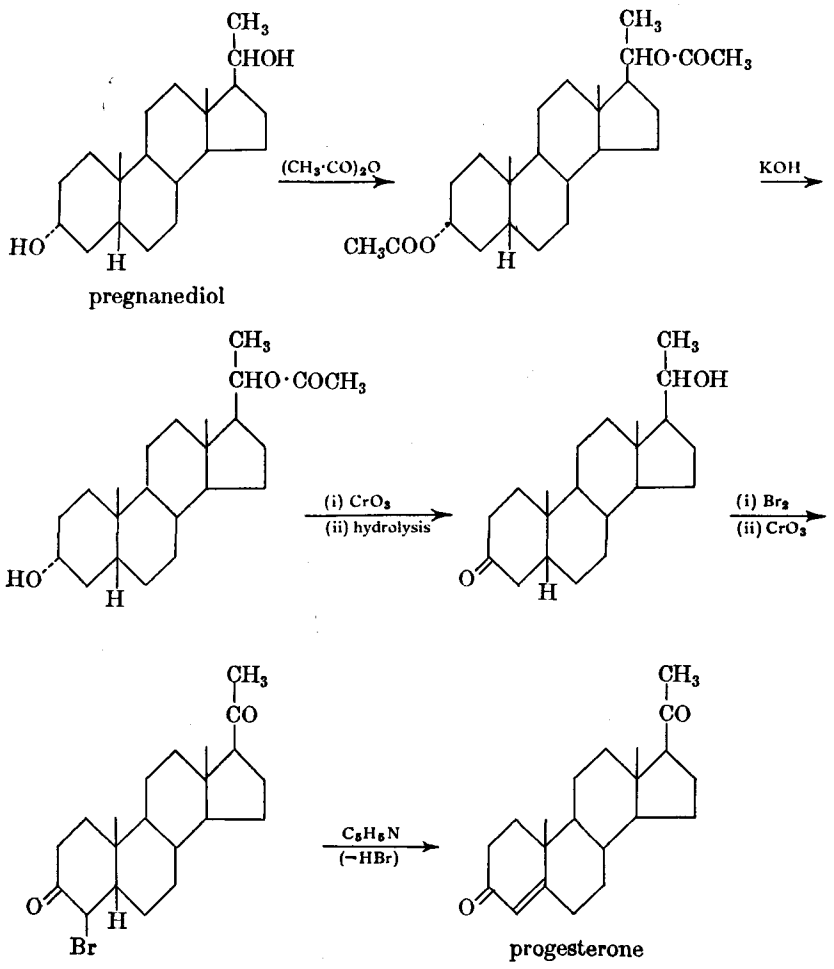


progesterone

Saponins and Sapogenins. Saponins are plant glycosides, and the aglycon is known as the sapogenin (*cf.* §24. VII). Saponins are very powerful emulsifiers, and derive their name from this property; they are used as detergents. There are two groups of saponins, the steroid and the triterpenoid saponins, and these two groups may be distinguished by the fact that only the former group gives Diels' hydrocarbon on distillation with selenium; the triterpenoid group gives mainly naphthalene or picene derivatives (*cf.* §1).

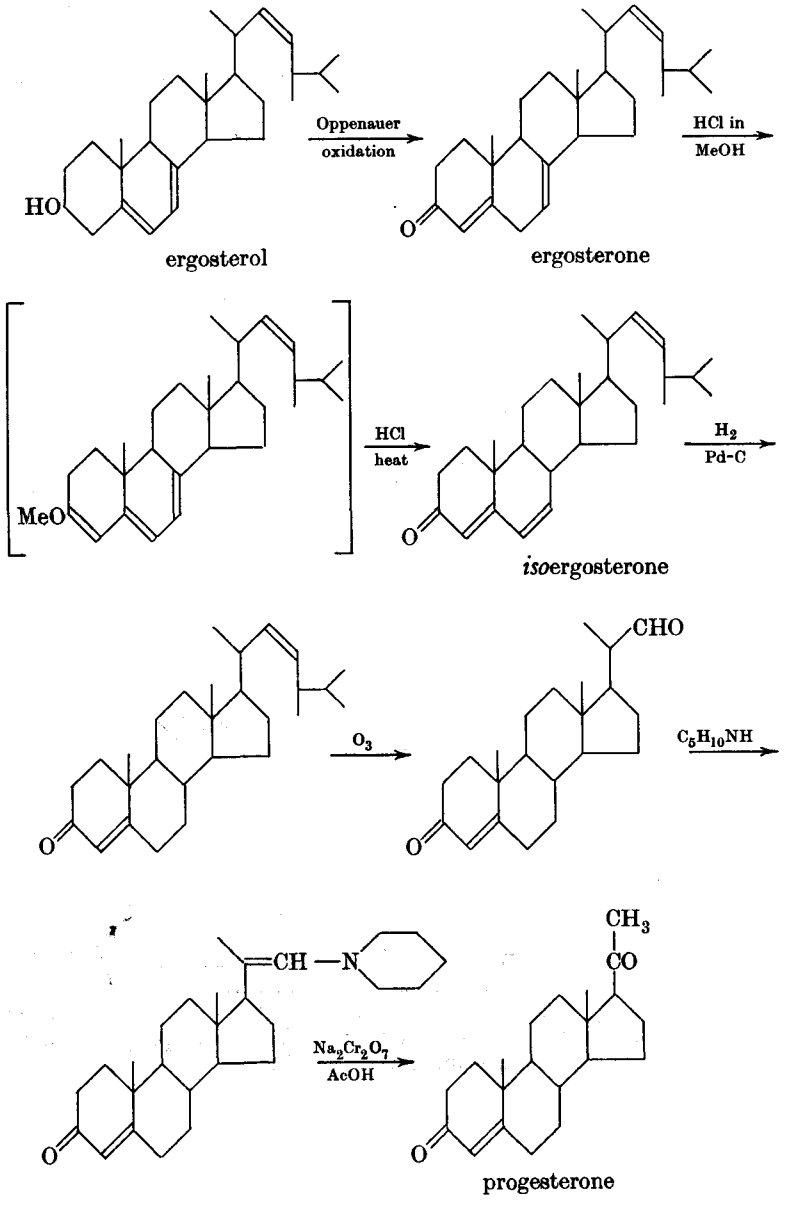
Digitonin is a steroid saponin; it causes hæmolysis of the red blood cells.

(iv) *Progesterone from pregnanediol* (Butenandt *et al.*, 1930).



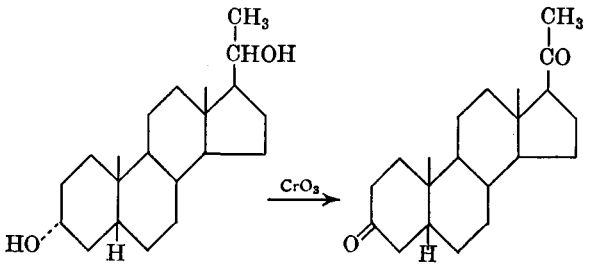
In the above reactions, bromination might have occurred in position 2; in this case the position of the double bond would have been 1 : 2. This is impossible, since the preparation of progesterone by methods (i) to (iii) shows that the double bond must be 4 : 5 or 5 : 6. Thus the preparation from pregnanediol proves that the double bond is 4 : 5.

(v) *Progesterone from ergosterol* (Shepherd *et al.*, 1955). This appears to be the most practical synthesis.

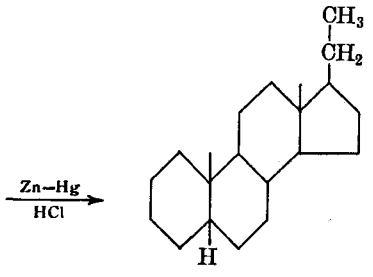


§20. **Pregnane-3α : 20α-diol,** $C_{21}H_{36}O_2$, was isolated from human pregnancy urine by Marrian (1929); it is biologically inactive, and is the main metabolic product of progesterone. The functional nature of the two oxygen atoms was shown to be secondary alcoholic, and since pregnanediol is saturated, the parent hydrocarbon is $C_{21}H_{36}$, and so the molecule is tetracyclic. Pregnanediol gives the haloform reaction; thus a $CH_3{\cdot}CHOH{\cdot}$ group is pre-

sent (see Vol. I). When oxidised, pregnanediol is converted into the diketone pregnanedione and this, on the Clemmensen reduction, forms pregnane, $C_{21}H_{36}$. This is identical with 17-ethylætiocholane, a compound of known structure. Thus pregnanediol contains the steroid nucleus, and the position of the side-chain is 17. Finally, the relationship between pregnanediol and progesterone shows that the former contains one hydroxyl group at position 3. Further work showed that the configuration of the 3-hydroxyl group is α. Thus:



pregnanediol pregnanedione

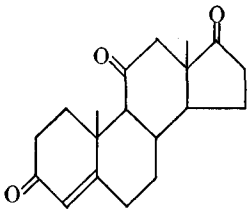


pregnane

ADRENAL CORTICAL HORMONES

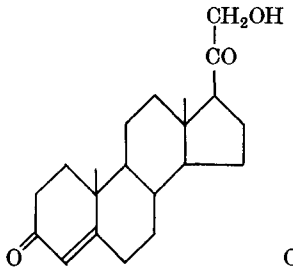
§21. Introduction. In the adrenal glands (of mammals) there are two regions, the *medulla* which produces adrenaline (see §12. XIV), and the *cortex* which produces steroid hormones. The production of these adreno-cortical hormones is controlled by the hormone produced in the anterior lobe of the pituitary, the so-called adrenocorticotrophic hormone, ACTH. The absence of the *corticoids* causes loss of sodium from the body.

§22. Adrenal cortical hormones. About 28 steroids have been isolated from the extract of the adrenal cortex, and their structures have been elucidated mainly by Kendall *et al.* (1935), Wintersteiner (1935–) and Reichstein *et al.* (1936–). Only six of these 28 compounds are physiologically active, fourteen are inactive and are produced by the reduction of the active hormones, and the remaining six are œstrone, progesterone, 17α-hydroxyprogesterone and adrenosterone, and two other compounds that are apparently produced by oxidation during the isolation of the hormones from the cortical extract. Adrenosterone is as shown, and possesses androgenic activity (see overleaf).

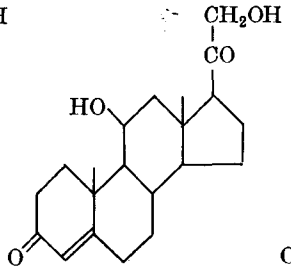


andrenosterone

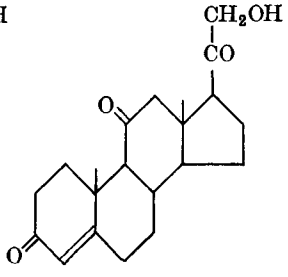
The six active compounds are as follows (they have been designated by letters as well as named systematically).



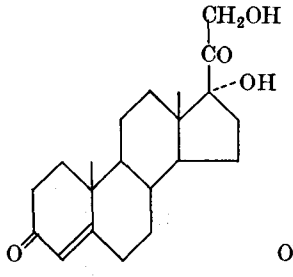
Substance Q;
11-Deoxycorticosterone;
21-Hydroxyprogesterone



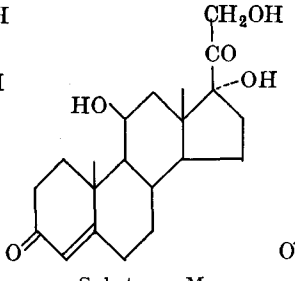
Substance H;
Corticosterone;
11 : 21-Dihydroxy-progesterone



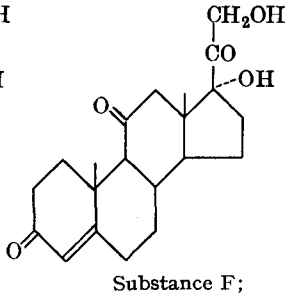
Compound A;
11-Dehydrocorticosterone;
21-Hydroxy-11-keto-progesterone



Substance S;
11-Deoxy-17-hydroxy-corticosterone



Substance M;
17-Hydroxy-corticosterone



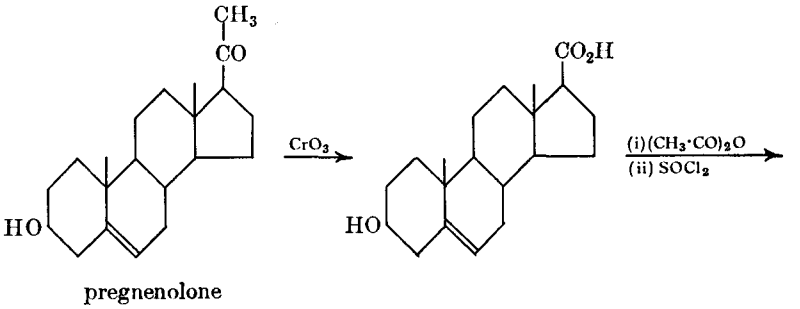
Substance F;
Compound E;
11-Dehydro-17-hydroxy-corticosterone;
cortisone

Owing to the presence of the α-hydroxyketone group, the adrenal cortical hormones are strong reducing agents. The hydroxyl group at position 21 behaves in the usual way, but the 11-keto group does not form an oxime or a phenylhydrazone. The 11-keto group is resistant to catalytic reduction in neutral solution, but can be reduced in acid solution; it is readily reduced to a hydroxyl group by lithium aluminium hydride, and to a methylene group by the Clemmensen reduction.

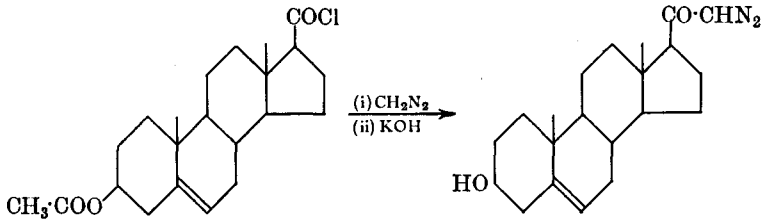
The keto-hormones are separated from non-keto compounds by means of Girard's reagents P and T (see Vol. I).

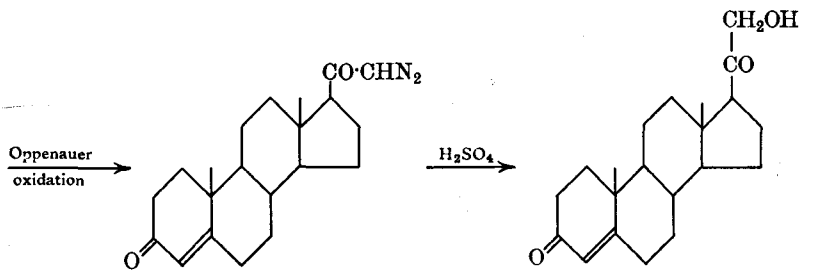
The structures of the cortical hormones have been elucidated by degrada-

tion and by partial syntheses from sterols of known structure, *e.g.*, deoxycorticosterone from stigmasterol (Reichstein *et al.*, 1937, 1940). The first step is the conversion of stigmasterol to pregnenolone (see §19 i).



pregnenolone

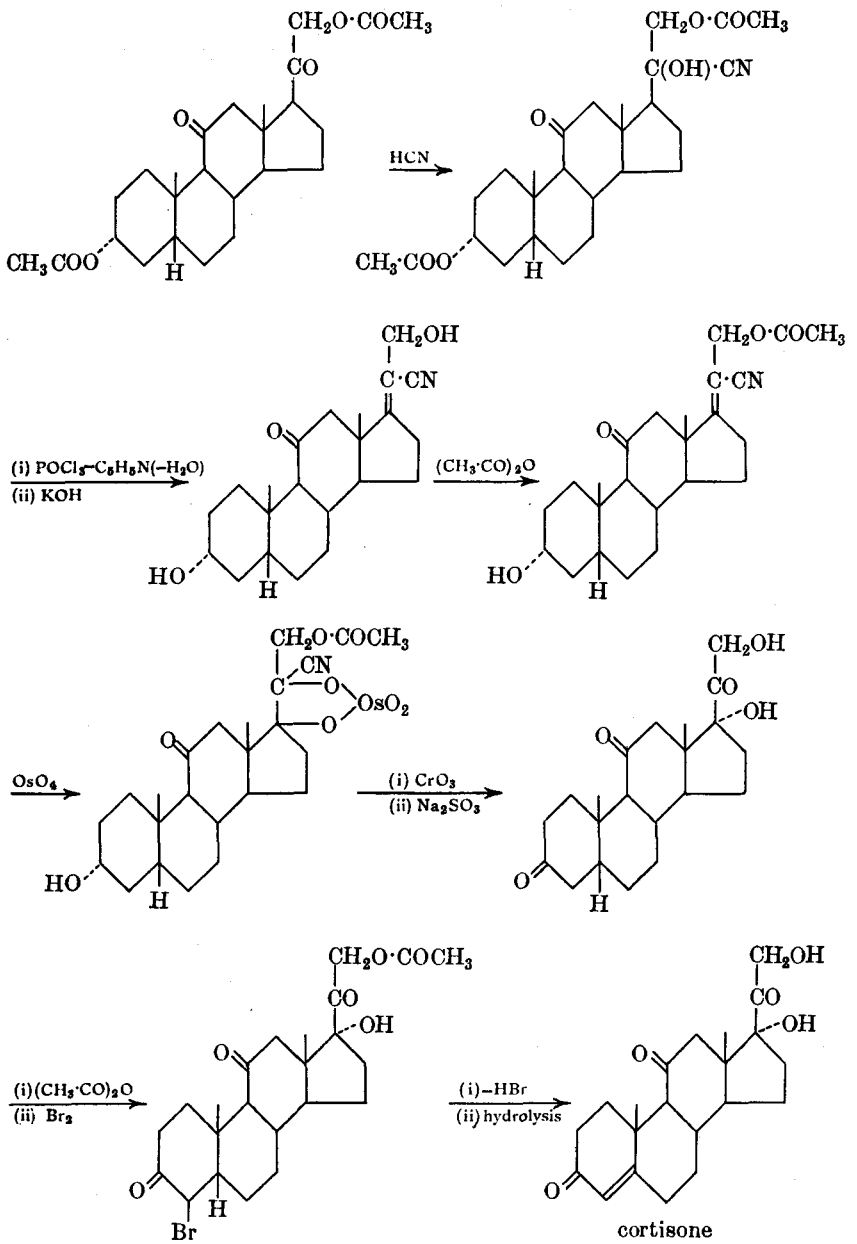




deoxycorticosterone

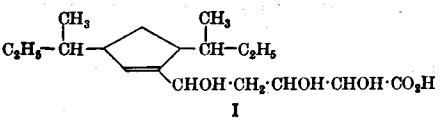
A very interesting point about the above synthesis is the unusual stability of the diazoketone.

Cortisone (Substance F, Compound E) has been used for the treatment of rheumatoid arthritis and rheumatic fever. Many partial syntheses are known, and there is also a total synthesis; *e.g.*, the following partial synthesis starts from 3α : 21-diacetoxypregnane-11 : 20-dione (Sarett, 1948) (see overleaf).

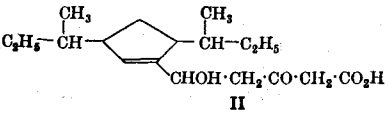


AUXINS

§23. It had been suggested for some years by botanists that various substances had plant growth-promoting properties, but it was not until 1933 that such compounds were actually isolated. In 1933, Kögl *et al.* isolated an active compound from human urine, and they named it **auxin *a*** and showed that its structure is I. Soon afterwards, Kögl *et al.* isolated **auxin *b*** (II) from maize germ oil.

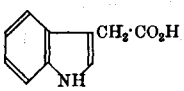


I
auxin *a*



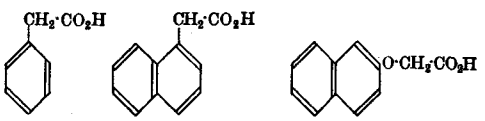
II
auxin *b*

The name **auxin** is now taken as the *generic* name for the plant hormones. Auxins generally occur in the plant kingdom, but are also present in urine, etc. Further work by Kögl *et al.* (1934) led to the isolation from urine of another growth-promoting substance which the authors named " hetero-auxin ", and subsequently showed that this compound is indole-3-acetic acid.



indole-3-acetic acid

The discovery that indole-3-acetic acid had plant growth-promoting properties led to the examination of compounds of related structure, and it was soon found that various derivatives of indole-3-acetic acid are also very active; it was also found that a number of arylacetic acids and aryloxyacetic acids are active, *e.g.*, phenylacetic acid, III, 1-naphthaleneacetic acid, IV, and 2-naphthoxyacetic acid, V.



III IV V

Recent work has suggested that indole-3-acetic acid is the natural plant hormone, and not auxins *a* and *b*. In fact, there now appears to be some doubt as to the existence of auxin *a* (*auxentriolic acid*) and auxin *b* (*auxenolonic acid*); neither of these compounds has been isolated since Kögl obtained them.

The relation between chemical structure and growth-promoting properties has still to be solved, but nevertheless much progress has been made in this direction. Koepli *et al.* (1938) believe that a plant hormone must have a ring structure containing at least one double bond, and a side-chain containing a carboxyl group (or a group capable of being converted into a carboxyl group) removed from the ring by at least one carbon atom (*cf.* compounds I–V). These requirements, however, have been modified by Veldestra (1944–).

READING REFERENCES

Fieser and Fieser, *Steroids*, Reinhold (1959).
Gilman (Ed.), *Advanced Organic Chemistry*, Wiley (1943, 2nd ed.). Ch. 19. The Steroids.
Rodd (Ed.), *Chemistry of Carbon Compounds*, Elsevier. Vol. IIB (1953). Ch. 17. Sterols and Bile Acids. Ch. 18. Sex Hormones; Adrenocortical Hormones.
Stewart and Graham, *Recent Advances in Organic Chemistry*, Longmans, Green. Vol. III (1948, 7th ed.). Ch. I. The Bile Acids and Sterols. Ch. III. The Hormones.
Vitamins and Hormones, Academic Press (Vol. I, 1943–).

Cook (Ed.), *Progress in Organic Chemistry*, Butterworth. Vol. II (1953). Ch. 4. The Partial Synthesis of Cortisone and Related Compounds from Accessible Steroids. Ch. 5. The Relationship of Natural Steroids to Carcinogenic Aromatic Compounds. Vol. III (1955). Ch. 1. Total Synthesis of Steroids. Vol. 5 (1961). Ch. 4. The Chemistry of the Higher Terpenoids.

Shoppee, *Chemistry of the Steroids*, Academic Press (1958).

Klyne, *The Chemistry of the Steroids*, Methuen (1957).

Lythgoe, Some Recent Advances in the Chemistry of the D-Vitamins, *Proc. Chem. Soc.*, **1959**, 141.

Butenandt, The Windaus Memorial Lecture, *Proc. Chem. Soc.*, **1961**, 131.

Loewenthal, Selective Reactions and Modifications of Functional Groups in Steroid Chemistry, *Tetrahedron*, **1959**, **6**, 269.

Handbook for Chemical Society Authors, Chem. Soc. (1960). Ch. 4. Nomenclature of Steroids.

Dodds, Synthetic Œstrogens, *J. Pharm. Pharmacol.*, **1949**, **1**, 137.

Wicker, The Mechanism of Catalytic Hydrogenation of Cyclic Compounds, *J.C.S.*, **1956**, 2165.

Popják, Chemistry, Biochemistry and Isotopic Tracer Technique, Royal Institute of Chemistry Monograph, No. 2 (1955).

Ciba Foundation Symposium on the Biosynthesis of Terpenes and Sterols, Churchill (1959).

Skoog (Ed.), *Plant Growth Substances*, University of Wisconsin (1951).

Pincus and Thimann (Ed.), *The Hormones*, Academic Press. Vol. I (1948). Plant Growth Hormones (p. 5).

Audus, *Plant Growth Substances*, Leonard Hill Ltd. (1953).

CHAPTER XII

HETEROCYCLIC COMPOUNDS CONTAINING TWO OR MORE HETERO-ATOMS

§1. Nomenclature. (i) When the heterocyclic compound contains two or more hetero-atoms, the starting point for numbering is the hetero-atom of as high a group in the periodic table and as low an atomic number in that group. Thus the order of naming will be O, S, Se, N, P, As, Sb, Si, Sn, Pb, Hg.

(ii) With the atom of the preferred kind as number 1, the ring is numbered in such a way that the hetero-atoms are given the lowest numbers possible.

(iii) Of two or more numberings conforming to rules (i) and (ii), the one that is chosen is that which assigns low numbers more nearly in the order of precedence established by rule (i).

(iv) Of two or more numberings conforming to rules (i)–(iii), the one that is chosen is that which gives hydrogen atoms the lowest numbers possible.

(v) When a heterocyclic compound containing at least one nitrogen atom does not end in *ine* and gives *basic* compounds on progressive hydrogenation, the latter derivatives will be indicated by the successive endings *ine, idine*; *e.g.*, pyrazole, pyrazoline, pyrazolidine.

The hetero-atoms in heterocyclic compounds are indicated by prefixes, *e.g.*, **O** by **oxa, S** by **thia, N** by **aza.**

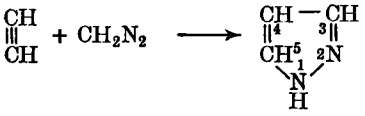
AZOLES

Azole is the suffix used for five-membered rings containing two or more hetero-atoms, at least one of which is nitrogen.

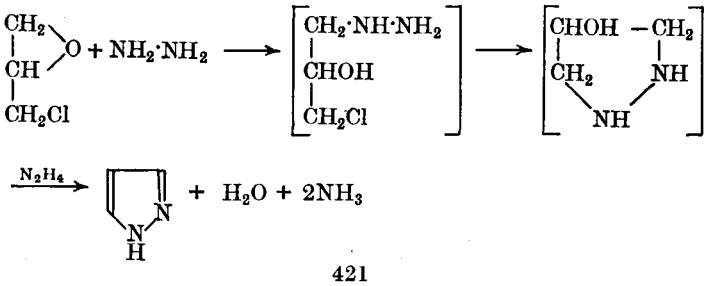
PYRAZOLE GROUP

§2. Pyrazole. Pyrazole may be synthesised in a number of ways, some of the more convenient methods being the following:

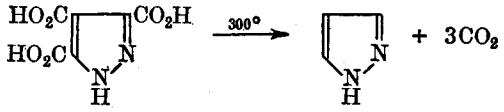
(i) By passing acetylene into a cold ethereal solution of diazomethane (von Pechmann, 1898).



(ii) By heating epichlorohydrin with hydrazine in the presence of zinc chloride (Balbiano, 1890).



(iii) By the decarboxylation of various pyrazolecarboxylic acids, *e.g.*, by heating pyrazole-3 : 4 : 5-tricarboxylic acid (see also §2a ii).



(iv) Jones (1949) has shown that pyrazole may be conveniently prepared by the condensation of 1 : 1 : 3 : 3-tetraethoxypropane,

$$(C_2H_5O)_2CH{\cdot}CH_2{\cdot}CH(OC_2H_5)_2,$$

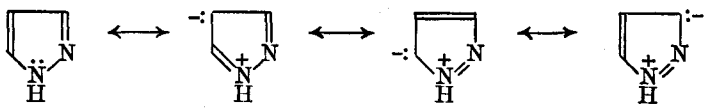
with hydrazine dihydrochloride.

Properties of pyrazole. Pyrazole is a colourless solid, m.p. 70°. It is a tautomeric substance; the existence of tautomerism cannot be demonstrated in pyrazole itself, but it can be inferred by the consideration of pyrazole derivatives. If pyrazole is tautomeric, then the positions 3 and 5 will be identical; if pyrazole is not tautomeric, then these positions are different. Now Knorr *et al.* (1893) showed that on oxidation, both 3-methyl-1-phenylpyrazole and 5-methyl-1-phenylpyrazole gave the *same* product, *viz.*, methylpyrazole. Thus positions 3 and 5 must be equivalent in pyrazole, and this

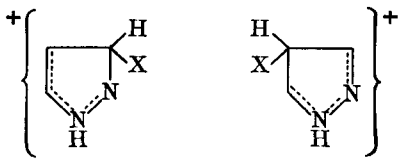
I II

can only be explained by assuming that pyrazole is tautomeric (I and II). It therefore follows that in pyrazole there can only be *two* carbon-alkyl derivatives, 3- (or 5-) and 4-. If, however, the imino hydrogen is replaced by an alkyl or aryl group, then *three* carbon-alkyl derivatives are possible, 3, 4 and 5, since tautomerism is now impossible, and so positions 3 and 5 are no longer equivalent.

Pyrazole exhibits aromatic properties, *e.g.*, it is readily halogenated, nitrated and sulphonated; the group enters at position 4. The following resonating structures are possible for pyrazole.



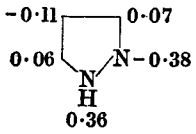
If these structures are contributed equally, then electrophilic attack should occur equally well at positions 3, 4 or 5 (in pyrazole itself, positions 3 and 5 are equivalent). As we have seen above, electrophilic attack occurs exclusively at position 4. The reason for this is not certain. Possibly the resonating structures are not contributed equally (as was assumed). On the other hand, Dewar (1949) has suggested that substitution occurs in the 4-position because the transition state for 4-substitution is more symmetrical,



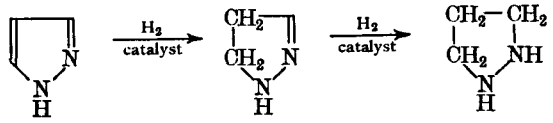
3-substitution 4-substitution

and consequently more stable, than the transition state for 3- (or 5-) sub-

stitution. Orgel *et al.* (1951), however, have calculated the electron distribution in pyrazole, and it can be seen from their results that 4-substitution will be favoured by electrophilic reagents. Brown (1955, 1960) has also calculated the electron densities in pyrazole.



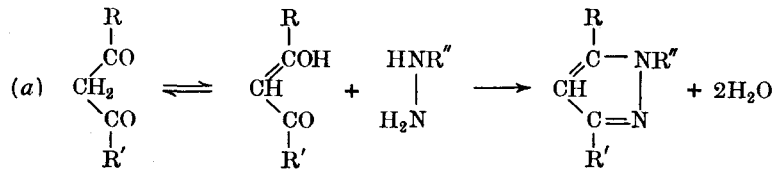
It is interesting to note that pyrazole-4-diazonium salts are stable to boiling water. Pyrazole is feebly basic, and forms salts with inorganic acids; the imino hydrogen may be replaced by an acyl group. Pyrazole is very resistant to oxidising and reducing agents, but may be hydrogenated catalytically, first to pyrazoline, and then to pyrazolidine. Both of these compounds are stronger bases than pyrazole.

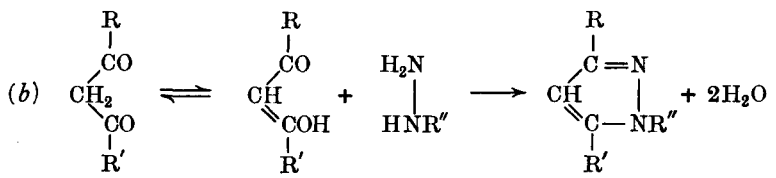


pyrazoline pyrazolidine

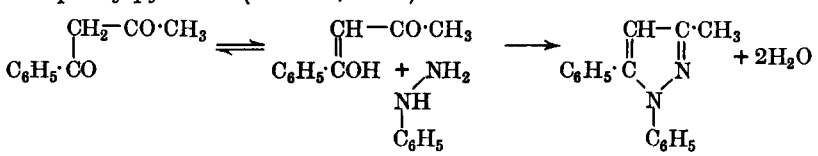
§2a. Synthesis of pyrazole derivatives.

(i) A very important method for preparing pyrazole derivatives is by the reaction between β-diketones (or β-ketoaldehydes) and hydrazines (Knorr *et al.*, 1883).

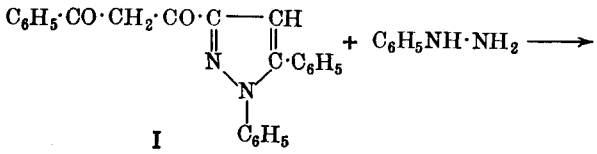




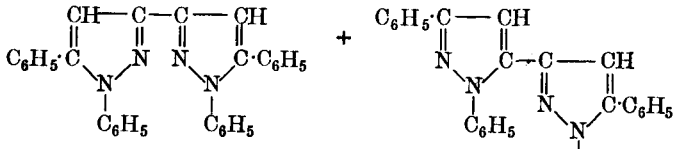
Thus, according to the above, a mixture of isomeric pyrazoles will be produced. Contrary to general opinion, the product is usually only one of the isomers, *e.g.*, benzoylacetone and phenylhydrazine form only 3-methyl-1 : 5-diphenylpyrazole (Drumm, 1931).



In a few cases, two isomers have been isolated, *e.g.*, 3-α-benzoylacetyl-1 : 5-diphenylpyrazole, I, reacts with phenylhydrazine to produce a mixture of 1 : 1′ : 5 : 5′-tetraphenyl-3 : 3′-dipyrazolyl, II, and 1 : 1′ : 3′ : 5-tetraphenyl-3 : 5′-dipyrazolyl, III (Finar, 1955).

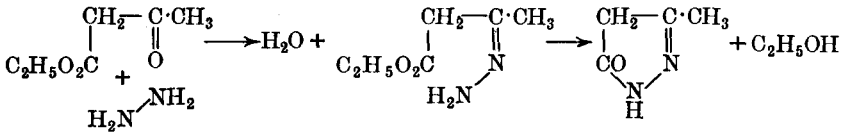


I

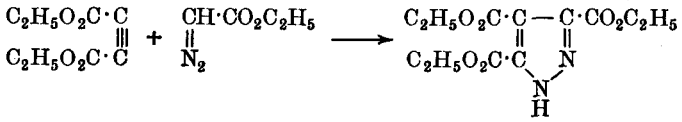


II III

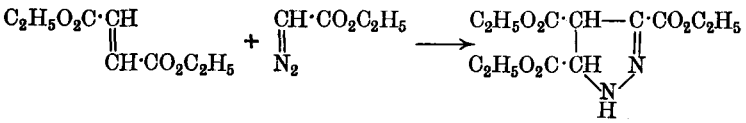
If β-ketoesters are used instead of β-diketones, then 5-pyrazolones are formed (Knorr *et al.*, 1883), *e.g.*, ethyl acetoacetate reacts with hydrazine to form 3-methylpyrazol-5-one.



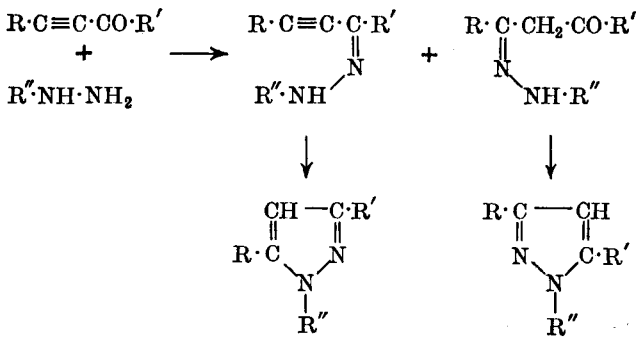
(ii) Pyrazolecarboxylic acids are produced by the reaction between diazoacetic ester and acetylenic compounds, *e.g.*, with ethyl acetylenedicarboxylate, ethyl pyrazole-3 : 4 : 5-tricarboxylate is formed.



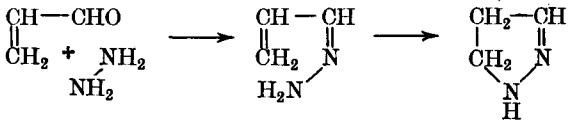
If an ethylenic compound is used instead of an acetylenic one, then a pyrazoline derivative is produced, *e.g.*, ethyl fumarate gives ethyl pyrazoline-3 : 4 : 5-tricarboxylate.



(iii) Pyrazoles are produced by the reaction between acetylenic carbonyl compounds and hydrazines (Moureu *et al.*, 1903); a mixture of isomers is said to be obtained.

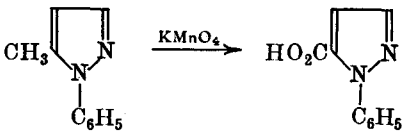


(iv) Pyrazolines are obtained by the condensation of $\alpha:\beta$-unsaturated ketones or aldehydes with hydrazines, *e.g.*, acraldehyde and hydrazine give pyrazoline.

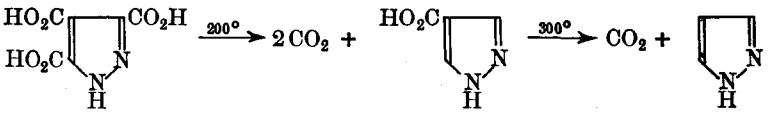


Pyrazolines may be oxidised to pyrazoles by bromine or mercuric oxide.

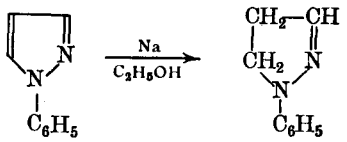
Properties of the pyrazole derivatives. Pyrazoles with substituent methyl groups may be oxidised by potassium permanganate to the corresponding pyrazolecarboxylic acids, *e.g.*,



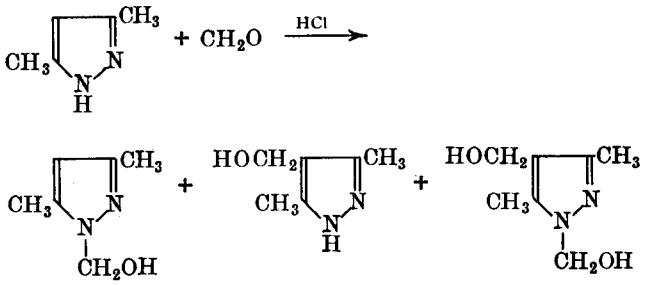
Pyrazole-3- and 5-carboxylic acids are readily decarboxylated by heating above their melting points; the pyrazole-4-carboxylic acids are more stable, but can nevertheless be decarboxylated at elevated temperatures, *e.g.*,



Although pyrazole itself is not reduced by sodium and ethanol, *N*-phenyl substituted pyrazoles are readily reduced to the corresponding pyrazolines, *e.g.*,

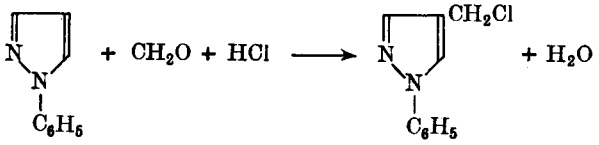


1-Unsubstituted pyrazoles apparently cannot be chloromethylated; carbinols are produced, *e.g.* (Dvoretzky *et al.*, 1950):



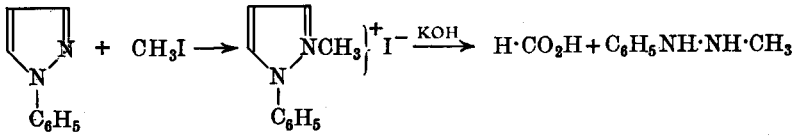
(main product)

On the other hand, 1-phenylpyrazole can readily be chloromethylated in the 4-position (Finar *et al.*, 1954).



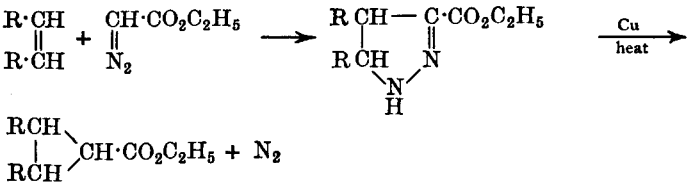
4-Chloromethyl-1-phenylpyrazole can be converted into 1-phenylpyrazole-4-aldehyde by means of the Sommelet reaction (see Vol. I). The 4-aldehyde is more conveniently prepared by the direct formylation of 1-phenylpyrazole with dimethylformamide and phosphoryl chloride (Finar *et al.*, 1957). 1-Phenylpyrazole can also be mercurated in the 4-position (Finar *et al.*, 1954).

When boiled with concentrated aqueous potassium hydroxide, quaternary pyrazoles are converted into hydrazines (Knorr *et al.*, 1906), *e.g.*,

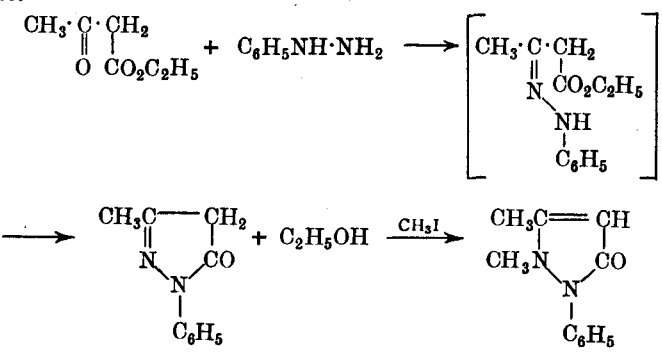


Knorr used this reaction to prepare *sym.*-disubstituted hydrazines; at the same time, this reaction proves the structure of the pyrazole-quaternary salts.

Esters of the pyrazolinecarboxylic acids eliminate nitrogen on heating to give *cyclo*propane derivatives; sometimes much better results are achieved if the compound is heated with copper powder.



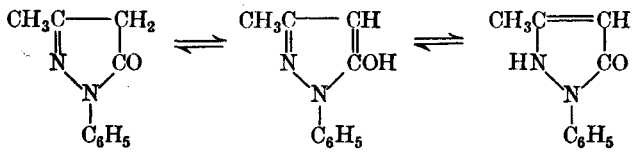
Antipyrine (2 : 3-dimethyl-1-phenylpyrazol-5-one), m.p. 127°, is very much used in medicine as a febrifuge. It is prepared industrially by condensing ethyl acetoacetate with phenylhydrazine, and methylating the product, 3-methyl-1-phenylpyrazole-5-one, with methyl iodide in alkaline ethanolic solution, or with methyl sulphate in the presence of sodium hydroxide.



3-methyl-1-phenylpyrazol-5-one antipyrine

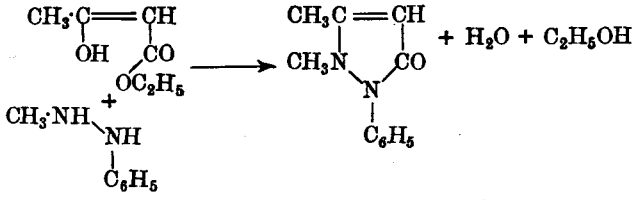
At first sight one might have expected to obtain the *O*-methyl or the 4-methyl derivative, since the tautomeric forms IV (keto) and V (enol) are theoretically

possible. Methylation of 3-methyl-1-phenylpyrazole-5-one with diazomethane results in the formation of the *O*-methyl derivative (this is also



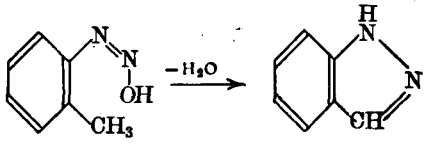
IV V VI

produced in a small amount when methyl iodide is used as the methylating reagent). This raised some doubts as to the structure of antipyrine, since for its formation, the tautomeric form VI must also be postulated. The structure of antipyrine was shown to be that given above by its synthesis from *sym.*-methylphenylhydrazine and ethyl acetoacetate.

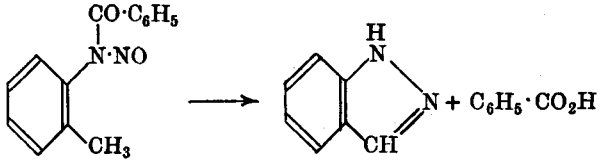


The pyrazole nucleus has always been considered to be a synthetic one, but Fowden *et al.* (1959) have now isolated α-amino-β-1-pyrazolylpropionic acid from water-melon seed; this acid has been synthesised in good yield by Finar *et al.* (1960).

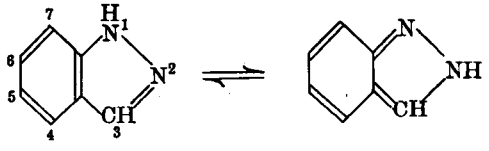
§2b. Indazoles (benzopyrazoles). Indazole may be prepared by the removal of a molecule of water from *o*-toluenediazohydroxide in neutral solution (the yield is very poor).



Indazole may conveniently be prepared by heating *o*-*N*-nitroso-*N*-benzoyltoluidine in benzene solution.



Indazole, m.p. 146°, exhibits the same type of tautomerism that exists in pyrazole, since *two* series of *N*-derivatives (1 and 2) are known:



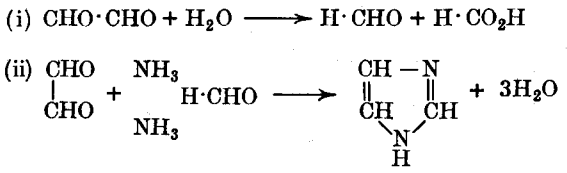
Nitration and sulphonation of indazole produce the 5-substitution product; bromination gives the 3 : 5-dibromo compound.

IMIDAZOLE GROUP

This group of compounds is also known as the *iminazoles* or the *glyoxalines*.

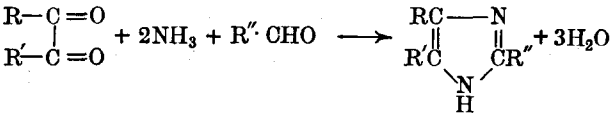
§3. Imidazole (iminazole, glyoxaline) is isomeric with pyrazole, and occurs in the purine nucleus and in the amino-acid histidine; 4-amino-imidazole-5-carboxamide occurs naturally as a riboside (or ribotide).

Imidazole may be prepared by the action of ammonia on glyoxal. The mechanism of this reaction is uncertain, but one suggestion is that one molecule of glyoxal breaks down into formic acid and formaldehyde, and then the latter reacts as follows:

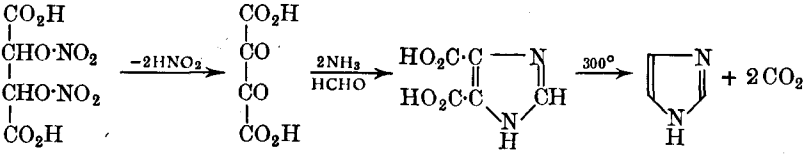


A certain amount of support for this mechanism is given by the fact that glyoxaline may be prepared directly from glyoxal, ammonia and formaldehyde.

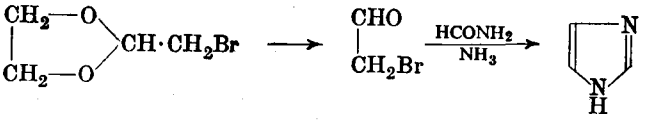
A general method for preparing imidazoles is by the reaction between an α-dicarbonyl compound, ammonia and an aldehyde (Radziszewsky, 1882).



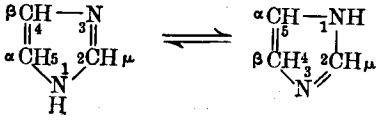
Imidazole itself is best prepared by the action of ammonia on a mixture of formaldehyde and tartaric acid dinitrate (" dinitrotartaric acid "), and then heating the dicarboxylic acid thereby produced.



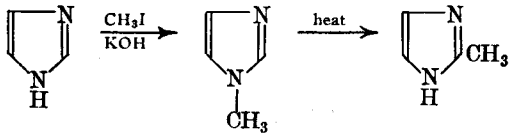
Another good method is to brominate paraldehyde in ethylene glycol and to heat the product, 2-bromomethyl-1 : 3-dioxalan, with formamide in the presence of ammonia (Bredereck *et al.*, 1958); bromoacetaldehyde is probably an intermediate:



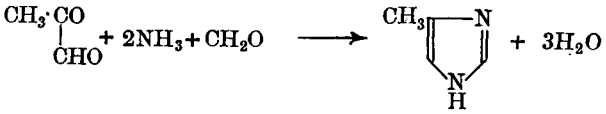
Imidazole, m.p. 90°, is a weak base, but it is more basic than pyrazole. Imidazole is a tautomeric substance, since positions 4 and 5 are equivalent (positions 5, 4 and 2 have also been designated α, β and μ, respectively).



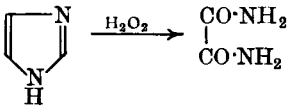
Methyl iodide attacks imidazole in potassium hydroxide solution to form 1-methylimidazole which, when strongly heated, isomerises to 2-methylimidazole (*cf.* the Hofmann rearrangement; see Vol. I).



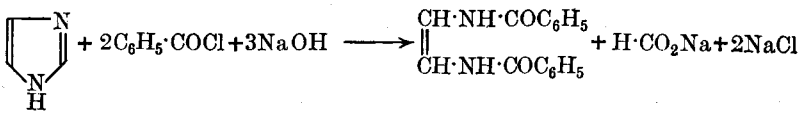
An interesting method of preparing 4(5)-methylimidazole is by the action of zinc hydroxide and ammonia on glucose; the reaction is assumed to occur *via* the breakdown of glucose into methylglyoxal and formaldehyde, which then react as follows:



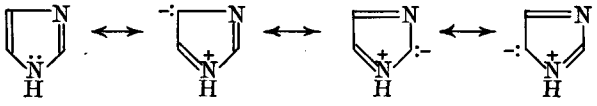
The imidazole ring is extremely stable towards oxidising and reducing agents; hydrogen peroxide, however, readily opens the ring to form oxamide.



Acetyl chloride and acetic anhydride have no action on imidazole, but benzoyl chloride in the presence of sodium hydroxide *opens* the ring to form dibenzoyldiaminoethylene.

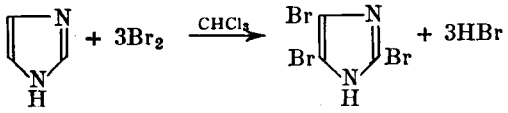


Nitration and sulphonation of imidazole produce the 4(5)-derivative. Electrophilic attack at positions 4 or 5 can be accounted for by the contributions of the resonating structures II and IV. Resonating structure III



I II III IV

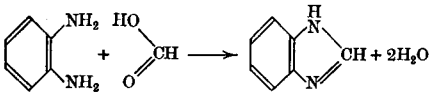
shows that position 2 should also be subject to electrophilic attack. This is found to be the case with halogenation, *e.g.*, bromine reacts with imidazole in chloroform solution to give 2 : 4 : 5-tribromoimidazole.



Imidazole couples with diazonium salts in the 2-position, but *N*-alkylimidazoles do not couple at all.

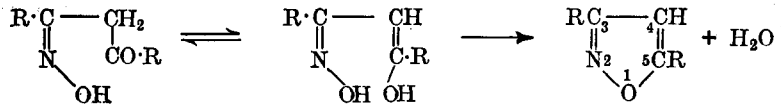
§3a. Benzimidazoles (benziminazoles). These are readily formed by heating *o*-phenylenediamines with carboxylic acids, *e.g.*, benzimidazole itself

(m.p. 170°) is produced by heating *o*-phenylenediamine with 90 per cent. formic acid.

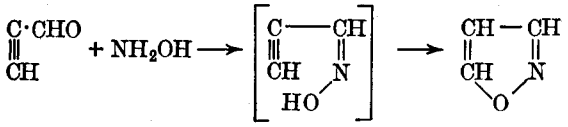


OXAZOLE GROUP

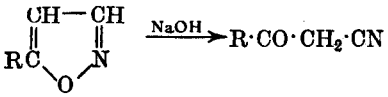
§4. *iso*-Oxazoles. *iso*-Oxazoles are formed by the dehydration of the monoximes of β-diketones or β-ketoaldehydes.



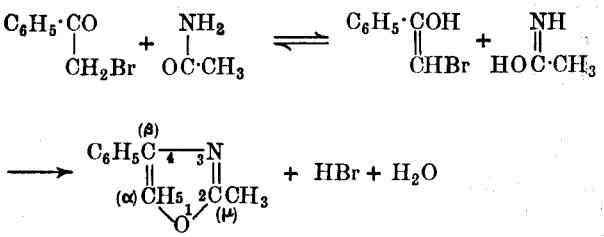
iso-Oxazole itself may be prepared by the action of hydroxylamine on propargylaldehyde.



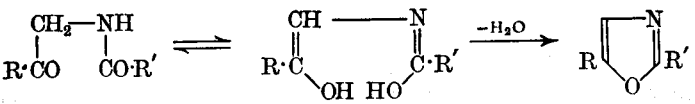
iso-Oxazole is a colourless liquid, b.p. 96°, and smells like pyridine; it is weakly basic. *iso*-Oxazoles, when substituted in the 3 : 5-positions, are stable to alkalis, but when the 3-position is vacant, the ring is opened to form ketonitriles (*cf.* oximes, §§2f, 2g. VI).



§4a. Oxazoles. Oxazoles may be prepared by the condensation of acid amides with α-halogenoketones, *e.g.*, acetamide and ω-bromoacetophenone form 2-methyl-4-phenyloxazole; the mechanism of the reaction is not certain but it may occur through the enol forms.



A better method of preparation is the dehydration of α-acylamidocarbonyl compounds with sulphuric acid or phosphorus pentachloride.

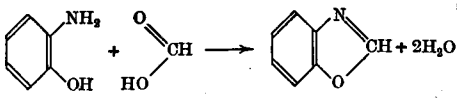


Oxazoles have basic properties similar to those of pyridine, but are less resistant to oxidation. They possess aromatic properties, and the stability of the ring towards hydrolytic reagents depends on the nature of the sub-

stituents in the ring (*cf. iso*-oxazoles). The parent compound, oxazole, has not yet been prepared.

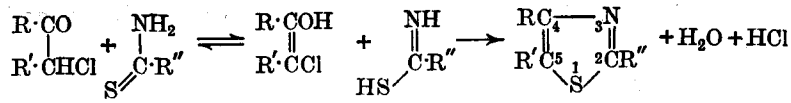
5-Oxazolones. The oxazolones are keto derivatives of the oxazolines, the most important group being the 5-oxazolones or **azlactones.** These azlactones are very important intermediates in the preparation of α-amino-acids (see §2 *va.* XIII) and keto-acids (see Vol. I).

§4b. Benzoxazoles. These may be prepared by the reaction between *o*-aminophenols and carboxylic acids, *e.g.*, *o*-aminophenol and formic acid form benzoxazole, m.p. 31°.

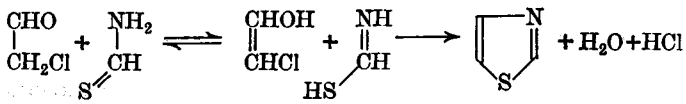


THIAZOLE GROUP

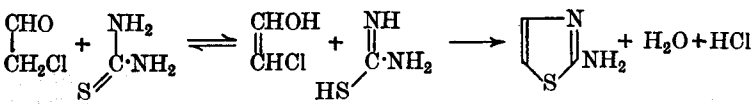
§5. Thiazoles. A general method for preparing thiazoles is the condensation between α-halogenocarbonyl compounds (particularly the chloro derivatives) and thioamides; the mechanism of the reaction is uncertain, but it may occur through the enol forms.

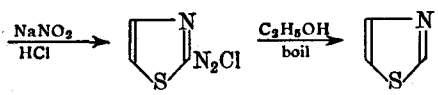


Thiazole itself may be prepared from chloroacetaldehyde and thioformamide.

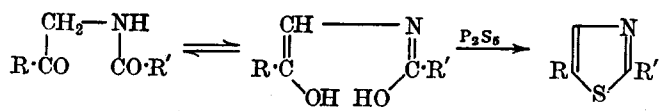


If thiourea or its substitution products are used instead of thioamides, then 2-aminothiazoles are produced, *e.g.*, thiazole may be prepared from chloroacetaldehyde and thiourea as follows:

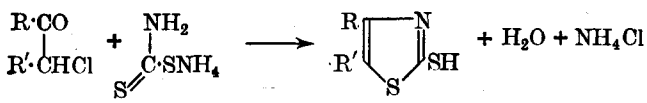




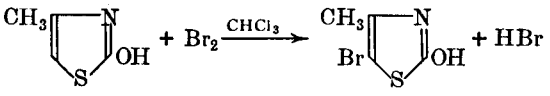
Another general method for preparing thiazoles is by the action of phosphorus pentasulphide on α-acylamidocarbonyl compounds.



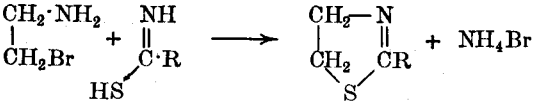
2-Mercaptothiazoles may be prepared by the condensation between α-chloroketones and ammonium dithiocarbamate.



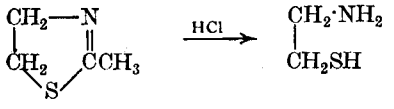
Thiazole is a weakly basic liquid, b.p. 117°; it occurs in vitamin B_1. It is a very stable compound, and is not affected by the usual reducing agents; sodium and ethanol, however, open the ring to form thiols (or hydrogen sulphide) and amines. Thiazole is very resistant to substitution reactions, but if a hydroxyl group or an amino group is in position 2, then the molecule is readily attacked by the usual electrophilic reagents to form 5-substitution products, *e.g.*, 2-hydroxy-4-methylthiazole is readily brominated in chloroform solution to give 5-bromo-2-hydroxy-4-methylthiazole.



§5a. Thiazolines. These may be prepared by the reaction between β-halogenoamines and thioamides, *e.g.*,

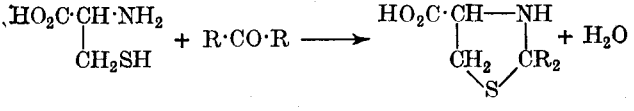


A characteristic reaction of the thiazoles is their ring opening by the action of acids, *e.g.*,



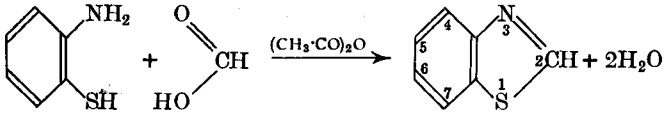
2-methylthiazoline 2-aminoethanethiol

§5b. Thiazolidines. These are readily formed by the condensation of carbonyl compounds with cysteine.

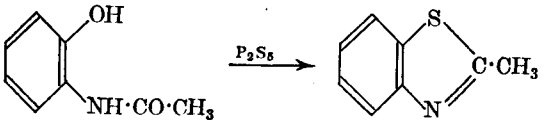


The thiazolidine ring is very easily opened, sometimes by boiling with water, or with an aqueous solution of mercuric chloride (see also penicillin, §6a. XVIII).

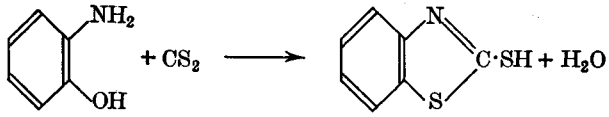
§5c. Benzothiazoles. These may be prepared by the action of acid anhydrides or chlorides on *o*-aminothiophenols, *e.g.*, benzothiazole from *o*-aminothiophenol and formic acid in the presence of acetic anhydride.



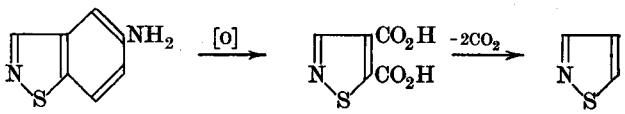
Benzothiazoles are also formed by the action of phosphorus pentasulphide on *o*-acylamidophenols, *e.g.*,



2-Mercaptobenzothiazole is a vulcanisation accelerator (§33a. VIII); it may be prepared as follows:



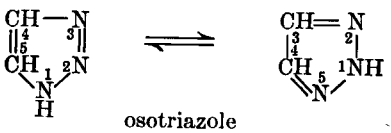
§5d. *iso*Thiazoles. Benz*iso*thiazoles have been known for many years, but no derivatives of *iso*thiazole itself have been obtained until very recently when Adams *et al.* (1956) prepared the parent compound and a number of its simple derivatives, *e.g.*,



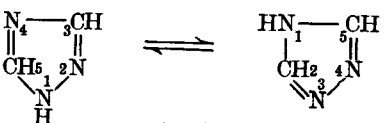
*iso*Thiazole is a colourless liquid which smells like pyridine.

TRIAZOLE GROUP

§6. Osotriazoles and triazoles. Triazoles are five-membered rings which contain two carbon and three nitrogen atoms. Two structural isomeric triazoles are known, the 1 : 2 : 3-(1 : 2 : 5-) and the 1 : 2 : 4- (1 ; 3 : 4-), the former being known as *osotriazole*, and the latter as *triazole*. Each exists in two dissimilar tautomeric forms.



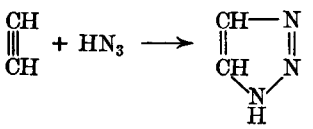
osotriazole



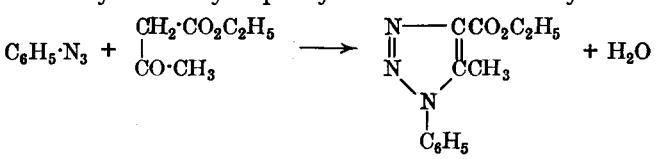
triazole

Replacement of the imino hydrogen atom by an alkyl or aryl group prevents tautomerism, and thereby gives rise to the possibility of *two* *N*-substituted triazoles and *two* *N*-substituted osotriazoles. All four types of compounds have been prepared.

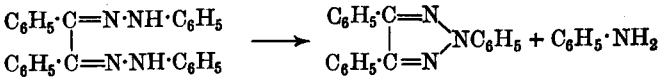
Osotriazole may be prepared by the reaction between acetylene and hydrazoic acid.



On the other hand, a general method for preparing osotriazoles is the condensation of azides with β-ketoesters, *e.g.*, phenyl azide and ethyl acetoacetate form ethyl 5-methyl-1-phenylosotriazole-4-carboxylate.

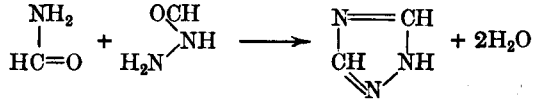


Derivatives of osotriazole may also be prepared by the oxidation of osazones with dichromate and sulphuric acid, or with dilute copper sulphate solution, *e.g.*, benzilosazone gives 1 : 3 : 4-triphenylosotriazole.

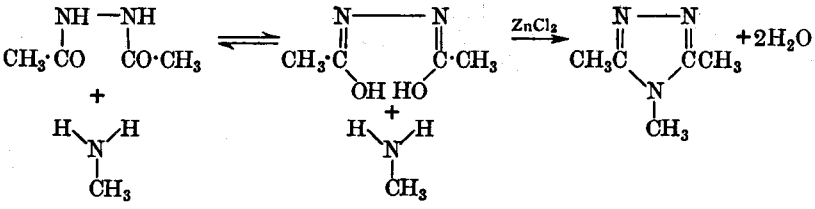


The formation of osotriazoles from sugar osazones provides a good derivative for the characterisation of sugars (see Vol. I).

Triazoles may be prepared by heating acid hydrazides with amides, *e.g.*, formyl hydrazide and formamide give triazole.

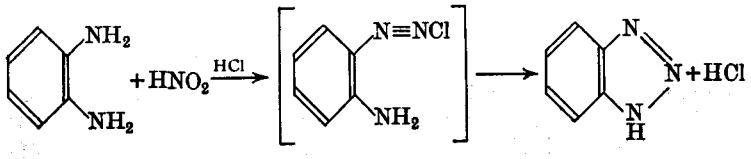


Triazoles are also formed when *sym.*-diacylhydrazines are heated with ammonia or amines in the presence of zinc chloride, *e.g.*, *sym.*-diacetylhydrazine and methylamine give 1 : 2 : 5-trimethyltriazole.

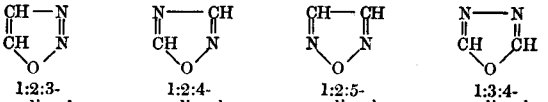


Both triazoles are weak bases, and are very stable compounds.

Benzotriazole is formed by the action of nitrous acid on *o*-phenylenediamine.

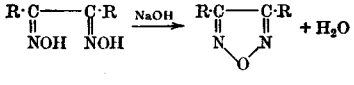


§7. Oxadiazoles. These are five-membered rings containing two carbon and two nitrogen atoms and one oxygen atom; four types are known.



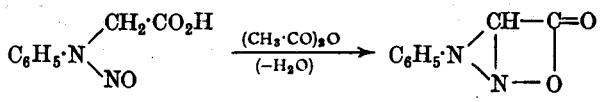
1:2:3-oxadiazole | 1:2:4-oxadiazole | 1:2:5-oxadiazole | 1:3:4-oxadiazole

The **furazans** (1 : 2 : 5-oxadiazoles) may be prepared by the action of sodium hydroxide on the dioximes of α-diketones.



§8. Sydnones. The sydnones were first prepared by Earl *et al.* (1935) by the action of cold acetic anhydride on *N*-nitroso-*N*-phenylglycines;

Earl formulated the reaction as follows:



Earl (1946) proposed the name *sydnone* for compounds of this type; thus the above compound is *N*-phenylsydnone.

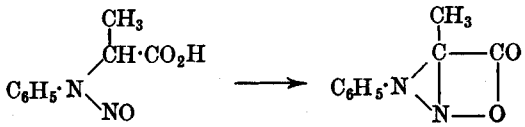
Sydnones are white or pale yellow crystalline compounds, which are hydrolysed by hot 5 per cent. sodium hydroxide to the original *N*-nitroso-*N*-arylglycine, and by moderately concentrated hydrochloric acid to an arylhydrazine, formic acid and carbon dioxide.

The structure proposed by Earl is similar to that of a β-lactone, but Baker *et al.* (1946, 1949) offered a number of objections to this structure, *e.g.*,

(i) A system containing fused three- and four-membered rings would be highly strained, and consequently is unlikely to be produced by dehydration with acetic anhydride; β-lactones are not produced under these conditions.

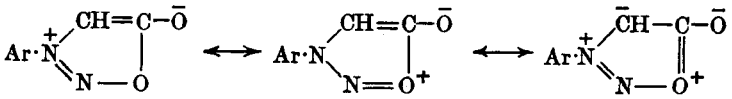
(ii) Many β-lactones are unstable to heat; sydnones are stable and so the β-lactone structure is unlikely.

(iii) If the β-lactone structure is correct, then sydnones should be capable of existing in optically active forms. Kenner and Baker (1946) prepared (+)-*N*-nitroso-*N*-phenylalanine, and when this was converted into a sydnone, the product was optically inactive. If Earl's structure were correct, then the sydnone would be expected to be optically active.



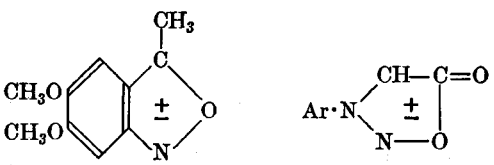
(iv) The aryl nucleus in sydnones is very resistant to substitution by electrophilic reagents. Since the above structure is similar to that of an arylhydrazine, this resistance is unexpected.

Baker *et al.* (1946) therefore proposed a five-membered ring which cannot be represented by any one purely covalent structure; they put forward a number of charged structures, the sydnone being a resonance hybrid, *e.g.*, three charged resonating structures are:



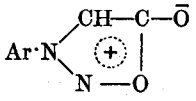
I II III

Now Simpson (1945) had proposed structure IV for 3-methyl-5 : 6-dimethoxyanthranil; Baker *et al.* (1949) adopted this ± sign and suggested that sydnones be represented by structure V. Baker also proposed the



IV V

term **meso-ionic** to describe the sydnone structure. Baker *et al.* (1955) have, however, revised the definition of the term meso-ionic, and have proposed formula V*a* instead of V. This is based on the fact that sydnones are aromatic in character, and the circle and plus sign represent the sextet

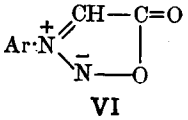


V*a*

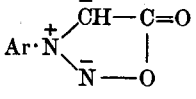
of π-electrons in association with a positive charge (the " aromatic sextet " is the essential feature of aromatic compounds).

Dipole moment measurements of various sydnones have shown that the positive end of the dipole is situated on the nitrogen atom attached to the aryl group (Sutton *et al.*, 1947, 1949; Le Fèvre *et al.*, 1947). This is in keeping with Baker's structure.

The meso-ionic structure would necessitate a planar, or almost planar molecule; such a molecule would not be optically active (*cf.* iii above). Earl (1953) has suggested that, from the available evidence, sydnones can be represented as a resonance hybrid, the two main contributing structures being VI and VII.



VI

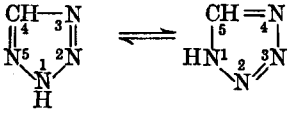


VII

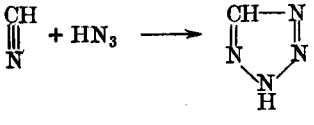
Sutton *et al.* (1949), however, have shown that VI probably contributes to the resonance hybrid, but to a lesser extent than I, II and III.

TETRAZOLE GROUP

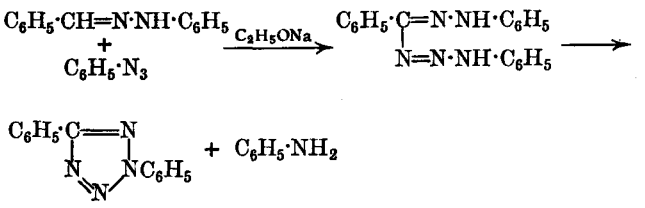
§9. Tetrazole. Tetrazole is a five-membered ring which contains one carbon and four nitrogen atoms. There are two tautomeric forms of tetrazole, and replacement of the imino hydrogen by, *e.g.*, an alkyl group gives rise to *two* *N*-alkyltetrazoles (*cf.* triazoles, §6).



Tetrazole may be prepared by heating hydrogen cyanide with hydrazoic acid in benzene solution at 100°.



Derivatives of tetrazole may be prepared by the condensation of phenyl azide with phenylhydrazones of aldehydes in the presence of ethanolic sodium ethoxide, *e.g.*, benzaldehyde phenylhydrazone and phenyl azide form 1 : 4-diphenyltetrazole.



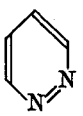
Tetrazole is a colourless solid, m.p. 156°; it has no basic properties, but the imino hydrogen is acidic, *e.g.*, tetrazole forms a silver salt $[CHN_4]^-Ag^+$.

AZINES

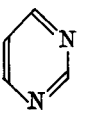
The suffix *azine* is used for six-membered rings which contain two or more hetero-atoms, at least one of which is nitrogen.

DIAZINE GROUP

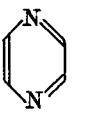
§10. Introduction. The diazines are six-membered rings containing two nitrogen atoms. Three isomeric diazines are theoretically possible, and all three are known.



o-diazine;
pyridazine



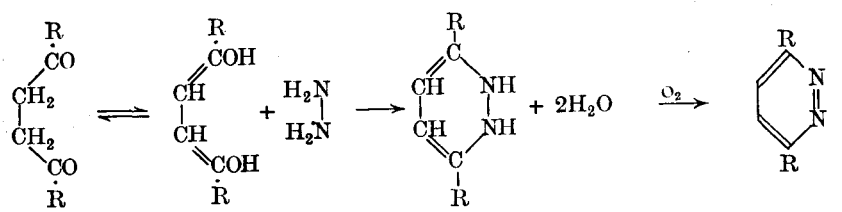
m-diazine;
miazine;
pyrimidine



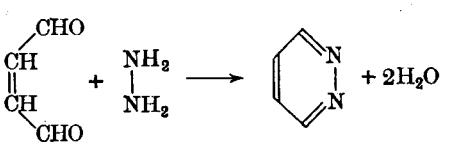
p-diazine;
piazine;
pyrazine

The above formulæ are now usually written with a nitrogen atom at the top, *i.e.*, the formulæ of pyridazine and pyrimidine are inverted.

§11. Pyridazines. These may be prepared by the action of hydrazine on 1 : 4-diketones, the intermediate dihydro compound being readily oxidised by atmospheric oxygen.



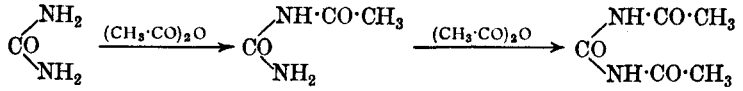
Pyridazine itself may be prepared from maleic dialdehyde and hydrazine hydrate.



Pyridazine is a colourless liquid, b.p. 208°.

PYRIMIDINES

§12. Ureides. Ureides are acylureas, and may be prepared by the action of an acid anhydride or acid chloride on urea, *e.g.*,



acetylurea diacetylurea

The simple ureides resemble the amides in properties.

Allophanic acid, $NH_2{\cdot}CO{\cdot}NH{\cdot}CO_2H$, is not known in the free state, but many of its esters have been prepared:

(i) By the action of chloroformates on urea.

$$NH_2{\cdot}CO{\cdot}NH_2 + Cl{\cdot}CO_2R \longrightarrow NH_2{\cdot}CO{\cdot}NH{\cdot}CO_2R + HCl$$

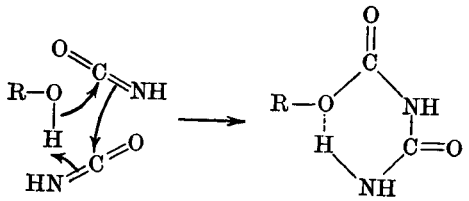
(ii) By the reaction between urethans and cyanic acid.

$$HNCO + NH_2{\cdot}CO_2R \longrightarrow NH_2{\cdot}CO{\cdot}NH{\cdot}CO_2R$$

The alkyl allophanates are well-defined crystalline compounds, and so are frequently used to identify alcohols. They are prepared by passing cyanic acid vapour into the dry alcohol; urethans are intermediate products.

$$ROH + HNCO \longrightarrow NH_2{\cdot}CO_2R \xrightarrow{HNCO} NH_2{\cdot}CO{\cdot}NH{\cdot}CO_2R$$

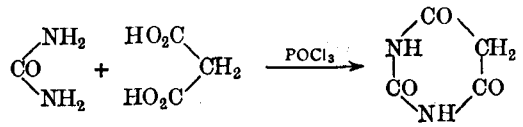
According to Close *et al.* (1953), allophanate formation occurs *via* a concerted attack of two molecules of cyanic acid to form a chelate intermediate.



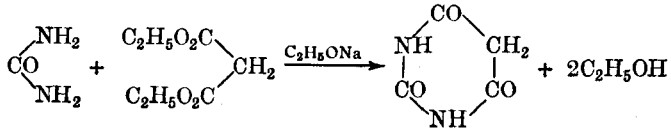
§13. Cyclic ureides. Many cyclic ureides are known; some occur naturally and others are synthetic (a number of cyclic ureides—alloxan, allantoin, parabanic acid and hydantoin—are discussed in §2. XVI, in connection with the purines, which are cyclic diureides).

The cyclic ureides containing a *six-membered ring* behave, in a number of ways, as pyrimidine derivatives.

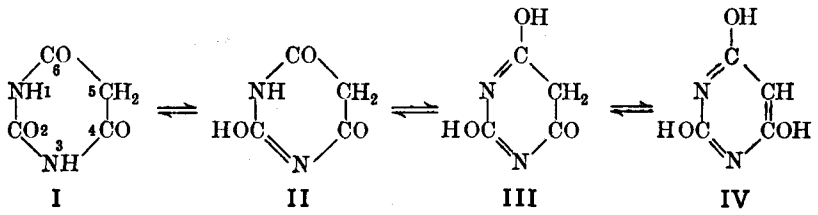
§13a. Barbituric acid. A very important pyrimidine derivative is barbituric acid (malonylurea). It was originally prepared by condensing urea with malonic acid in the presence of phosphoryl chloride (Grimaux, 1879).



A much better synthesis is to reflux ethyl malonate with urea in ethanolic solution in the presence of sodium ethoxide.



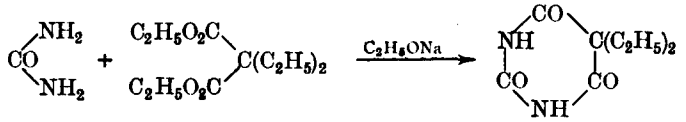
Barbituric acid is a solid, m.p. 253°, and is not very soluble in water. It is strongly acidic due to enolisation (lactam-lactim tautomerism); some possible lactim forms are II–IV. Structure IV represents barbituric acid



I II III IV

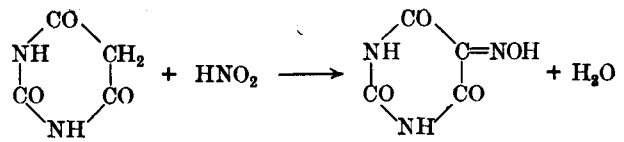
as 2 : 4 : 6-trihydroxypyrimidine, and this structure has been proposed because of the acidic nature of barbituric acid. On the other hand, barbituric acid contains an active methylene group, since it readily forms an oximino derivative with nitrous acid. Thus barbituric acid behaves as if it had structure I, II or III. Furthermore, it is very difficult to acylate hydroxypyrimidines containing hydroxyl groups in the 2-, 4- or 6-positions, thus indicating that structure I is more probable than II or III. This is supported by the fact that methylation of hydroxypyrimidines with, *e.g.*, methyl iodide in the presence of sodium hydroxide, results in the formation of *N*-methyl derivatives; this indicates the probable presence of imino groups. On the other hand, it is possible to replace three hydroxyl groups by three chlorine atoms by means of phosphoryl chloride; this suggests barbituric acid behaves as IV. Barbituric acid also forms *O*-alkyl derivatives, thereby indicating structures II, III and IV.

Barbituric acid can be nitrated and brominated in the 5-position, and also forms metallic derivatives (at position 5). By means of the sodio derivative, one or two alkyl groups may be introduced at position 5 (this reaction is characteristic of the —CH_2·CO— group). Barbituric acid and 5 : 5-dimethylbarbituric acid have no hypnotic action. On the other hand, 5 : 5-diethylbarbituric acid (*Barbitone, Veronal*) has a strong hypnotic action; it is best prepared as follows:



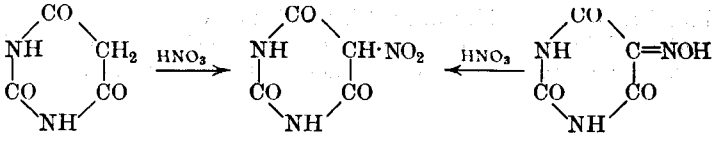
5-*cyclo*Hexyl-3 : 5-dimethylbarbituric acid (*Evipan*) is a better hypnotic than *Barbitone* and is not so toxic. 5-Ethyl-5-phenylbarbituric acid (*Luminal*) is also used in medicine.

§13b. Derivatives of barbituric acid. Violuric acid (5-oximinobarbituric acid) is formed when barbituric acid is treated with nitrous acid; it is the oxime of alloxan (see §2. XVI). Violuric acid gives a violet colour



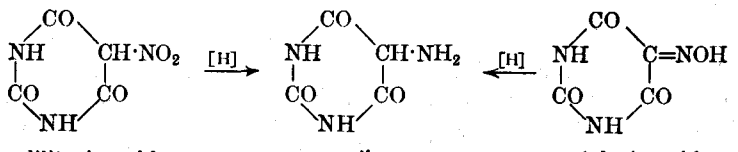
in water, and forms deeply coloured salts with various metals, *e.g.*, the potassium salt is blue and the magnesium and barium salts are purple.

Dilituric acid (5-nitrobarbituric acid) may be prepared by nitrating barbituric acid with fuming nitric acid, or by the oxidation of violuric acid with nitric acid.



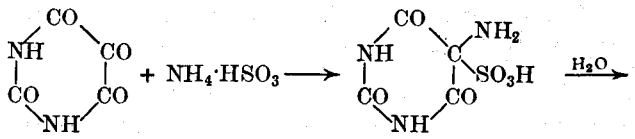
barbituric acid dilituric acid violuric acid

Uramil (5-aminobarbituric acid) is formed by the reduction of either dilituric acid or violuric acid.

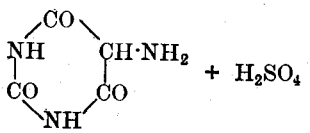


dilituric acid uramil violuric acid

Uramil may also be prepared by the action of ammonium hydrogen sulphite on alloxan, and then boiling the product, **thionuric acid,** with water.

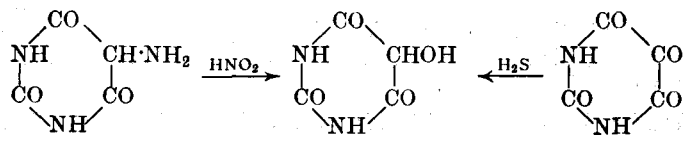


alloxan thionuric acid



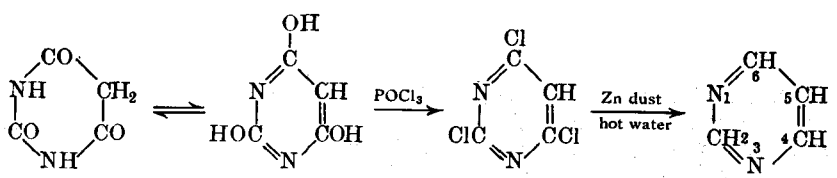
uramil

Dialuric acid (5-hydroxybarbituric acid) is produced by the action of nitrous acid on uramil; it is also formed when alloxan is reduced with hydrogen sulphide or with zinc and hydrochloric acid.



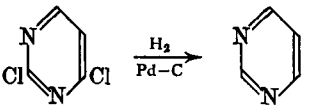
uramil dialuric acid alloxan

§14. **Pyrimidine,** m.p. 22·5°, b.p. 124°/758 mm., was first prepared from barbituric acid as follows (Gabriel, 1900).

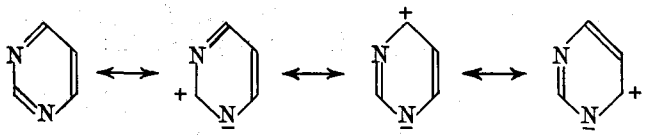


pyrimidine

Pyrimidine may also be prepared by the oxidation of alkylpyrimidines, followed by decarboxylation. A recent preparation is the catalytic reductive dechlorination of 2 : 4-dichloropyrimidine; the latter is heated with hydrogen under pressure in the presence of Pd—C and magnesium oxide (Whittaker, 1953).



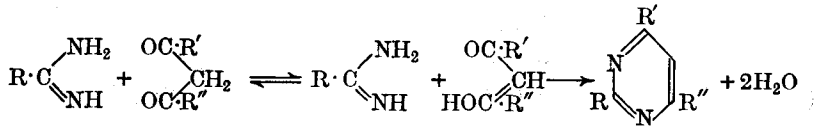
Pyrimidine is neutral in solution, but forms salts with acids. Pyrimidine is probably a resonance hybrid of the following resonating structures:



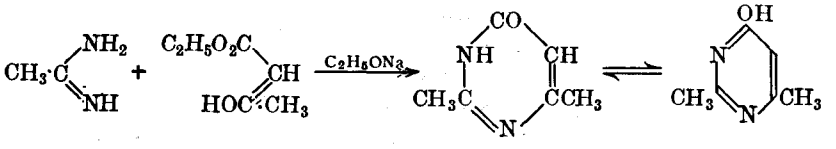
Thus the ring is deactivated, and position 5 has the greatest electron density (*cf.* nitrobenzene and pyridine, Vol. I). It can therefore be expected that attack by electrophilic reagents will be difficult, but attack by nucleophilic reagents (at positions 2, 4 and 6) will be facilitated. Chlorine atoms at 2, 4 or 6 are readily replaced by hydroxyl or amino groups, and an amino group in position 2 or 6 is readily replaced by a hydroxyl group merely on boiling with water (*cf.* vitamin B_1, §3. XVII).

When a hydroxyl or an amino group is present in the pyrimidine nucleus, the compound no longer behaves entirely as an aromatic derivative. The introduction of hydroxyl or amino groups into positions 2, 4 and 6 progressively diminishes the aromatic properties of the compound (*cf.* barbituric acid, §13a, and uracil, §15).

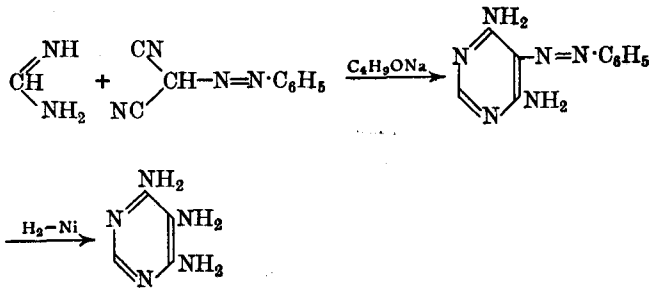
Pyrimidine derivatives. A very important general method for preparing pyrimidines is the condensation between β-carbonyl compounds of the type R·CO·CH_2·CO·R′, where R and R′ = H, R, OR, CN, and compounds having the amidine structure R·C(= NH)·NH_2, where R = R (an amidine), OH (urea), SH or SR (thiourea or its *S*-derivative), NH_2 (guanidine); the condensation is carried out in the presence of sodium hydroxide or sodium ethoxide. Thus:



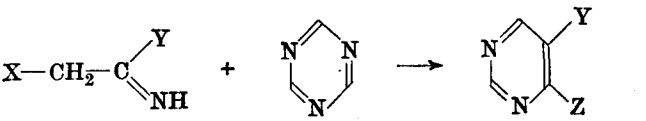
This general reaction may be illustrated by the condensation of acetamidine (R = CH_3) with ethyl acetoacetate (R′ = OC_2H_5, and R″ = CH_3) to form 6-hydroxy-2 : 4-dimethylpyrimidine.



4 : 5-Diaminopyrimidines, which are intermediates in purine synthesis (see §4. XVI), may be prepared by condensing formamidine with phenylazomalononitrile (Todd *et al.*, 1943).



Schaeffer *et al.* (1962) have shown that *s*-triazine reacts with amidines, amidine salts and imidates having α-acidic methylene groups to produce 4 : 5-disubstituted pyrimidines (yield: 51–100 per cent.):

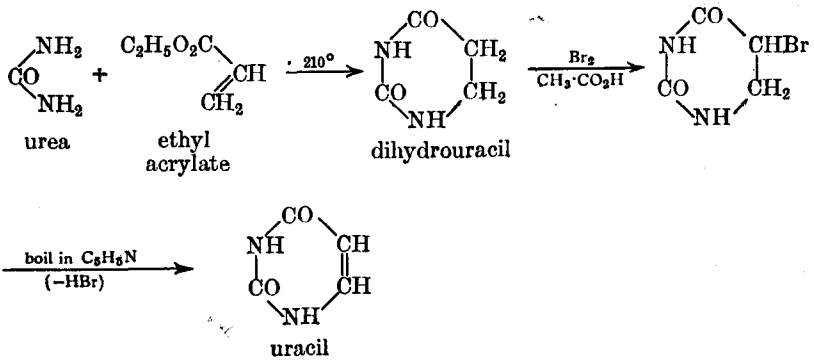


X = CO_2R, $CONH_2$, CN, COPh
Y = NH_2, OR, SR

Z = Y or NH_2

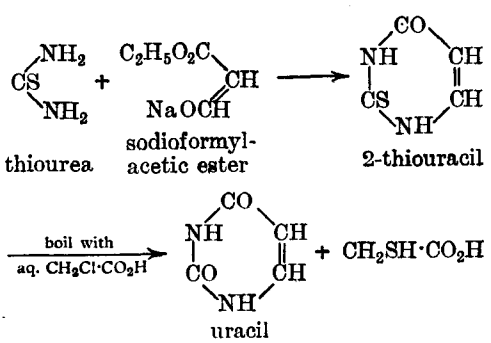
§15. Uracil (2 : 6-dihydroxypyrimidine) is a hydrolytic product of the nucleic acids (§§13, 13b. XVI). It has been synthesised in many ways, *e.g.*,

(i) Fischer and Roeder (1901).



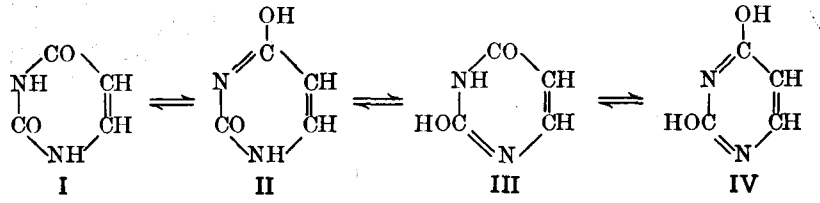
urea — ethyl acrylate — dihydrouracil

(ii) Wheeler and Liddle (1908).



thiourea — sodioformyl-acetic ester — 2-thiouracil — uracil

Four tautomeric structures are theoretically possible for uracil.

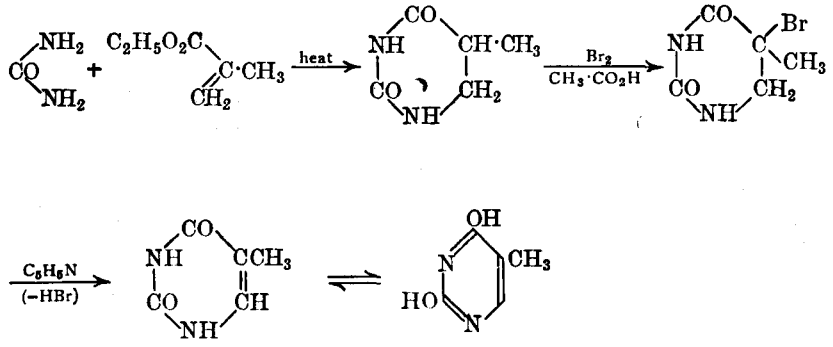


I II III IV

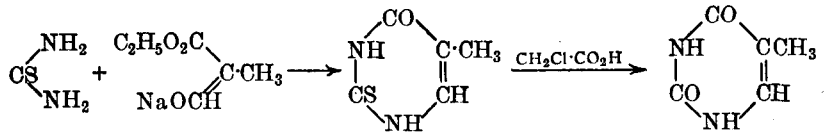
The ultraviolet absorption spectrum of uracil (in ethanol) is different from that of 1 : 3-dimethyluracil (a derivative of I), from that of 6-methoxy-3-methyluracil (a derivative of II), and from that of 2 : 6-diethoxyuracil (a derivative of IV). Thus uracil is probably III, and this is supported by the fact that the ultraviolet absorption spectrum of 1-methyluracil (a derivative of III) is similar to that of uracil (Austin, 1934) (but see also §13b. XVI).

§16. Thymine (5-methyluracil, 2 : 6-dihydroxy-5-methylpyrimidine) is a hydrolytic product of the nucleic acids. It has been synthesised by methods similar to those used for uracil.

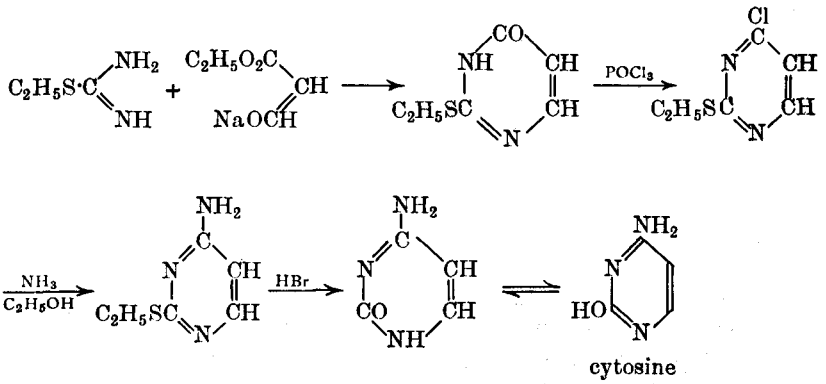
(i) Fischer and Roeder (1901); in this case ethyl methacrylate is used instead of ethyl acrylate.



(ii) Wheeler and Liddle (1908); in this case sodioformylpropionic ester is used instead of sodioformylacetic ester.

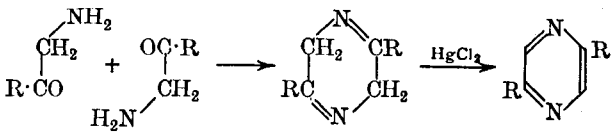


§17. Cytosine (6-aminouracil, 6-amino-2-hydroxypyrimidine) is a hydrolytic product of the nucleic acids. It has been synthesised by Wheeler and Johnson (1903) starting from *S*-ethyl*iso*thiourea and sodioformylacetic ester (see also §13b. XVI).



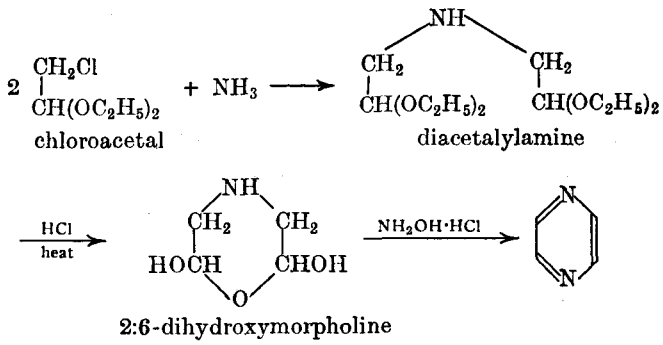
Pyrazines

§18. Pyrazines may be prepared by the self-condensation of an α-aminoketone in the presence of an oxidising agent such as mercuric chloride; the intermediate dihydro compound is readily oxidised to the pyrazine (Gabriel *et al.*, 1893).



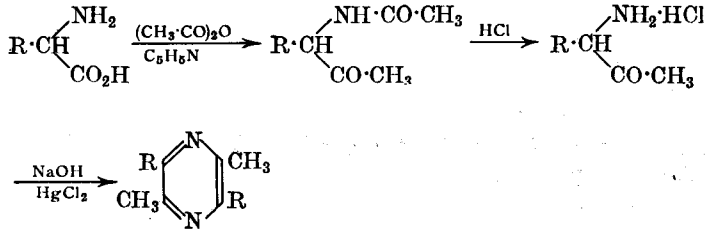
Actually, only the salts of α-aminoketo compounds are known; addition of alkali liberates the free base which immediately forms a pyrazine in the presence of mercuric chloride.

Pyrazine itself may be prepared from aminoacetaldehyde (R = H in the above equations). The best method, however, for preparing pyrazine is as follows (Wolff *et al.*, 1908).

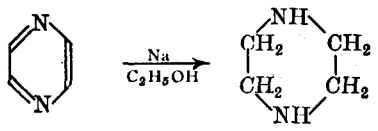


2:6-dihydroxymorpholine

A convenient general method for preparing pyrazines is to heat an α-amino-acid with acetic anhydride in the presence of pyridine, hydrolyse the product (an acetamidoketone) with acid and then warm with sodium hydroxide in the presence of mercuric chloride (Dakin *et al.*, 1928). This method is thus similar to the first general method given above, but offers a convenient method of preparing α-aminocarbonyl compounds.



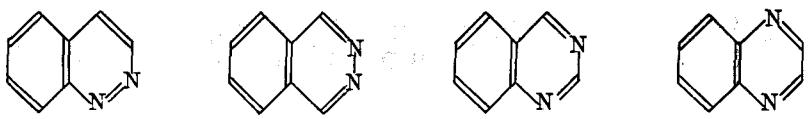
Pyrazine is a solid, m.p. 55°; pyrazines (and pyrazine) are readily reduced by sodium and ethanol to hexahydropyrazines or **piperazines.** Piperazine, m.p. 104°, is a strong diacid base. 2 : 5-Diketopiperazines are produced from α-amino-acids (see §4 C. XIII).



piperazine

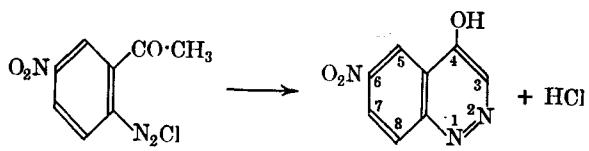
BENZODIAZINES

§19. The following benzodiazines are theoretically possible, and all are known; the first two are derived from pyridazine, the third from pyrimidine and the fourth from pyrazine.

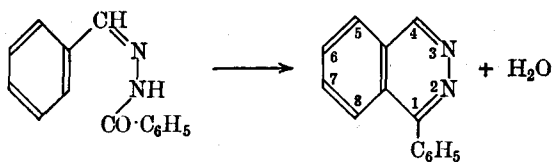


cinnoline phthalazine quinazoline quinoxaline

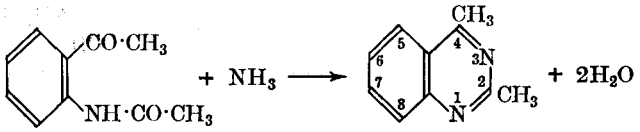
Cinnolines may be prepared by the cyclisation of diazotised *o*-amino-acetophenones (Schofield *et al.*, 1948), *e.g.*,



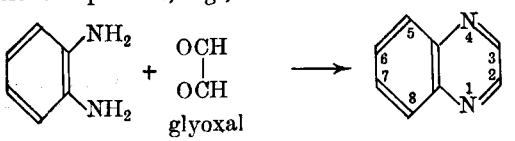
Phthalazines are formed by heating the benzoyl derivative of benzaldehyde hydrazones, *e.g.*,



Quinazolines may be prepared by the action of ammonia on acylated *o*-aminobenzaldehydes *or* *o*-aminoacetophenones (Isensee *et al.*, 1948), *e.g.*,

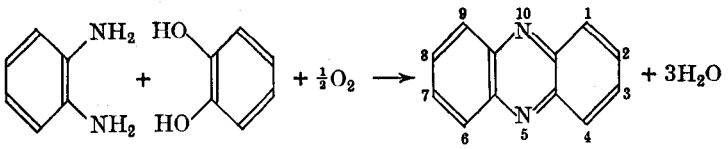


Quinoxalines are formed by the condensation of *o*-phenylenediamines with α-dioxo compounds, *e.g.*,



The formation of quinoxalines is used to identify aromatic *o*-diamines and 1 : 2-diketones (see, *e.g.*, §9. XVII).

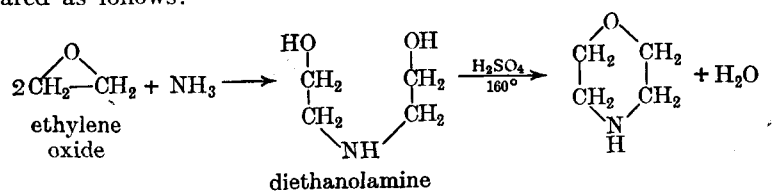
Of the *dibenzodiazines*, only the **phenazines** (*dibenzopyrazines*) are important. Phenazine, m.p. 171°, may be prepared by condensing *o*-phenylenediamine with catechol in the presence of air.



Phenazine forms unstable salts (coloured red or yellow) in excess of strong acids. Many dyes are derived from phenazine, *e.g.*, the safranines (see Vol. I).

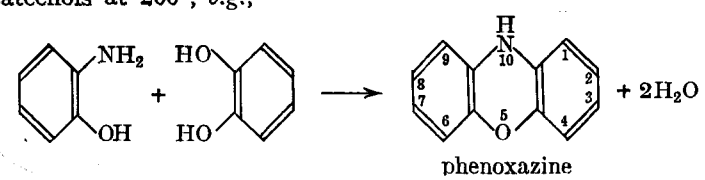
DIAZINES CONTAINING ONE NITROGEN ATOM AND AN OXYGEN OR SULPHUR ATOM

§20. Oxazines. *Morpholine* is tetrahydro-1 : 4-oxazine, and it may be prepared as follows:

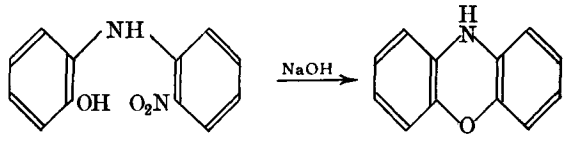


Morpholine is a liquid, b.p. 128°, and is strongly basic. It is miscible with water in all proportions, and is widely used as a solvent.

§21. Phenoxazines. These are formed by condensing *o*-aminophenols with catechols at 260°, *e.g.*,

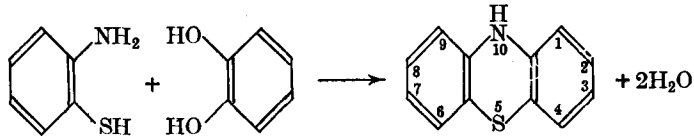


Phenoxazines are also produced by the action of alkali on 2-hydroxy-2'-nitrodiphenylamines, *e.g.*,



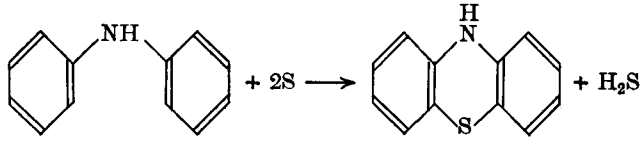
Phenoxazine is a solid, m.p. 156°; it is the parent substance of a number of dyes, *e.g.*, Meldola's Blue (see Vol. I).

§22. Thiazines. *Phenothiazines* may be prepared by heating *o*-aminothiophenols with catechols, *e.g.*,



phenothiazine

Phenothiazine may also be prepared by fusing diphenylamine with sulphur.



Phenothiazine, m.p. 185°, is used as an insecticide; it is the parent substance of a number of dyes, *e.g.*, Methylene Blue (see Vol. I).

TRIAZINES AND TETRAZINES

§23. Triazines. Three triazines are theoretically possible; the parent compounds are unknown, but derivatives of each have been prepared.

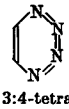
1:2:3-triazine; *vic.*-triazine; β-triazine

1:2:4-triazine; *as.*-triazine; α-triazine

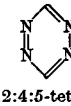
1:3:5-triazine; *sym.*-triazine; cyanidine

Cyanuric acid, cyamelide and hexamethylenetetramine are derivatives of *sym.*-triazine (see Vol. I).

§24. Tetrazines. Only derivatives of two tetrazines are known.

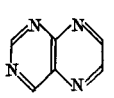


1:2:3:4-tetrazine; osotetrazine

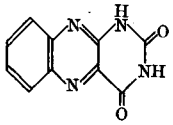


1:2:4:5-tetrazine; *sym.*-tetrazine

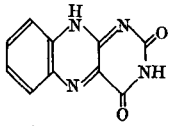
§25. Some important condensed systems containing *two* fused heterocyclic systems are:



pteridine



alloxazine



*iso*alloxazine

These occur in natural products (see Ch. XVII, Vitamins). It appears that *iso*alloxazine, the tautomer of alloxazine, does not exist as such; only when the hydrogen atom is substituted is the *iso*alloxazine form retained (see §6. XVII).

READING REFERENCES

Gilman (Ed.), *Advanced Organic Chemistry*, Wiley. Vol. IV (1953). Ch. 8. Heterocyclic Chemistry.

Morton, *The Chemistry of Heterocyclic Compounds*, McGraw-Hill (1946).

Acheson, *An Introduction to the Chemistry of Heterocyclic Compounds*, Interscience (1960).

Badger, *The Chemistry of Heterocyclic Compounds*, Academic Press (1961).

Rodd (Ed.), *Chemistry of the Carbon Compounds*, Elsevier. Vol. IVA, B and C (1958–1960). Heterocyclic Compounds.

Elderfield (Ed.), *Heterocyclic Compounds*, Wiley (1951–).

Patterson and Capell, *The Ring Index*, Reinhold (1940).

Handbook for Chemical Society Authors, Chem. Soc. (1960). Pp. 90–106. Heterocyclic Systems.

Finar and Simmonds, The Reaction between Aroylacetones and Arylhydrazines, *J.C.S.*, **1958**, 200.

Wright, The Chemistry of the Benzimidazoles, *Chem. Reviews*, 1951, **48**, 397.

Wiley, The Chemistry of the Oxazoles, *Chem. Reviews*, 1945, **37**, 401.

Organic Reactions, Wiley. Vol. VI (1951). Ch. 8. The Preparation of Thiazoles.

Benson and Savell, The Chemistry of the Vicinal Triazoles, *Chem. Reviews*, 1950, **46**, 1.

Potts, The Chemistry of 1,2,4-Triazoles, *Chem. Reviews*, 1961, **61**, 87.

Baker and Ollis, Meso-ionic Compounds, *Quart. Reviews* (*Chem. Soc.*), 1957, **11**, 15.

Benson, The Chemistry of the Tetrazoles, *Chem. Reviews*, 1947, **41**, 1.

Nineham, The Chemistry of Formazans and Tetrazolium Salts, *Chem. Reviews*, 1955, **55**, 355.

Franklin, Heterocyclic Nitrogen Compounds; Part I. Pentacyclic Compounds, *Chem. Reviews*, 1935, **16**, 305.

Johnson and Hahn, Pyrimidines; Their Amino and Amino-oxy Derivatives, *Chem. Reviews*, 1933, **13**, 193.

Shriner and Neumann, The Chemistry of the Amidines, *Chem. Reviews*, 1944, **35**, p. 395; The formation of substituted pyrimidines.

Lythgoe, Some Aspects of Pyrimidine and Purine Chemistry, *Quart. Reviews* (*Chem. Soc.*), 1949, **3**, 181.

Krems and Spoerri, The Pyrazines, *Chem. Reviews*, 1947, **40**, 279.

Leonard, The Chemistry of the Cinnolines, *Chem. Reviews*, 1945, **37**, 269.

Vaughan, The Chemistry of the Phthalazines, *Chem. Reviews*, 1948, **43**, 447.

Gates, The Chemistry of the Pteridines, *Chem. Reviews*, 1947, **41**, 63.

King, Three- and Four-Membered Heterocyclic Rings, *J.C.S.*, **1949**, 1318.

CHAPTER XIII

AMINO-ACIDS AND PROTEINS

§1. Classification of the amino-acids. When hydrolysed by acids, alkalis or enzymes, proteins (§6) yield a mixture of amino-acids. Acid hydrolysis destroys certain amino-acids, particularly tryptophan. On the other hand, alkaline hydrolysis causes complete racemisation and also the destruction of a number of amino-acids, *e.g.*, serine, threonine, cysteine, etc. Enzymic hydrolysis has also difficulties, particularly the long time that is usually needed and the fact that the hydrolysis is often not complete. Thus acid hydrolysis is the most satisfactory, but enzymic hydrolysis is very useful for the isolation of tryptophan. Gurnani *et al.* (1955) have introduced an improved method for the hydrolysis of proteins. The tissue is first dissolved in 85 per cent. formic acid and then 2N hydrochloric acid is added; all the amino-acids, except tryptophan, are liberated within two hours. The number of amino-acids so far obtained from proteins appears to be about twenty-five, all of which except two are α-amino-acids; the two exceptions are proline and hydroxyproline, which are imino-acids (see list of amino-acids below). Ten of the amino-acids are essential acids, *i.e.*, a deficiency in any one prevents growth in young animals, and may even cause death. The amino-acids are classified in several ways; the table on pages 452 and 453 shows a convenient classification; the letters *g*, *l* and *e* which follow the name of the acids indicate that the acid is respectively of general occurrence, lesser occurrence and essential (to man).

The α-amino-acids listed in the table have been isolated from proteins. Plants have continued to provide new amino-acids of diverse structure; between 1950 and 1960 about fifty amino- or imino-acids have been identified as components of higher plants. About 20 more have been recognised as constituents of micro-organisms or have been obtained as fragments of the antibiotics excreted by the micro-organisms. These discoveries are the result of the application of paper and ion-exchange chromatography to the examination of plant extracts.

§2. General methods of preparation of the amino-acids. There are many general methods for preparing α-amino-acids, but usually each method applies to a small number of particular acids; many acids are also synthesised by methods special to an individual. It should also be noted that very often a synthesis is a more convenient way of preparing an amino-acid than preparing it from natural sources.

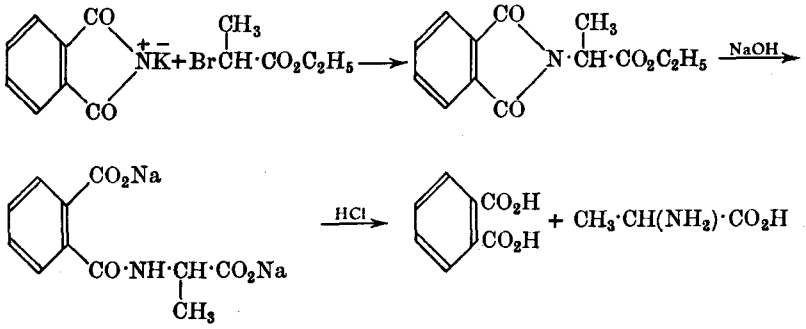
(i) **Amination of α-halogenated acids** (Perkin *et al.*, 1858).

(*a*) An α-chloro- or bromo-acid is treated with concentrated ammonia, *e.g.*,

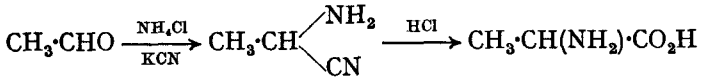
$$CH_2Cl{\cdot}CO_2H + 2NH_3 \longrightarrow CH_2(NH_2){\cdot}CO_2H + NH_4Cl$$

This method is convenient for the preparation of glycine, alanine, serine, threonine, valine, leucine and norleucine.

(*b*) The yields obtained by the above method are variable because of side-reactions. Better yields are obtained by using *Gabriel's phthalimide synthesis* (1889) with α-halogeno-acids (see also Vol. I), *e.g.*,

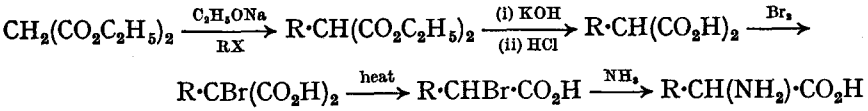


(ii) **Strecker synthesis** (1850). A cyanohydrin is treated with concentrated ammonia, and the resulting amino-nitrile is then hydrolysed with acid. In practice the amino-nitrile is usually prepared from the oxo compound in one step by treating the latter with an equimolecular mixture of ammonium chloride and potassium cyanide (this mixture is equivalent to ammonium cyanide), *e.g.*,



This method is useful for preparing the following amino-acids: glycine, alanine, serine, valine, methionine, glutamic acid, leucine, norleucine and phenylalanine.

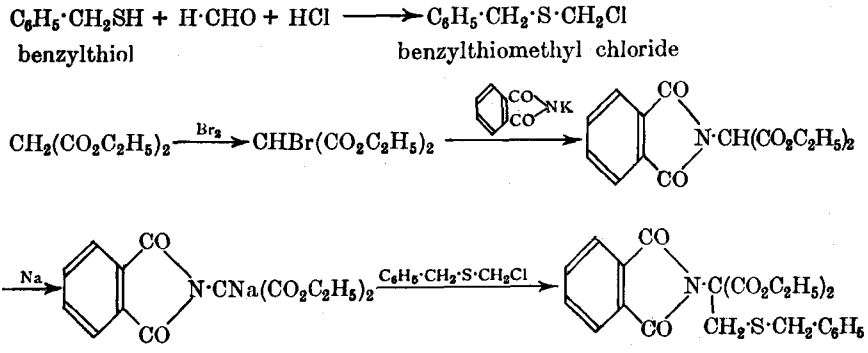
(iii*a*) **Malonic ester synthesis.** This method is really an extension of (i) *a*; it offers a means of preparing α-halogeno-acids, *e.g.*,

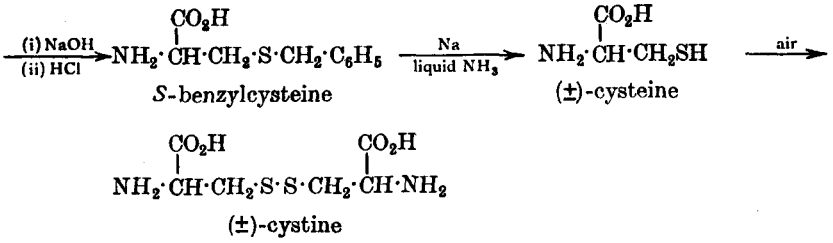


This method offers a means of preparing, from readily accessible materials, the following acids: phenylalanine, proline, leucine, *iso*leucine, norleucine and methionine.

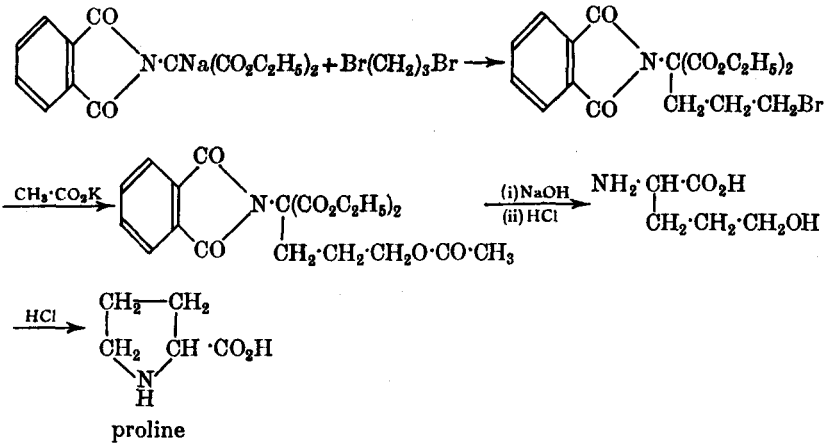
The malonic ester synthesis may also be combined with the Gabriel phthalimide synthesis to prepare phenylalanine, tyrosine, proline, cystine, serine, aspartic acid, methionine and lysine, *e.g.*,

Cystine.

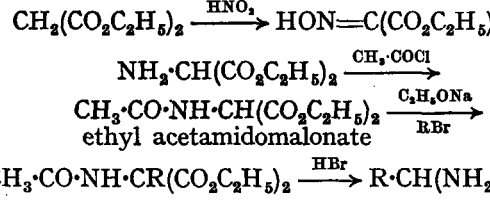




Proline.

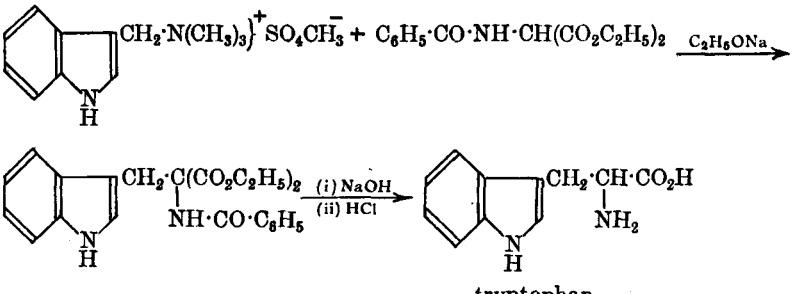


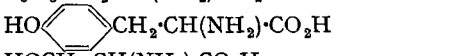
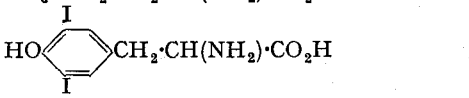
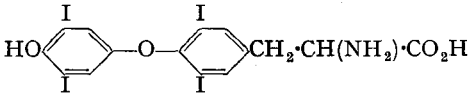
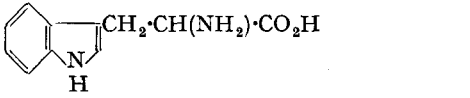
Acylamido derivatives of malonic ester may also be used to synthesise amino-acids; the usual derivative employed is ethyl acetamidomalonate (Albertson, 1946).

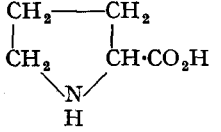
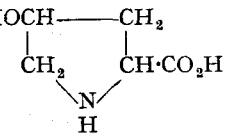
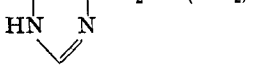


The following acids may be prepared by this method: serine, leucine, valine, methionine, lysine, glutamic acid and ornithine.

A special application of this method is the preparation of tryptophan from benzamidomalonic ester and gramine methosulphate (Albertson *et al.*, 1945; Tishler *et al.*, 1945).



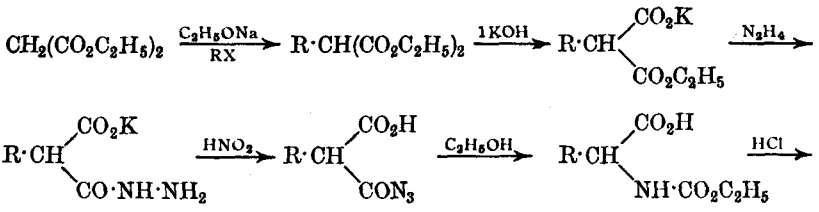
Name	Systematic Name	Formula
Neutral Amino-acids (one amino-group and one carboxyl group)		
1. Glycine (*g*)	Aminoacetic acid	$CH_2(NH_2)\cdot CO_2H$
2. Alanine (*g*)	α-Aminopropionic acid	$CH_3\cdot CH(NH_2)\cdot CO_2H$
3. Valine (*g*, *e*)	α-Amino*iso*valeric acid	$(CH_3)_2CH\cdot CH(NH_2)\cdot CO_2H$
4. Leucine (*g*, *e*)	α-Amino*iso*caproic acid	$(CH_3)_2CH\cdot CH_2\cdot CH(NH_2)\cdot CO_2H$
5. *iso*Leucine (*g*, *e*)	α-Amino-β-methyl-*n*-valeric acid	$CH_3\cdot CH_2\cdot CH(CH_3)\cdot CH(NH_2)\cdot CO_2H$
6. Norleucine[1] (*l*)	α-Amino-*n*-caproic acid	$CH_3\cdot (CH_2)_3\cdot CH(NH_2)\cdot CO_2H$
7. Phenylalanine (*g*, *e*)	α-Amino-β-phenylpropionic acid	$C_6H_5\cdot CH_2\cdot CH(NH_2)\cdot CO_2H$
8. Tyrosine (*g*)	α-Amino-β-(*p*-hydroxyphenyl)propionic acid	
9. Serine (*g*)	α-Amino-β-hydroxypropionic acid	$HOCH_2\cdot CH(NH_2)\cdot CO_2H$
10. Cysteine[2] (*g*)	α-Amino-β-mercaptopropionic acid	$HSCH_2\cdot CH(NH_2)\cdot CO_2H$
11. Cystine (*g*)	Bis-(α-aminopropionic acid)-β-disulphide	$[-SCH_2\cdot CH(NH_2)\cdot CO_2H]_2$
12. Threonine (*g*, *e*)	α-Amino-β-hydroxy-*n*-butyric acid	$CH_3\cdot CHOH\cdot CH(NH_2)\cdot CO_2H$
13. Methionine (*g*, *e*)	α-Amino-γ-methylthio-*n*-butyric acid	$CH_3SCH_2\cdot CH_2\cdot CH(NH_2)\cdot CO_2H$
14. Iodogorgic acid (*l*)	3 : 5-Di-iodotyrosine	
15. Thyroxine (*l*)	β-3: 5-Di-iodo-4-(3′ : 5′-di-iodo-4′-hydroxy)phenyl-α-aminopropionic acid	
16. Tryptophan (*g*, *e*)	α-Amino-β-indolepropionic acid	

17. Proline (*g*)	Pyrrolidine-α-carboxylic acid	
18. Hydroxyproline (*l*)	γ-Hydroxypyrrolidine-α-carboxylic acid	
Acidic Amino-acids (one amino-group and two carboxyl groups)		
19. Aspartic acid (*g*)	α-Aminosuccinic acid	$CO_2H{\cdot}CH_2{\cdot}CH(NH_2){\cdot}CO_2H$
20. Asparagine (*l*)	α-Aminosuccinamic acid	$CONH_2{\cdot}CH_2{\cdot}CH(NH_2){\cdot}CO_2H$
21. Glutamic acid (*g*)	α-Aminoglutaric acid	$HO_2C{\cdot}CH_2{\cdot}CH_2{\cdot}CH(NH_2){\cdot}CO_2H$
22. β-Hydroxyglutamic acid[3]	α-Amino-β-hydroxyglutaric acid	$HO_2C{\cdot}CH_2{\cdot}CHOH{\cdot}CH(NH_2){\cdot}CO_2H$
23. Glutamine (*l*)	α-Aminoglutaramic acid	$CONH_2{\cdot}CH_2{\cdot}CH_2{\cdot}CH(NH_2){\cdot}CO_2H$
Basic Amino-acids (two amino-groups and one carboxyl group)		
24. Ornithine[4]	α : δ-Diamino-*n*-valeric acid	$NH_2{\cdot}CH_2{\cdot}CH_2{\cdot}CH_2{\cdot}CH(NH_2){\cdot}CO_2H$
25. Arginine (*g*, *e*)	α-Amino-δ-guanidino-*n*-valeric acid	$NH{=}C(NH_2){\cdot}NH{\cdot}(CH_2)_3{\cdot}CH(NH_2){\cdot}CO_2H$
26. Lysine (*g*, *e*)	α : ε-Diaminocaproic acid	$NH_2{\cdot}(CH_2)_4{\cdot}CH(NH_2){\cdot}CO_2H$
27. Histidine (*g*, *e*)	α-Amino-β-imidazolepropionic acid	

[1] The occurrence of norleucine in proteins is uncertain.
[2] Cysteine has not yet definitely been shown to be present in proteins, but its presence is inferred from various tests.
[3] The occurrence of β-hydroxyglutamic acid in proteins is uncertain.
[4] Ornithine is probably not present in proteins, but is formed by the hydrolysis of arginine.

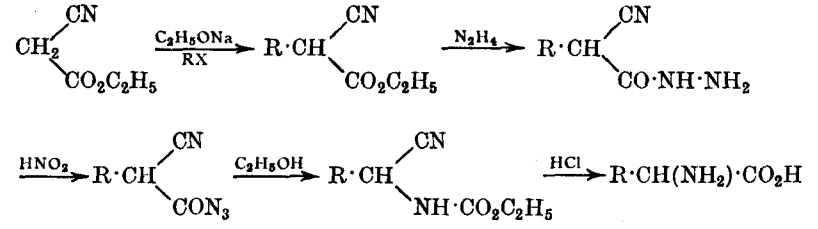
A more recent method of preparing ethyl acetamidomalonate is to reduce oximinomalonic ester in a mixture of acetic anhydride, pyridine and sodium acetate with hydrogen in the presence of Raney nickel (Vignau, 1952).

(iii*b*) α-Amino-acids may be synthesised by means of the **Curtius reaction** (see also Vol. I).



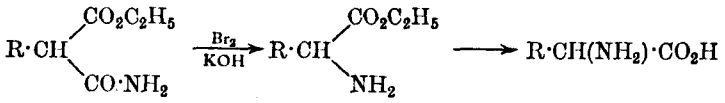
Glycine, alanine, phenylalanine and valine can be prepared by this method.

Instead of malonic ester, the starting material can be ethyl cyanoacetate.

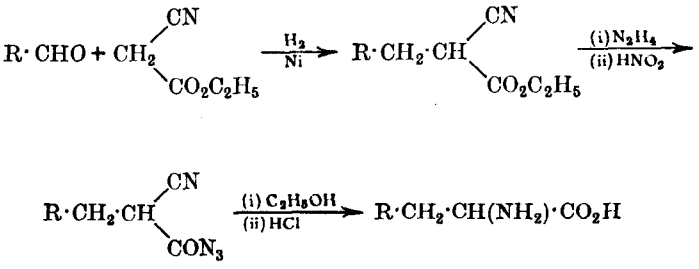


Phenylalanine and tyrosine are conveniently prepared by this method.

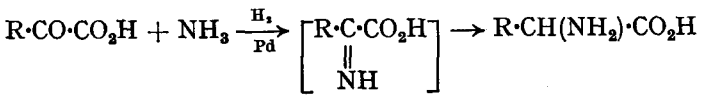
Another variation is the use of the Hofmann degradation on ester amides (see also Vol. I).



(iii*c*) **The Darapsky synthesis** (1936). In this method an aldehyde is condensed with ethyl cyanoacetate and simultaneously hydrogenated; the product, an alkylcyanoacetic ester, is then treated as above (for the cyanoacetic ester method).

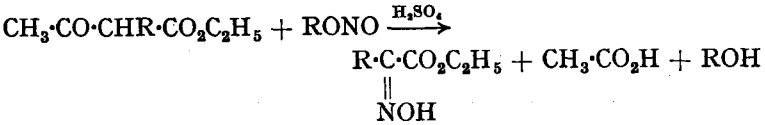


(iv) Amino-acids may be prepared by reducing α-ketonic acids in the presence of ammonia; the reduction may be performed catalytically, or with sodium and ethanol. The mechanism of the reaction is not certain, but it probably occurs *via* the imino-acid.

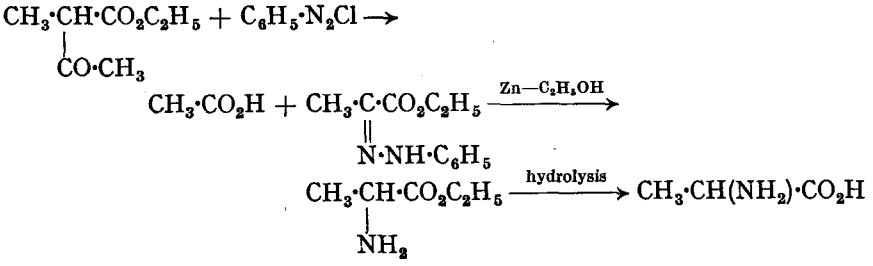


This method works well for alanine and glutamic acid.

Oximes of α-keto-acids may also be reduced to α-amino-acids. The advantage of this method is that the oximes may readily be prepared in good yield by the action of sulphuric acid on a mixture of an alkylacetoacetic ester and an alkyl nitrite (Hartung *et al.*, 1942).

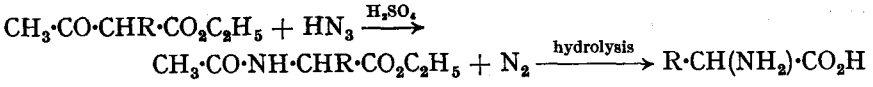


The reduction of phenylhydrazones made by the action of a diazonium salt on an alkylacetoacetic ester also may be used to prepare α-amino-acids (*cf.* the Japp–Klingermann reaction, Vol. I); *e.g.*,

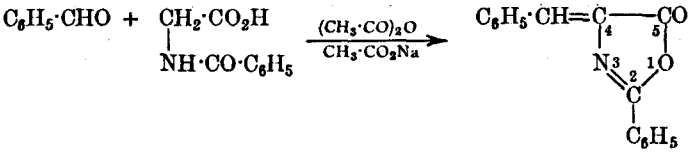


Thus alanine, phenylalanine, leucine, *iso*leucine, valine and hydroxyproline may be prepared in this way.

Alkylacetoacetic esters may also be converted into α-amino-acids by means of the Schmidt reaction (see also Vol. I).

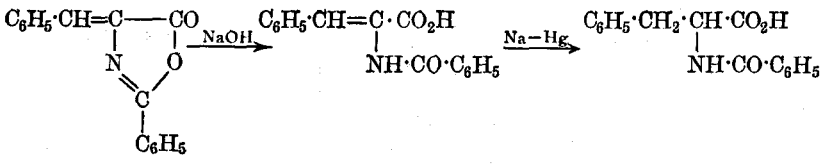


(v*a*) **The Azlactone synthesis** (*Erlenmeyer synthesis*, 1893). Azlactones are usually prepared by heating an aromatic aldehyde with hippuric acid (benzoylglycine) in the presence of acetic anhydride and sodium acetate, *e.g.*, benzaldehyde forms benzoyl-α-aminocinnamic azlactone (4-benzylidene-2-phenyloxazol-5-one).



This reaction is usually referred to as the **Erlenmeyer azlactone synthesis.** Aceturic acid (acetylglycine) may also be used instead of hippuric acid. Furthermore, it has been found that aliphatic aldehydes may condense with hippuric acid to form azlactones if lead acetate is used instead of sodium acetate (Finar *et al.*, 1949).

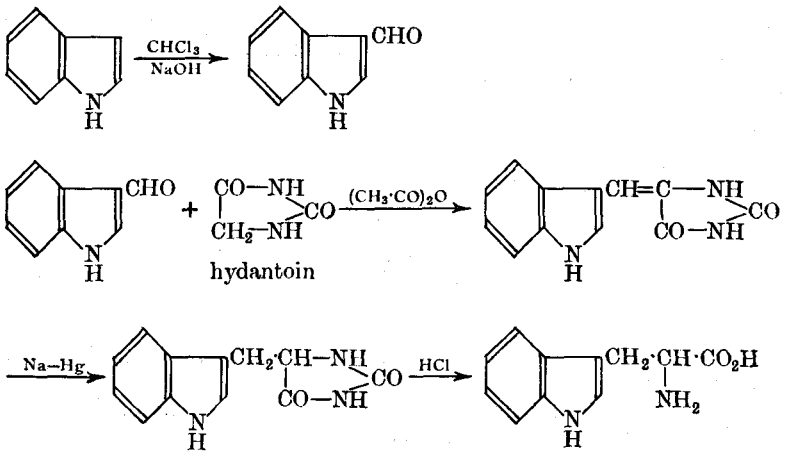
When azlactones are warmed with one per cent. sodium hydroxide solution, the ring is opened, and if the product is reduced with sodium amalgam followed by hydrolysis with acid, an α-amino-acid is produced, *e.g.*,



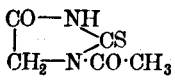
$\xrightarrow{HCl}$ $C_6H_5\cdot CH_2\cdot CH(NH_2)\cdot CO_2H + C_6H_5\cdot CO_2H$

The azlactone synthesis offers a convenient means of preparing phenylalanine, tyrosine, tryptophan and thyroxine.

(v*b*) Aromatic aldehydes also condense with hydantoin, and reduction of the product with sodium amalgam or ammonium hydrogen sulphide, followed by hydrolysis, gives an α-amino-acid, *e.g.*, tryptophan may be prepared by first converting indole into indole-3-aldehyde by means of the Reimer–Tiemann reaction (see Vol. I).



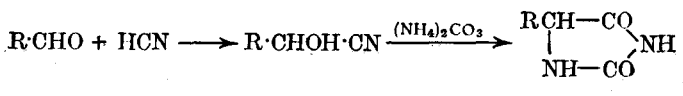
This method has been improved by using acetylthiohydantoin instead of hydantoin.



acetylthiohydantoin

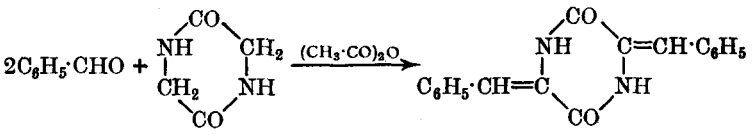
The above method may be used to prepare phenylalanine, tyrosine, tryptophan and methionine.

Another modification of the hydantoin synthesis is the **Bücherer hydantoin synthesis** (1934). In this method an oxo compound is converted into the cyanohydrin and this, on treatment with ammonium carbonate, produces a 5-substituted hydantoin which, on hydrolysis, gives an α-amino-acid.



$\xrightarrow{HCl}$ $R\cdot CH(NH_2)\cdot CO_2H$

(v*c*) Aromatic aldehydes may be condensed with diketopiperazine, and the product converted into an amino-acid by heating with hydriodic acid and red phosphorus, *e.g.*,



$$\xrightarrow[P]{HI} \quad 2C_6H_5{\cdot}CH_2{\cdot}CH(NH_2){\cdot}CO_2H$$

Phenylalanine, tyrosine and methionine may be prepared by this method.

§3. Isolation of amino-acids from protein hydrolysates. Many amino-acids can be detected colorimetrically, and these colour reactions have now been developed for quantitative estimation. Also, amino-acids containing a benzene or pyrrolidine nucleus have characteristic absorption spectra; thus the presence of such acids can readily be ascertained.

The actual *quantitative isolation* of amino-acids from their mixtures is a difficult problem. The earliest method was the fractional distillation of the amino-acid esters *in vacuo* (Fischer, 1901). This method is very little used now, and is only satisfactory for the neutral amino-acids (*i.e.*, those containing one amino-group and one carboxyl group).

Neutral amino-acids may be extracted by *n*-butanol saturated with water, and then separated by fractional crystallisation or by the fractional distillation of the esters. After the butanol extraction, the residue may be treated with phosphotungstic acid, whereupon the *basic* amino-acids are precipitated (Dakin *et al.*, 1913).

A number of individual amino-acids can be obtained by means of selective precipitation as salts, *e.g.*, lysine is precipitated by picric acid.

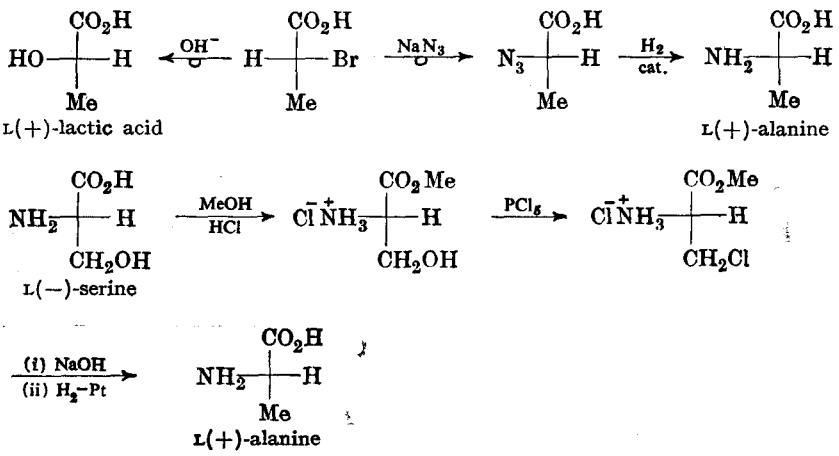
Mixtures of amino-acids may be separated into fractions consisting of the neutral, basic and acidic acids by means of the *electrical transport method.* In this method a P.D. is applied to the mixture at the proper *p*H; the basic acids (positively charged) migrate to the cathode compartment, the acidic acids (negatively charged) migrate to the anode compartment, and the neutral acids remain in the centre compartment.

The most satisfactory method of analysing amino-acid mixtures is partition chromatography carried out on paper (Martin *et al.*, 1944). The mixture of amino-acids is partitioned between a stationary water phase adsorbed on a strip or sheet of filter paper and a moving phase of some organic solvent (butanol, phenol, etc.). The moving phase either ascends or descends the paper strip (according to the way the experiment is performed). A small amount of the amino-acid solution is applied to one end of the paper, the strip then placed in a suitable glass container containing the organic solvent saturated with water, and when the solvent front has progressed a suitable distance, the distance moved by the solvent is measured, the strip dried, and then sprayed with a dilute solution of ninhydrin in butanol (see also §4C). Coloured spots are produced at the positions of the various amino-acids. The ratio of the distance travelled by the amino-acid to the distance travelled by the solvent is characteristic of each amino-acid, and is known as the R_F value (this value depends on the experimental conditions).

A very interesting analytical method is the *microbiological assay.* This depends on the fact that micro-organisms can be " trained " to feed on a specific amino-acid in the nutrient medium. The rate of growth of the micro-organism is first measured by breeding in a medium containing the particular amino-acid, and then the rate of growth is measured in the mixture of amino-acids to be analysed. In this way it is possible to determine the amounts of various amino-acids in protein hydrolysates *without* isolation of the acids. Another method of analysis is that of *isotopic dilution.*

Suppose the amount of glycine is to be estimated. A weighed amount of labelled glycine is added to the hydrolysate, and then glycine is isolated by one of the standard methods. The amount of *labelled* glycine in this specimen is now measured, *e.g.*, say it is 1 per cent. Thus for every 1 g. of labelled glycine there are 99 g. of ordinary glycine. Since the weight of the added labelled glycine is known, the total weight of glycine in the mixture can therefore be calculated (see also Vol. I).

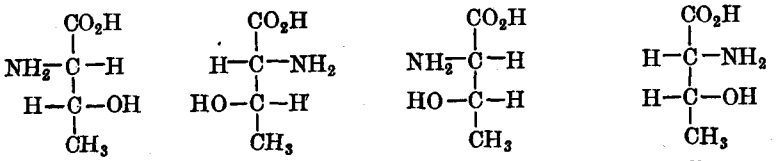
§4. General properties of the amino-acids. The amino-acids are colourless crystalline compounds which are generally soluble in water but sparingly soluble in organic solvents; most melt with decomposition, but Gross *et al.* (1955) have shown that sublimation is possible with a number of amino-acids. All except glycine contain at least one asymmetric carbon atom, and all (except glycine) occur naturally in their optically active forms. It has been mentioned in §5b. II that natural (−)-serine was chosen as the arbitrary standard for correlating the configurations of amino-acids, the relationship to this acid being indicated by D$_s$ or L$_s$. It has now been shown that L$_a$ ≡ L$_s$, *i.e.*, natural (−)-serine belongs to the L-series (with glyceraldehyde as absolute standard). The correlation between the two standards was established as follows. (+)-Alanine has been correlated with L(+)-lactic acid (for the correlation of the latter with L(−)-glyceraldehyde see §5b i. II); and L(+)-alanine has been correlated with L(−)-serine:



A new method for determining the configuration of an α-amino-acid is by studying the rotatory dispersion curves of the *N*-alkylthio derivatives. L-Compounds show a positive Cotton effect, whereas the D-compounds show a negative effect (see §12a. I). It has been shown that the α-carbon atom, *i.e.*, the carbon atom attached to the amino-group, has, in almost all the amino-acids, the same configuration as L(−)-glyceraldehyde. The specific rotation of the amino-acids depends on the *p*H of the solution, the temperature, the presence of salts and the nature of the solvent (see §12. I). The racemic amino-acids may be resolved by first formylating and then resolving the formyl derivatives *via* the salt with an optically active base, and finally removing the formyl group by hydrolysis (see also C i). Alternatively, racemic amino-acids may be resolved by means of enzymes (see §10 iv. II). A more recent method is the selective destruction of one or other enantiomorph of a racemate by a specific D- or L-oxidase (Parikh *et al.*, 1958); the optical purity of the product is greater than 99·9 per cent. As pointed out above, most natural amino-acids are L; these are obtained by acid or

enzymic hydrolysis of proteins. Alkaline hydrolysis of proteins gives the DL-amino-acids (§1), and so does the synthetic preparation; it is by resolution of the synthetic racemic modification that the D-amino-acids are frequently prepared.

The symbols D and L are used for the configuration of the α-carbon atom (see above), and the symbols (+) and (−) are used to indicate the direction of the rotation (*cf.* §5. II). When two asymmetric centres are present, then D and L still refer to the α-carbon atom, and the *naturally occurring acid* is known as the L-amino-acid. The *allo*-form is the name given to that form in which the configuration of the *second asymmetric carbon atom is inverted, e.g.*, L(−)-threonine (the naturally occurring form), D(+)-threonine, L-*allo*threonine and D-*allo*threonine.



L(−)-threonine D(+)-threonine L-*allo*threonine D-*allo*threonine

Since they contain amino and carboxyl groups, the amino-acids possess the properties of both a base and an acid, *i.e.*, they are amphoteric.

A. Reactions due to the amino-group.

(i) The amino-acids form salts with strong inorganic acids, *e.g.*,

$$\overset{-}{Cl}\{H_3\overset{+}{N}\cdot CH_2\cdot CO_2H.$$

These salts are usually sparingly soluble in water, and the free acid may be liberated from its salt by means of a strong organic base, *e.g.*, pyridine.

(ii) Amino-acids may be acetylated by means of acetyl chloride or acetic anhydride.

$$R\cdot CH(NH_2)\cdot CO_2H + (CH_3\cdot CO)_2O \longrightarrow R\cdot CH(NH\cdot CO\cdot CH_3)\cdot CO_2H + CH_3\cdot CO_2H$$

Similarly, benzoyl chloride produces the benzoyl derivative. These acetylated derivatives are acidic, the basic character of the amino-group being effectively eliminated by the presence of the negative group attached to the nitrogen. It should also be noted that the carboxyl group of one molecule can react with the amino-group of another molecule of an amino-acid to form a peptide (see §9). Sanger (1945) has shown that 1-fluoro-2 : 4-dinitrobenzene combines with amino-acids to form dinitrophenyl derivatives (see §11).

(iii) Nitrous acid liberates nitrogen from amino-acids.

$$R\cdot CH(NH_2)\cdot CO_2H + HNO_2 \longrightarrow R\cdot CHOH\cdot CO_2H + N_2 + H_2O$$

The nitrogen is evolved quantitatively, and this forms the basis of the **van Slyke method** (1911) for analysing mixtures of amino-acids.

(iv) Nitrosyl chloride (or bromide) reacts with amino-acids to form chloro- (or bromo) acids.

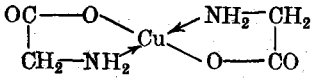
$$R\cdot CH(NH_2)\cdot CO_2H + NOCl \longrightarrow R\cdot CHCl\cdot CO_2H + N_2 + H_2O$$

(v) When heated with hydriodic acid at 200°, the amino-group is eliminated with the formation of a fatty acid.

$$R\cdot CH(NH_2)\cdot CO_2H \xrightarrow[200°]{HI} R\cdot CH_2\cdot CO_2H + NH_3$$

B. Reactions due to the carboxyl group.

(i) Amino-acids form salts; the salts of the heavy metals are chelate compounds, *e.g.*, the copper salt of glycine (deep blue needles) is formed by heating copper oxide with an aqueous solution of glycine.



The amino-acids may be liberated from their alkali salts by treatment in ethanolic solution with ethyl oximinocyanoacetate (Galat, 1947).

(ii) When heated with an alcohol in the presence of dry hydrogen chloride, amino-acids form ester hydrochlorides, *e.g.*,

$$NH_2{\cdot}CH_2{\cdot}CO_2H + C_2H_5OH + HCl \longrightarrow \bar{Cl}\{H_3\overset{+}{N}{\cdot}CH_2{\cdot}CO_2C_2H_5 + H_2O$$

The free ester may be obtained by the action of aqueous sodium carbonate on the ester salt. The esters are fairly readily hydrolysed to the amino-acid by aqueous sodium hydroxide (even at room temperature). These esters may be reduced to the amino-alcohols by means of sodium and ethanol, or hydrogenated in the presence of Raney nickel. Amino-acids may be reduced directly to the amino-alcohol with lithium aluminium hydride, and in this case no racemisation occurs (Vogel *et al.*, 1952).

$$R{\cdot}CH(NH_2){\cdot}CO_2H \xrightarrow{LiAlH_4} R{\cdot}CH(NH_2){\cdot}CH_2OH$$

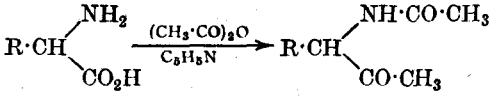
(iii) When suspended in acetyl chloride and then treated with phosphorus pentachloride, amino-acids form the hydrochloride of the acid chloride.

$$R{\cdot}CH(NH_2){\cdot}CO_2H + PCl_5 \longrightarrow \bar{Cl}\{H_3\overset{+}{N}{\cdot}CHR{\cdot}COCl + POCl_3$$

(iv) Dry distillation, or better by heating with barium oxide, decarboxylates amino-acids to amines.

$$R{\cdot}CH(NH_2){\cdot}CO_2H \longrightarrow R{\cdot}CH_2{\cdot}NH_2 + CO_2$$

(v) When heated with acetic anhydride in pyridine solution, amino-acids are converted into methyl α-acetamidoketones (Dakin *et al.*, 1928; see also §18. XII); this reaction is often referred to as the **Dakin–West reaction.**



C. Reactions due to both the amino and carboxyl groups.

(i) When measured in aqueous solution, the dipole moment of glycine (and other amino-acids) is found to have a large value. To account for this large value it has been suggested that glycine exists, in solution, as an *inner salt*:

$$NH_2{\cdot}CH_2{\cdot}CO_2H + H_2O \rightleftharpoons \overset{+}{N}H_3{\cdot}CH_2{\cdot}\bar{C}O_2 + H_2O$$

Such a doubly charged ion is also known as a *zwitterion, ampholyte* or a *dipolar ion*. This dipolar ion structure also accounts for the absence of acidic and basic properties of an amino-acid (the carboxyl and amino-groups of the *same* molecule neutralise each other to form a salt). The properties of crystalline glycine, *e.g.*, its high melting point and its insolubility in hydrocarbon solvents, also indicate that it exists as the inner salt in the solid state.

Each amino-acid has a definite *p*H at which it does not migrate to either electrode when a P.D. is applied. This *p*H is known as the **isoelectric point,** and at this point the amino-acid has its lowest solubility.

Owing to their amphoteric character, amino-acids cannot be titrated directly with alkali. When formalin solution is added to glycine, methyleneglycine is formed.

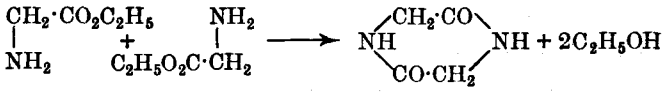
$$NH_2{\cdot}CH_2{\cdot}CO_2H + H{\cdot}CHO \longrightarrow CH_2{=}N{\cdot}CH_2{\cdot}CO_2H + H_2O$$

Although some methyleneglycine is probably formed, it appears that the reaction is more complex; the main product appears to be dimethylolglycine.

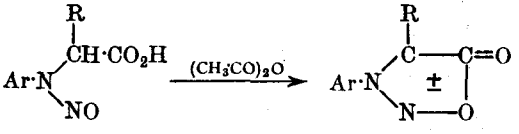
$$NH_2{\cdot}CH_2{\cdot}CO_2H + 2H{\cdot}CHO \longrightarrow (CH_2OH)_2N{\cdot}CH_2{\cdot}CO_2H$$

These glycine derivatives are strong acids (the basic character of the amino-group being now suppressed), and can be titrated with alkali. This method is known as the *Sörensen formol titration.*

(ii) When heated, α-amino-acids form 2 : 5-diketopiperazines; esters give better yields; *e.g.*, diketopiperazine from glycine ester.



(iii) *N*-alkyl or arylamino-acids form *N*-nitroso derivatives with nitrous acid, and these may be dehydrated to sydnones by means of acetic anhydride (see §8. XII).



(iv) **Betaines.** These are the trialkyl derivatives of the amino-acids; betaine itself may be prepared by heating glycine with methyl iodide in methanolic solution. The betaines exist as dipolar ions; thus the formation of betaine may be written:

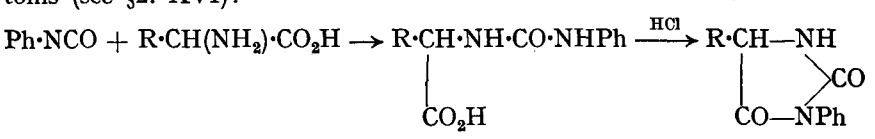
$$H_3\overset{+}{N}{\cdot}CH_2{\cdot}\overset{-}{CO_2} + 3CH_3I \longrightarrow (CH_3)_3\overset{+}{N}{\cdot}CH_2{\cdot}\overset{-}{CO_2} + 3HI$$

Betaine is more conveniently prepared by warming an aqueous solution of chloroacetic acid with trimethylamine.

$$(CH_3)_3N + ClCH_2{\cdot}CO_2H \longrightarrow (CH_3)_3\overset{+}{N}{\cdot}CH_2{\cdot}\overset{-}{CO_2} + HCl$$

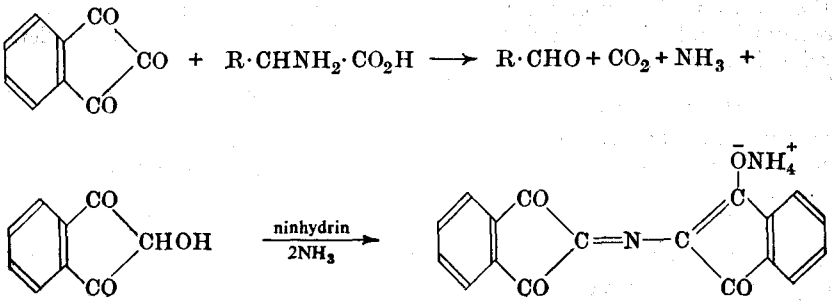
Betaine is a solid, m.p. 300° (with decomposition). It occurs in nature, especially in plant juices. It behaves as a base, *e.g.*, with hydrochloric acid it forms the stable crystalline hydrochloride, $\overset{-}{Cl}\{(CH_3)_3\overset{+}{N}{\cdot}CH_2{\cdot}CO_2H$.

(v) Amino-acids react with phenyl *iso*cyanate to form phenylhydantoic acids, and these, on treatment with hydrochloric acid, readily form hydantoins (see §2. XVI):



If phenyl *iso*thiocyanate is used instead of the *iso*cyanate, then thiohydantoins are produced (see §11).

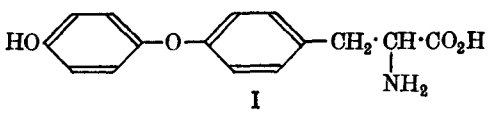
(vi) **Ninhydrin reaction.** Ninhydrin (indane-1 : 2 : 3-trione hydrate) reacts with amino-acids as follows:



The amino-acid is oxidised to aldehyde and the ninhydrin is reduced to 1 : 3-diketoindan-2-ol. The latter then reacts with another molecule of ninhydrin and with ammonia (which is produced in the first reaction) to form a coloured product. This reaction is the basis of a colorimetric method for estimating amino-acids.

§5. Thyroxine (thyroxin). Thyroxine is a hormone; it is the active principle of the thyroid gland and was first isolated by Kendall (1919). It was later isolated by Harington (1930) as a white crystalline solid, m.p. 235°, with a lævorotation.

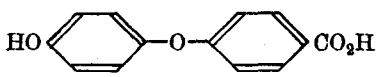
The structure of thyroxine was established by Harington (1926). This author showed that the molecular formula of thyroxine is $C_{15}H_{11}O_4NI_4$. When treated in alkaline solution with hydrogen in the presence of colloidal palladium, the iodine in thyroxine is replaced by hydrogen to form thyronine (thyronin), $C_{15}H_{15}O_4N$. This behaves as a phenol and an α-amino-acid. On fusion with potassium hydroxide in an atmosphere of hydrogen, thyronine gives a mixture of *p*-hydroxybenzoic acid, quinol, oxalic acid and ammonia. When fused with potassium hydroxide at 250°, thyronine gives *p*-hydroxybenzoic acid, quinol and a compound with the molecular formula $C_{13}H_{12}O_2$ (II). A structure for thyronine which would give all these products is I.



I

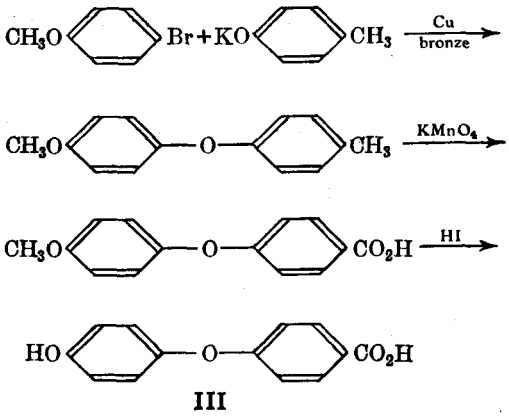
thyronine

Thyronine (provisionally structure I) was subjected to the Hofmann exhaustive methylation (see §4. XIV) and the product thereby obtained was then oxidised. The final product would be III (on the assumption that I is thyronine).



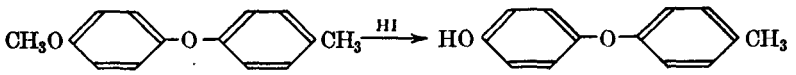
III

The structure of III was confirmed by synthesis, starting from *p*-bromoanisole and *p*-cresol.



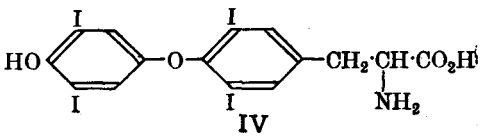
III

Furthermore, when 4-methoxy-4'-methyldiphenyl ether is heated with hydriodic acid, compound II ($C_{13}H_{12}O_2$; see above) is obtained; thus the structure of II is also established.



II

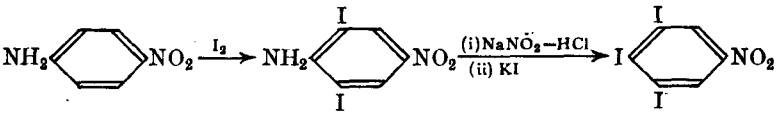
Now when thyroxine is fused with potassium hydroxide, no *p*-hydroxybenzoic acid is obtained; instead, compounds of the pyrogallol type are formed. These facts suggest that two atoms of iodine are adjacent to the hydroxyl group, and that the two remaining iodine atoms are in the other benzene ring. This, together with the analogy with di-iodotyrosine, leads to the suggestion that thyroxine is IV.

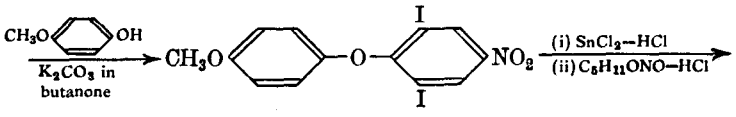


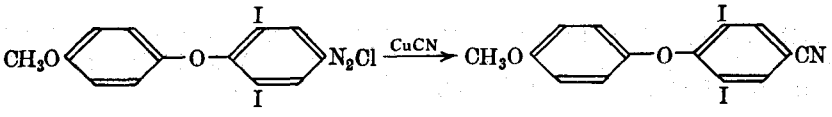
IV

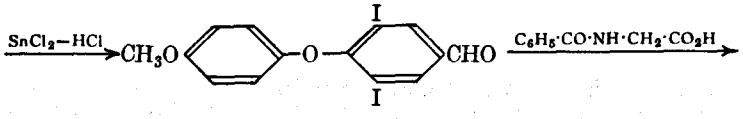
thyroxine

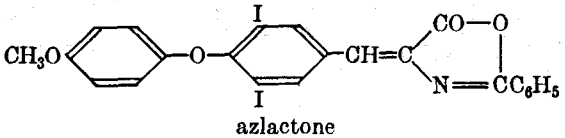
This structure for thyroxine has been confirmed by synthesis (Harington *et al.*, 1927).



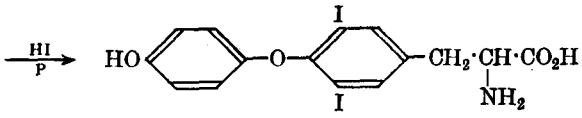


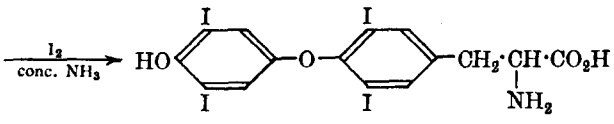






azlactone

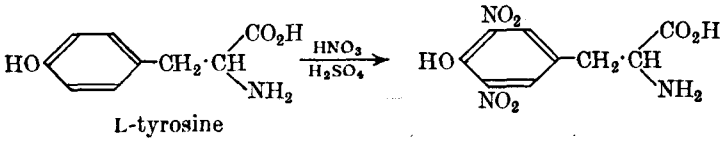




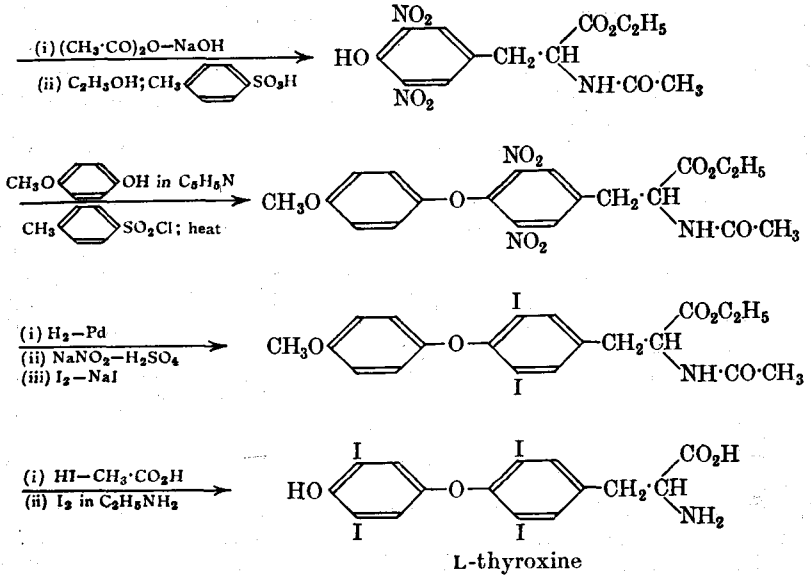
(±)-thyroxine

The racemic modification was resolved *via* the formyl derivative (Harington, 1938).

The synthesis of thyroxine has been improved, *e.g.*, by Hems *et al.* (1949).



L-tyrosine



L-thyroxine

The thyroid gland also contains 3 : 5-di-iodotyrosine, and this compound is believed to be the precursor of thyroxine. Deficiency in thyroxine causes myxœdema.

PROTEINS

§6. General nature of proteins. The name *protein* was introduced by Mulder (1839), who derived it from the Greek word *proteios* (meaning *first*). Proteins are nitrogenous substances which occur in the protoplasm of all animal and plant cells. Their composition varies with the source; an approximate composition may be given as: carbon, 47–50%; hydrogen, 6–7%; oxygen, 24–25%; nitrogen, 16–17%; sulphur, 0·2–0·3%. Other elements may also be present, *e.g.*, phosphorus (nucleoproteins), iron (hæmoglobin).

Proteins are colloids and have no characteristic melting points; some have been obtained in crystalline form. All proteins are optically active (lævorotatory), their activity arising from the fact that they are complex substances built up of amino-acids. It appears likely that all enzymes are proteins (see §12); many hormones are also proteins, *e.g.*, insulin.

Proteins may be coagulated, *i.e.*, precipitated irreversibly, by heat and by strong inorganic acids and bases, etc. When proteins are precipitated irreversibly, they are said to be *denatured*, but the chemical changes that occur in this process are still uncertain. The results of denaturation may be a change in any of the following properties: solubility, molecular shape and size, biological activity, or susceptibility to enzymic reactions. One point that appears to be reasonably certain is that a critical number of hydrogen bonds must be broken before irreversible denaturation can occur. Proteins may be precipitated by ethanol or concentrated solutions of ammonium sulphate or sodium chloride. In this case, the precipitation is reversible, *i.e.*, the precipitated proteins may be redissolved; thus they are not denatured by these reagents. Proteins are also precipitated by the salts of the heavy metals, *e.g.*, mercuric chloride, copper sulphate, etc., and they give many characteristic colour reactions with various reagents, *e.g.*,

(i) *Biuret reaction.* Addition of a very dilute solution of copper sulphate to an alkaline solution of a protein produces a red or violet colour.

(ii) *Millon's reaction.* When a solution of mercuric nitrate containing nitrous acid is added to a protein solution, a white precipitate is formed and slowly turns pink.

(iii) *Xanthoproteic reaction.* Proteins produce a yellow colour when treated with concentrated nitric acid.

Proteins are amphoteric, their behaviour as an anion or a cation depending on the *p*H of the solution. At some definite *p*H, characteristic for each protein, the solution contains equal amounts of anion and cation. In this condition the protein is said to be at its *isoelectric point*, and at this *p*H the protein has its least solubility, *i.e.*, it is most readily precipitated (*cf.* amino-acids, §4 C i). The osmotic pressure and viscosity of the protein solution are also a minimum at the isoelectric point. The amphoteric nature of proteins is due to the presence of a large number of free acidic and basic groups arising from the amino-acid units in the molecule. These groups can be titrated with alkali or acid, and by this means it has been possible to identify acidic and basic groups belonging to the various amino-acid units (see also §11).

The molecular weights of proteins have been determined by means of the ultracentrifuge, osmotic pressure measurements, X-ray diffraction, light scattering effects and by chemical analysis. Chemical methods are based on the estimation of a particular amino-acid, *e.g.*, casein contains cystine; hence the estimation of the percentage of this amino-acid and of sulphur will lead to the evaluation of the molecular weight of casein. The most reliable values of the molecular weights are those obtained by the ultracentrifuge method; the values recorded vary considerably for the individual proteins, ranging from about 40,000 for egg albumin to about 5,000,000 for hæmocyanin.

§7. **Classification of proteins.** Several arbitrary classifications of the proteins are in use. One method is based mainly on physical properties, particularly solubility.

A. Simple proteins. These give only amino-acids on hydrolysis.

(i) *Albumins.* These are soluble in water (and in acids and alkalis), and are coagulated by heat. They are precipitated by saturating their solutions with ammonium sulphate.

Albumins are usually low or deficient in glycine; some albumins are serum albumin, egg albumin and lactalbumin.

(ii) *Globulins.* These are insoluble in water, but are soluble in dilute salt solution and in dilute solutions of strong inorganic acids and alkalis. They are precipitated by half saturating their solutions with ammonium sulphate, and they are coagulated by heat.

Globulins usually contain glycine; some typical globulins are serum globulin, tissue globulin and vegetable globulin.

(iii) *Prolamines.* These are insoluble in water or salt solution, but are soluble in dilute acids and alkalis, and in 70–90 per cent. ethanol.

Prolamines are deficient in lysine, and contain large amounts of proline; some prolamines are zein (from maize), gliadin (from wheat) and hordein (from barley).

(iv) *Glutelins.* These are insoluble in water or dilute salt solution, but are soluble in dilute acids and alkalis; they are coagulated by heat.

Some glutelins are glutenin (from wheat) and oryzenin (from rice).

(v) *Scleroproteins* (*albuminoids*). These are insoluble in water or salt solution, but are soluble in strong acids or alkalis.

Examples: keratin (from hair, hoof), fibroin (from silk); these are not attacked by enzymes.

Submembers of the scleroproteins are:

(*a*) *Collagens* (in skin, tendons and bones); these form gelatin (a water-soluble protein) when boiled with water. Collagens are attacked by pepsin or trypsin.

(*b*) *Elastins* (in tendons and arteries); these are not converted into gelatin, and are attacked slowly by trypsin.

(vi) *Basic proteins.* These are strongly basic, and fall into two groups.

(*a*) *Histones.* These are soluble in water or dilute acids, but are insoluble in dilute ammonia. They are not coagulated by heat, and contain large amounts of histidine and arginine. Histones are the proteins of the nucleic acids, hæmoglobin, etc.

(*b*) *Protamines.* These are more basic than the histones and have a simpler structure. They are soluble in water, dilute acids and dilute ammonia; they are not coagulated by heat, and are precipitated from solution by ethanol. They contain large amounts of arginine, and occur in various nucleic acids.

B. Conjugated proteins are proteins which contain a non-protein group (*i.e.*, a compound not containing amino-acid residues) attached to the protein part. The non-protein group is known as the *prosthetic group*, and it may be separated from the protein part by careful hydrolysis.

(i) *Nucleoproteins.* The prosthetic group is a nucleic acid.

(ii) *Chromoproteins.* These are characterised by the presence of a metal, *e.g.*, iron, magnesium, copper, manganese, cobalt, etc. Chromoproteins may also contain a coloured prosthetic group. Examples: chlorophyll and hæmoglobin.

(iii) *Glycoproteins.* In these the prosthetic group contains a carbohydrate or a derivative of the carbohydrates.

(iv) *Phosphoproteins.* These are conjugated proteins in which the prosthetic group contains phosphoric acid in some form other than in the nucleic acids or in the lipoproteins.

(v) *Lipoproteins.* In these the prosthetic group is lecithin, kephalin, etc.

(vi) *Metalloproteins.* These are heavy metal-protein complexes; all the heavy metals can form complex ions with proteins, *e.g.*, calcium caseinate occurs in blood.

C. Derived proteins are degradation products obtained by the action of acids, alkalis or enzymes on proteins.

Protein ⟶ *Denatured proteins;* insoluble proteins formed by the action of heat, etc., on proteins.

↓

Primary proteoses (*metaproteins*)*;* insoluble in water or dilute salt solution, but are soluble in acids or alkalis. They are precipitated by half-saturation with ammonium sulphate.

↓

Secondary proteoses; soluble in water, not coagulated by heat, and are precipitated by saturation with ammonium sulphate.

↓

Peptones

↓

Polypeptides } These are soluble in water, not coagulated by heat, and are not precipitated by saturation with ammonium sulphate.

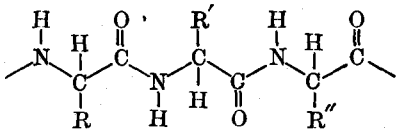
↓

Simple peptides

↓

Amino-acids

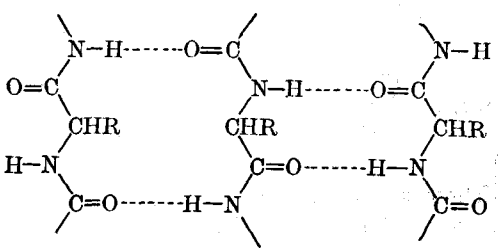
§8. Structure of the proteins. Proteins are hydrolysed by acids, alkalis or by suitable enzymes to a mixture of amino-acids. About twenty-five acids have been definitely isolated; all or only some of these acids may be present in a given protein, and their proportions vary from protein to protein. It appears that, in general, three or four types of amino-acid residues make up the bulk of a given protein molecule, and minor amounts of fifteen or more other acids are also present. Fischer (1902) and Hofmeister (1902) suggested that amino-acids in proteins are joined in a *linear* fashion by *peptide* linkages, *i.e.*, by the —CO·NH— group, the carboxyl group of one amino-acid molecule forming an amide by combination with the amino-group of the next amino-acid molecule, etc. When a relatively small number of amino-acids are linked together (as amides), the resulting molecule is called a *peptide*. When a relatively large number of amino-acid residues are present in the molecule, then that compound is called a *polypeptide*. Proteins are far more complex than the polypeptides. Thus, on this basis, a protein molecule may be represented as a *linear polymer* of amino-acid molecules.



The examination of the infra-red absorption spectra of various synthetic polypeptides has shown the presence of the peptide (*i.e.*, amide) link, and that these links are at positions expected for them. Furthermore, it has been shown that proteins of the keratin type have bands characteristic of the peptide link (Darmon *et al.*, 1947).

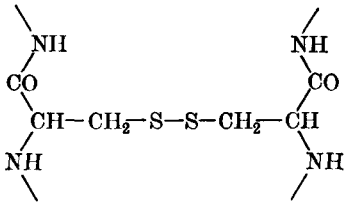
Since some amino-acids contain two amino or two carboxyl groups, it is therefore possible to have *free* amino and carboxyl groups at various positions along the chain, *i.e.*, the group R may contain a free amino or carboxyl group. Since the hydrolysis of certain proteins leads to the formation of ammonia, it has been concluded that in addition to free amino and carboxyl groups, there are also some carbonamide groups, $—CONH_2$. X-ray analysis has confirmed the existence of these polypeptide chains and has shown that the amide group is generally planar. Furthermore, these chains are arranged in a three-dimensional lattice, the chains being held together, to a large extent, by hydrogen bonds. Infra-red studies have shown the presence of hydrogen bonding with NH groups and that the configuration of the substituted amide group is *trans* (see above structure).

On the other hand, when the protein contains cystine, the chains are

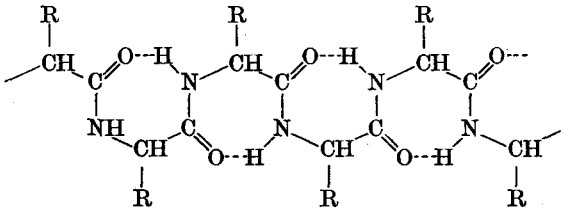


cross-linked via sulphur (*cf.* vulcanisation of rubber, §33a. VIII). The presence of this disulphide linkage has been definitely established, and it has been shown that this link may be broken by oxidation with performic acid

(Hirs, 1956), or by reduction (Sela *et al.*, 1957, 1959). In both cases, the rest of the molecule is unaffected.



Proteins have been found to be of two types, *fibrous* and *globular*. In fibrous proteins the polypeptide chains are extended. In some cases, however, the chains are apparently " coiled ", and these may be extended by the application of a force. The nature of the coiled structure is uncertain, but two configurations have been proposed which agree reasonably well with information obtained from infra-red spectra, X-ray data, bond lengths and bond energies. According to Ambrose *et al.* (1949), the polypeptide chain is folded into a series of seven-membered rings, the folds of the chain being stabilised by hydrogen bonding; in the natural fibre, a number of these folded chains are cross-linked (see above).



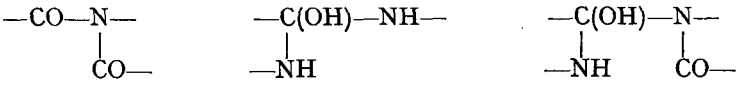
On the other hand, Pauling *et al.* (1951) have proposed a coiled chain in the form of a helix containing either 3·7 or 5·1 acid residues.

The folded (coiled) form of a fibrous protein is known as the α-form, and the extended as the β-form. Elliott *et al.* (1951, 1953) have observed that the frequency of the CO stretching mode in synthetic polypeptides and natural proteins depends on the configuration of the polypeptide chain. Thus this offers a means of distinguishing between the α- and β-forms.

It has been shown that in the solid state many synthetic polypeptides form stable helical structures which correspond closely to the α-helix form. With other synthetic polypeptides, this α-helical configuration appears to be less stable (Elliott *et al.*, 1960; Blout *et al.*, 1960). It has been shown that in poly-β-benzyl-L-aspartate, steric interference between the side-chain and main-chain makes the α-helical configuration fairly unstable (Elliott *et al.*, 1959, 1962). When this compound is heated, it adopts a new helical form, which has been termed the ω-helix. Fraser *et al.* (1962) have prepared another polypeptide which, although not identical in form with that of the aspartate polymer, also is conveniently described as an ω-helix.

The foregoing account of the structure of proteins is based on the long-accepted hypothesis that the peptide structure is universally valid. Recently, however, evidence has been obtained which indicates that this peptide hypothesis is inadequate. Wrinch (1957, 1960), from her examination of small peptides, showed that various observations that are anomalous in the peptide system can be accounted for by a hypothesis consisting of two postulates: (i) the amino-acid residues of peptides are united not only

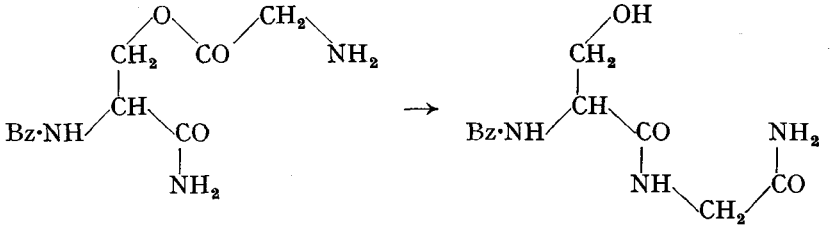
in one-bond peptide grouping, —CO·NH—, but also in the two-bond and three-bond peptide groupings:



(ii) Various reactive side-groups make ring-closures; these in the case of the hydroxy and thiol amino-acids introduce two-bond groupings of the form:

This is known as the *cyclol hypothesis* (Wrinch, 1961).

Another point that complicates the problem of protein structure is the work of Brenner (1958, 1959), who has shown that rearrangements may occur between peptides, *e.g.*, between O·Gly·*N*·Bz·Ser·NH_2 and *N*·Bz·Ser·Gly·NH_2 (see §11 for the meanings of these symbols).



The globular (corpuscular) proteins are more compact than the fibrous proteins, but their shape is not spherical; *e.g.*, X-ray studies have shown that hæmoglobin has a cylindrical shape. The chains in globular proteins are folded many times, and in order to account for certain properties, this folding must follow some definite pattern. An interesting point in connection with globular proteins is that infra-red methods may be used to detect the presence of carboxylate groups in them at the isoelectric point (Ehrlich *et al.*, 1954).

Of the two types of proteins, it is only the globular which have been obtained crystalline; the fibrous proteins lack the characteristics necessary for crystallisation. It appears that all protein crystals grown from solution contain solvent, the removal of which causes the protein to become less crystalline. The solvent has been shown to be interstitial and not " solvent of crystallisation ".

One other point about the nature of these polypeptide chains will now be mentioned briefly. Let us consider a dipeptide composed of two *different* amino-acids, **A** and **B**. These may be combined in two different ways:

$$NH_2\text{—}\mathbf{A}\text{—}CO\text{—}NH\text{—}\mathbf{B}\text{—}CO_2H$$

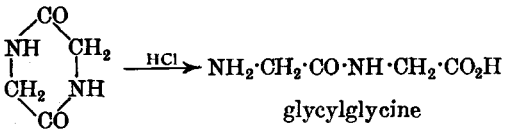
and

$$HO_2C\text{—}\mathbf{A}\text{—}NH\text{—}CO\text{—}\mathbf{B}\text{—}NH_2$$

Three different amino-acids may be combined in six different ways. In general, with n different acids, there will be $n!$ different combinations possible. Had not the naturally occurring amino-acids (excluding glycine) been all of the L-series, the total number of possible combinations would have been very much larger still. It is therefore of great interest to ascertain the " order " in which amino-acids are combined in proteins. Some progress has been made in this direction (see §11).

§9. Synthesis of polypeptides. Various methods have been introduced, *e.g.*,

(i) The partial hydrolysis of a diketopiperazine with hydrochloric acid gives a dipeptide (Fischer, 1901), *e.g.*,



2:5-diketopiperazine

Glycylglycine was the first peptide to be synthesised. The method is very limited in application, since only dipeptides may be prepared. If a "mixed" diketopiperazine is used, then hydrolysis can proceed in two different ways; the nature of the product depends on the hydrolysing agent used.

(ii) The methyl esters of di- and tri-peptides tend to eliminate methanol to form a higher peptide (Fischer, 1901).

$$2NH_2{\cdot}CH_2{\cdot}CO{\cdot}NH{\cdot}CH_2{\cdot}CO_2CH_3 \longrightarrow NH_2{\cdot}CH_2{\cdot}CO{\cdot}(NH{\cdot}CH_2{\cdot}CO)_2{\cdot}NH{\cdot}CH_2{\cdot}CO_2CH_3$$

By means of this reaction, Frankel *et al.* (1942) prepared polypeptides containing up to 110 glycyl units.

(iii) When two different amino-acids are joined to form a dipeptide, two possibilities occur (*cf.* §8); thus, if glycine and alanine are linked together, the two possibilities are:

$$NH_2{\cdot}CH_2{\cdot}CO{\cdot}NH{\cdot}CH(CH_3){\cdot}CO_2H$$

and

$$NH_2{\cdot}CH(CH_3){\cdot}CO{\cdot}NH{\cdot}CH_2{\cdot}CO_2H$$

In order to condense the two amino-acids in a known manner, Fischer (1901, 1903) "blocked" the amino-group of one molecule by first reacting that compound with ethyl chloroformate; thus:

$$C_2H_5O_2C{\cdot}CH_2{\cdot}NH_2 + Cl{\cdot}CO_2C_2H_5 \longrightarrow C_2H_5O_2C{\cdot}NH{\cdot}CH_2{\cdot}CO_2C_2H_5$$

$$\xrightarrow[\text{heat}]{CH_3{\cdot}CH(NH_2){\cdot}CO_2C_2H_5} C_2H_5O_2C{\cdot}NH{\cdot}CH_2{\cdot}CO{\cdot}NH{\cdot}CH(CH_3){\cdot}CO_2C_2H_5$$

Hence by starting with glycylglycine, the diglycylalanine derivative may be prepared. The difficulty with this method, however, is that it is not possible to remove the *N*-carbethoxyl group by hydrolysis without also hydrolysing the peptide link. This difficulty was overcome by Fischer (1915) by using *p*-toluenesulphonyl chloride as the "blocking agent" instead of ethyl chloroformate; the former group can be removed (as the thiophenol) by warming with hydriodic acid, without hydrolysis of the peptide link, *e.g.*,

$$CH_3{\cdot}C_6H_4{\cdot}SO_2Cl + NH_2{\cdot}CH_2{\cdot}CO_2C_2H_5 \longrightarrow$$

$$CH_3{\cdot}C_6H_4{\cdot}SO_2{\cdot}NH{\cdot}CH_2{\cdot}CO_2C_2H_5 \xrightarrow[\text{heat}]{NH_2{\cdot}CH_2{\cdot}CO_2C_2H_5}$$

$$CH_3{\cdot}C_6H_4{\cdot}SO_2{\cdot}NH{\cdot}CH_2{\cdot}CO{\cdot}NH{\cdot}CH_2{\cdot}CO_2C_2H_5 \xrightarrow[50\text{–}60°]{HI}$$

$$CH_3{\cdot}C_6H_4{\cdot}SH + NH_2{\cdot}CH_2{\cdot}CO{\cdot}NH{\cdot}CH_2{\cdot}CO_2H$$

Bergmann (1932) found that benzyl chloroformate is a very good blocking agent, and its application is much wider than that of *p*-toluenesulphonyl chloride. Benzyl chloroformate is readily prepared by the action of carbonyl

chloride on benzyl alcohol in toluene solution (benzyl chloroformate is also known as carbobenzyloxy or benzyloxycarbonyl chloride):

$$C_6H_5{\cdot}CH_2OH + COCl_2 \longrightarrow C_6H_5{\cdot}CH_2O{\cdot}COCl + HCl$$

The procedure is then as follows:

$$C_6H_5{\cdot}CH_2O{\cdot}COCl + R{\cdot}CH(NH_2){\cdot}CO_2H \longrightarrow$$

$$C_6H_5{\cdot}CH_2O{\cdot}CO{\cdot}NH{\cdot}CHR{\cdot}CO_2H \xrightarrow{PCl_5}$$

$$C_6H_5{\cdot}CH_2O{\cdot}CO{\cdot}NH{\cdot}CHR{\cdot}COCl \xrightarrow[NaOH]{R'{\cdot}CH(NH_2){\cdot}CO_2H}$$

$$C_6H_5{\cdot}CH_2O{\cdot}CO{\cdot}NH{\cdot}CHR{\cdot}CO{\cdot}NH{\cdot}CHR'{\cdot}CO_2H \xrightarrow{H_2—Pd}$$

$$C_6H_5{\cdot}CH_3 + CO_2 + NH_2{\cdot}CHR{\cdot}CO{\cdot}NH{\cdot}CHR'{\cdot}CO_2H$$

If the amino-acid contains sulphur, then catalytic reduction cannot be used, since the sulphur poisons the catalyst; the removal of the blocking group, however, may be successfully accomplished by means of phosphonium iodide or sodium in liquid ammonia.

A more recent and convenient method of removing the benzyloxycarbonyl group is to treat the derivative with hydrogen bromide in acetic acid or nitromethane (Ben-Ishai *et al.*, 1952; Anderson *et al.*, 1952):

$$C_6H_5{\cdot}CH_2O{\cdot}CO{\cdot}NH{\cdot}CHR{\cdot}CO{\cdot}NH{\cdot}CHR'{\cdot}CO_2H \xrightarrow{HBr}$$

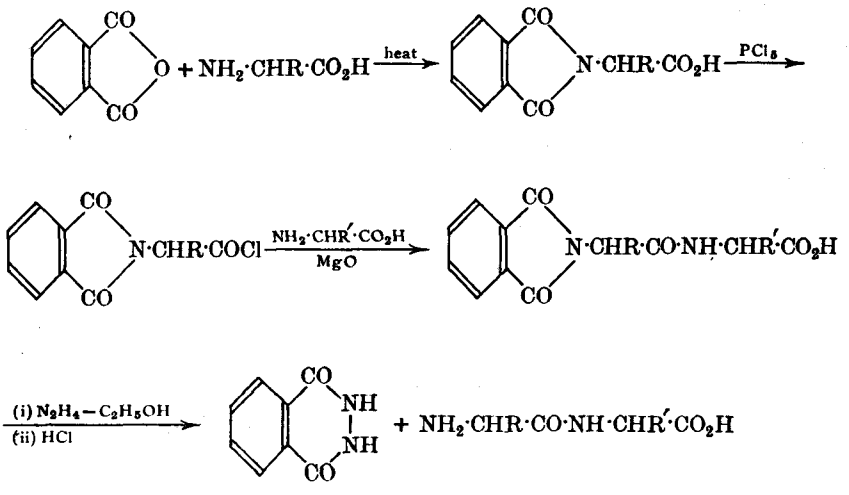
$$C_6H_5{\cdot}CH_2Br + CO_2 + \overset{-}{Br}\overset{+}{N}H_3{\cdot}CHR{\cdot}CO{\cdot}NH{\cdot}CHR'{\cdot}CO_2H$$

According to Weygand *et al.* (1959), boiling trifluoroacetic acid removes an *N*-benzyloxycarbonyl group without splitting peptide bonds or causing racemisation.

Weisblat *et al.* (1953) have also shown that the *p*-toluenesulphonyl group can be removed by means of hydrogen bromide in acetic acid containing phenol.

Stevens *et al.* (1950) have used allyl chloroformate instead of benzyl chloroformate, and then removed the carboallyloxy-group by means of sodium in liquid ammonia.

Sheehan *et al.* (1949) have used the following method for blocking the amino-group of one amino-acid residue (*cf.* Gabriel's phthalimide synthesis, §2 *ib.*).



(iv) Polypeptides may be synthesised by combining an α-halogenoacid chloride with an amino-acid ester and then proceeding as follows (Fischer, 1903).

$$ClCH_2{\cdot}COCl + NH_2{\cdot}CH_2{\cdot}CO_2C_2H_5 \longrightarrow$$

$$ClCH_2{\cdot}CO{\cdot}NH{\cdot}CH_2{\cdot}CO_2C_2H_5 \xrightarrow[NaOH]{\text{dilute}}$$

$$ClCH_2{\cdot}CO{\cdot}NH{\cdot}CH_2{\cdot}CO_2H \xrightarrow{PCl_5}$$

$$ClCH_2{\cdot}CO{\cdot}NH{\cdot}CH_2{\cdot}COCl \xrightarrow{CH_3{\cdot}CH(NH_2){\cdot}CO_2H}$$

$$ClCH_2{\cdot}CO{\cdot}NH{\cdot}CH_2{\cdot}CO{\cdot}NH{\cdot}CH(CH_3){\cdot}CO_2H \xrightarrow{NH_3}$$

$$NH_2{\cdot}CH_2{\cdot}CO{\cdot}NH{\cdot}CH_2{\cdot}CO{\cdot}NH{\cdot}CH(CH_3){\cdot}CO_2H$$

glycylglycylalanine

(v) A variation of the previous method is to convert an amino-acid into its corresponding acid chloride by means of phosphorus pentachloride in acetyl chloride, and then to treat the acid chloride with another molecule of an amino-acid (Fischer, 1907). In the formation of the acid chloride, hydrogen chloride is also produced, and this combines with the amino-group to form the group $Cl^-\{H_3\overset{+}{N}{\cdot}CHR{\cdot}$, which is *not* acetylated by the acetyl chloride present; *e.g.*,

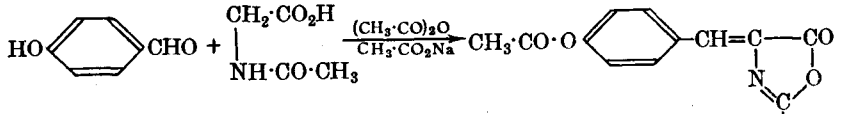
$$NH_2{\cdot}CH_2{\cdot}CO_2H \xrightarrow[CH_3{\cdot}COCl]{PCl_5} NH_2{\cdot}CH_2{\cdot}COCl \xrightarrow{NH_2{\cdot}CH_2{\cdot}CO_2H}$$

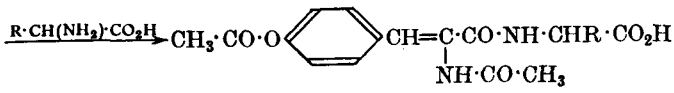
$$NH_2{\cdot}CH_2{\cdot}CO{\cdot}NH{\cdot}CH_2{\cdot}CO_2H$$

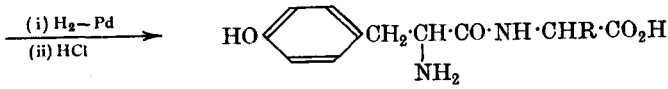
glycylglycine

By this means Fischer (1907) succeeded in synthesising an octadecapeptide (of molecular weight 1213), and Abderhalden (1916) synthesised a nonadecapeptide (of molecular weight 1326).

(vi) The above methods involving the intermediate formation of an acid chloride cannot be applied to hydroxyamino- and di-amino-acids, since these acids react with phosphorus pentachloride in a complicated fashion and do not give the desired halogen compounds. In such cases Bergmann (1926) successfully applied the azlactone synthesis, *e.g.*,







On the other hand, Beyerman *et al.* (1961) have used the *t*-butoxy group to protect the hydroxyl group in the synthesis of peptides containing hydroxy-amino-acids. This group is removed readily by acid without fission of the

peptide bond, and the optical activity is completely maintained during the process. The *t*-butoxy group is conveniently introduced by the acid-catalysed addition of *iso*butene to the hydroxy group of the *N*-acylated hydroxyamino-acid.

(vii) A very recent method of building up peptides is that of Schwyzer *et al.* (1955); this method involves the use of chloroacetonitrile as follows:

$$NH_2{\cdot}CHR{\cdot}CO_2Na + CH_2Cl{\cdot}CN \longrightarrow$$
$$NaCl + NH_2{\cdot}CHR{\cdot}CO_2CH_2{\cdot}CN \xrightarrow{NH_2{\cdot}CHR'{\cdot}CO_2H} NH_2{\cdot}CHR{\cdot}CO{\cdot}NH{\cdot}CHR'{\cdot}CO_2H$$

As we have seen (§4), all the amino-acids except glycine contain at least one asymmetric carbon atom. Furthermore, the α-acylamino-acids are readily racemised, and hence a very important point about the syntheses described above is that racemisation will occur during the syntheses. The actual extent of racemisation depends on the nature of the acyl group and the type of condensation used. According to Boissonas *et al.* (1955), the benzyloxycarbonyl group gives very resistant derivatives (to racemisation) and is therefore the best one to use.

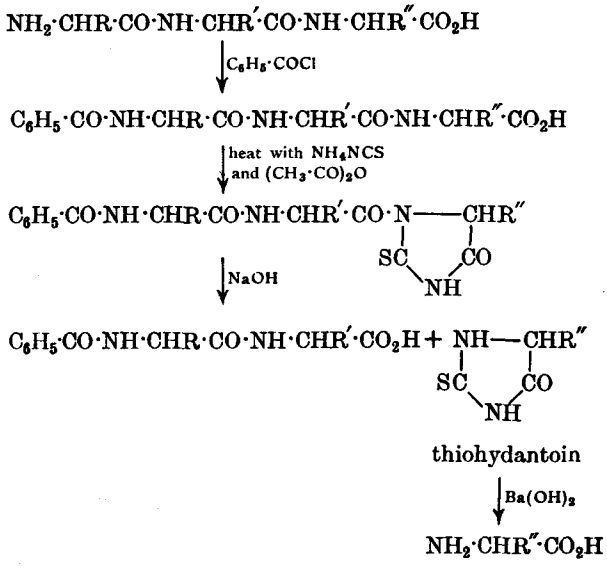
§10. Properties of the polypeptides. The polypeptides are solids which usually decompose when heated to 200–300°. They are soluble in water, but are insoluble in ethanol, and have a bitter taste similar to that of the proteins. They are hydrolysed by acids, alkalis and enzymes, and they very closely resemble the polypeptides actually obtained by the partial hydrolysis of proteins. Polypeptides (synthetic) also give the biuret test. Many peptides have been found as the products of metabolism of micro-organisms.

§11. Degradation of the polypeptides. It has already been pointed out that a necessary requirement for the elucidation of the structure of proteins is a knowledge of the " order " of the amino-acid residues in the molecule (§8). Chemical methods have been introduced whereby the *terminal* amino-acid residue of a polypeptide may be removed in a stepwise fashion. Consideration of the following structure of a polypeptide shows that the two ends of the molecule are not alike; the end on the left-hand side is known as the " amino-end ", and that on the right-hand side as the " carboxyl-end "; the former is said to be *N*-terminal and the latter *C*-terminal.

$$\underset{\text{amino-end}}{NH_2{\cdot}CHR{\cdot}CO{\cdot}NH{\cdot}CHR'{\cdot}CO{\cdot}NH{\cdot}CHR''{\cdot}CO{\cdot}NH \ldots CO{\cdot}NH{\cdot}CHR'''{\cdot}CO_2H}_{\text{carboxyl-end}}$$

Methods have been introduced for degrading either the carboxyl-end or the amino-end of the polypeptide chain, *e.g.*,

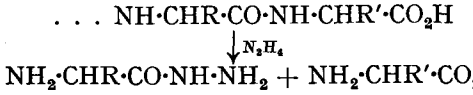
Carboxyl-end degradation. The following method is due to Schlack and Kumpf (1926).



Thus the terminal amino-acid can be identified, and the process can now be repeated on the degraded peptide.

Reduction of proteins with lithium aluminium hydride (or lithium borohydride) converts the free terminal carboxyl group to a primary alcoholic group (*cf.* §4b). Hydrolysis produces an amino-alcohol, which is then identified.

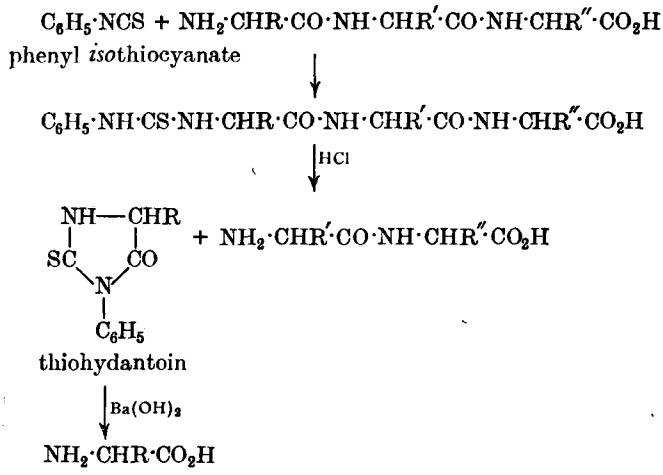
Hydrazinolysis is also used (Akabori *et al.*, 1956). Treatment of a protein with hydrazine converts all amino-acids, except the C-terminal one, into hydrazides.



Treatment of the product with benzaldehyde converts the hydrazides into hydrazones. The terminal amino-acid, which is unaffected by this treatment, is converted into its " DNP " derivative (see below).

Another method makes use of the enzyme carboxypeptidase. This enzyme attacks proteins at the end which contains the free carboxyl group. Thus the chain is gradually degraded.

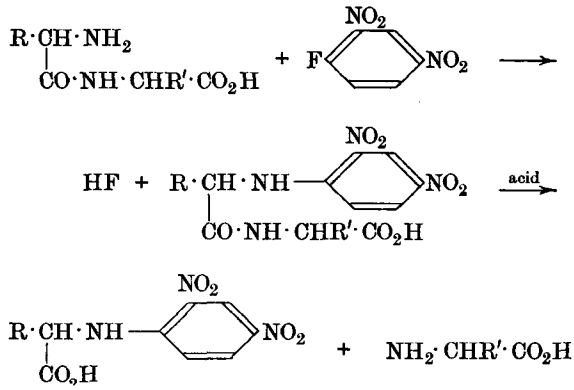
Amino-end degradation. The following method is due to Edman (1950) (see overleaf).



Thus the terminal amino-acid can be identified, and the process can now be repeated on the degraded peptide.

More recently, Asai *et al.* (1955) have investigated the infra-red spectra of polypeptides and have shown that certain bands depend largely on the *sequence* of the amino-acids in the chain. These authors have concluded that the crystalline part of silk fibroin contains glycine and alanine residues arranged alternately.

Another interesting point about the structure of polypeptides is the nature of the amino-acids in the chain which have free amino groups (*cf.* §6). Sanger (1945) has developed the "DNP" method for solving this problem. He showed that 1-fluoro-2 : 4-dinitrobenzene reacts readily only with amino groups and forms derivatives which are stable to acids, *e.g.*,



Thus, when a peptide is first treated with the reagent and then the product hydrolysed with acid, a number of amino-acids will be obtained as their dinitrophenyl derivatives (which can be separated by chromatography; *cf.* §3).

The methods described above for determining the sequence are chemical, but enzymic methods have also given a great deal of information on this problem, *e.g.*, trypsin attacks peptide bonds to which an L-arginine or L-lysine residue has contributed the carboxyl group. Enzymic and chemical methods used together have been extremely valuable.

The exact sequence of amino-acid residues has been worked out only for the hormone insulin, the enzyme ribonuclease, and for the unit protein of the tobacco mosaic virus. The arrangement of the acid residues is random, and consequently synthesis is made difficult.

In protein chemistry, to facilitate writing out the amino-acid sequence, the general practice is to use the first three letters of the names of the acids as abbreviations. When the sequence is not known, the abbreviations are enclosed in brackets, but the *N*- and *C*-terminal residues may be differentiated from residues within the chain by H and OH respectively, *e.g.*,

H·Ala·(Gly·Val·Leu)·OH.

ENZYMES

§12. General nature of enzymes. Enzymes are biological catalysts which bring about chemical reactions in living cells. They are produced by the living organism, and are usually present in only very small amounts in the various cells (about 0·01 per cent.). They can also exhibit their activity even when they have been extracted from their source. The enzymes are all organic compounds, and a number of them have been obtained in a crystalline form. Those so far obtained crystalline are proteins and have very high molecular weights. Most enzymes are colourless solids, but some are yellow, blue, green or greenish-brown; most are soluble in water or dilute salt solution. Some enzymes are purely protein in nature, but many contain a prosthetic group (see §7 B) which has a relatively low molecular weight. The prosthetic group of some enzymes is readily separated (*e.g.*, by dialysis) from the protein part and the latter, in this condition, is known as an *apoenzyme*, *e.g.*, peroxidase is composed of hæmatin (prosthetic group; see §2. XIX) linked with the protein (the apoenzyme). The prosthetic group is often referred to as the co-enzyme (when dealing with enzymes); both parts must be present for the "enzyme" to act. The conjugated protein, *i.e.*, apoenzyme + prosthetic group, has been designated as the *holoenzyme*.

§13. Nomenclature. The systematic method of naming enzymes is to add the suffix *ase* to the name of the *substrate*, *i.e.*, the substance being acted upon, *e.g.*, esterase acts on esters, amylase on starch (amylum), protease on proteins, urease on urea, etc. Some enzymes, however, have retained their trivial names, *e.g.*, emulsin, pepsin, trypsin, etc. Names are also used for *particular* enzymes, *e.g.*, urease, amylase, or as *general* names for *groups* of enzymes, *e.g.*, esterases, proteases, etc. Enzymes of various species are quite often similar, and the reactions catalysed by them are identical. Even so, it does not necessarily follow that these enzymes are identical chemically, *e.g.*, amylases from different sources have different *p*H optima (see below).

§14. Classification of enzymes. Enzymes are usually classified on the type of reaction which they catalyse. There are two main groups:

(i) *Hydrolytic enzymes.* These bring about hydrolysis, *e.g.*, proteases (proteins), lipases (esters), carbohydrases (carbohydrates), etc.

(ii) *Oxidative enzymes.* Most oxidative enzymes function by transferring hydrogen from the substrate (or a modified form, *e.g.*, a hydrated form) to themselves, *i.e.*, they behave as hydrogen acceptors. These enzymes are known as *dehydrogenases*. There are also a few enzymes which oxidise the substrate directly with molecular oxygen; these are known as *oxidases*, *e.g.*, *ascorbic acid oxidase* catalyses the oxidation of ascorbic acid to dehydroascorbic acid by molecular oxygen (*cf.* §11. VII).

Some other types of enzymes are isomerising enzymes, transferring enzymes (*e.g.*, transaminases catalyse the transfer of an amino group of an

amino-acid to a keto group of a keto-acid), and " splitting enzymes " (*e.g.*, decarboxylases catalyse decarboxylation).

§15. Conditions for enzyme action. A number of factors influence enzyme activity: the concentration of the enzyme, the concentration of the substrate, the *p*H of the solution and the temperature. The optimum conditions for a particular enzyme must be found experimentally. The optimum *p*H varies considerably for individual enzymes, and for a given enzyme, with the nature of the substrate. The optimum temperature for animal enzymes is usually between 40° and 50°, and that for plant enzymes 50° and 60°. Most enzymes are irreversibly destroyed when heated above 70–80°.

Many enzymes have been shown to be reversible in their action, *i.e.*, they can both degrade and synthesise. The optimum conditions, however, for degradation are very often totally different from those for synthesis. Furthermore, it does not follow that synthesis in the organism is effected by the same enzyme which produces degradation, *e.g.*, urea is hydrolysed by urease in plants, but is formed in animals by the action of arginase on the amino-acid arginine.

§16. Specificity of enzyme action. One of the most characteristic properties of enzymes is their specificity of action. This specificity may be manifested in one of three ways:

(i) Specificity for a particular reaction or a particular type of reaction, *e.g.*, urease will hydrolyse only urea; esterases hydrolyse only esters. Enzymes may also be specific within a group, *e.g.*, phosphatases (a group of esterases) only hydrolyse esters in which the acid component is phosphoric acid.

(ii) Many enzymes exhibit a *relative specificity, e.g.*, esterases, although hydrolysing all esters, hydrolyse the various esters at *different* speeds; pepsin hydrolyses the peptide link, but is most active for those links in which, among other things, the amino group belongs to an aromatic amino-acid and the carboxyl group is one of a dicarboxylic amino-acid.

(iii) Many enzymes are *stereospecific, e.g.*, maltase hydrolyses α-glycosides but not β-glycosides, whereas emulsin hydrolyses the latter but not the former (*cf.* §3. VII).

It should be noted, however, that a given enzyme can exhibit more than one of the specificities, *e.g.*, esterases, while hydrolysing only esters, may also hydrolyse one enantiomorph (of an optically active ester) more rapidly than the other.

Another point of interest here is that the general type of reaction catalysed by an enzyme depends on the nature of the prosthetic group, and the specificity of the enzyme depends on the nature of the apoenzyme (protein).

§17. Mechanism of enzyme action. According to one view, enzymes initiate the reaction, but according to another view, the reaction catalysed by an enzyme is capable of proceeding at a very slow rate in the absence of the enzyme (*cf.* the theories of catalysis).

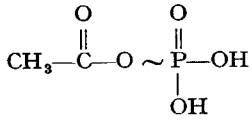
The details of the mechanism of the catalysis effected by enzymes are still not certain. A highly favoured theory is that the enzyme passes through a transition state by combination with its substrate, and then the enzyme is regenerated with the simultaneous formation of the products (*cf.* the transition state, Vol. I). A number of these transition states have been shown to exist from, *e.g.*, spectroscopic evidence; during the reaction the absorption spectrum of the enzyme is altered. It is also believed that the protein part of the enzyme has an " active centre ", and it is this which combines with the substrate. Assuming this be the case, it is now necessary to explain why neither the protein part of the enzyme nor the prosthetic group can act separately, but both must be present (apparently in

combination). The answer to this question has been given in certain cases, *e.g.*, with dehydrogenases it has been suggested that the function of the prosthetic group is to act as a hydrogen acceptor, and that the function of the protein part of the enzyme is to " facilitate " the transfer of the hydrogen from the activated complex. It appears that the usual dehydrogenase action occurs in a number of steps involving different enzymes acting as hydrogen acceptors. Thus each enzyme undergoes reversible reduction and oxidation, and finally the last step is catalysed by cytochrome which is reduced, and this is reoxidised to cytochrome (and water) by molecular oxygen by means of *cytochrome oxidase*. The sequence may therefore be represented:

$$ZH_2 + E_1 \longrightarrow Z + E_1H_2$$
$$E_1H_2 + E_2 \longrightarrow E_1 + E_2H_2 \dashrightarrow E_nH_2 \longrightarrow Cyt.\ H_2$$
$$Cyt.\ H_2 + \tfrac{1}{2}O_2 \xrightarrow[\text{oxidase}]{\text{cytochrome}} Cyt. + H_2O$$

Most of the dehydrogenases contain a prosthetic group (which is the hydrogen acceptor). Thus a number of vitamins (§1. XVII) function as part of prosthetic groups, *e.g.*, pyridino-enzymes (pyridine nucleus), flavo-enzymes (riboflavin), etc. On the other hand, cytochrome and catalase are hæm-containing enzymes, and ascorbic acid oxidase and phenolase are copper protein enzymes.

When small molecules are converted into large molecules containing more energy than the units from which they were built, then energy must be supplied to bring about these syntheses. Enzymes are involved in these syntheses, and it is believed that certain organic compounds contain *energy-rich phosphate bonds*, and when dephosphorylation occurs energy is liberated, *e.g.*, acetyl phosphate contains such a bond (the symbol $\sim$ is used to represent an energy-rich bond):



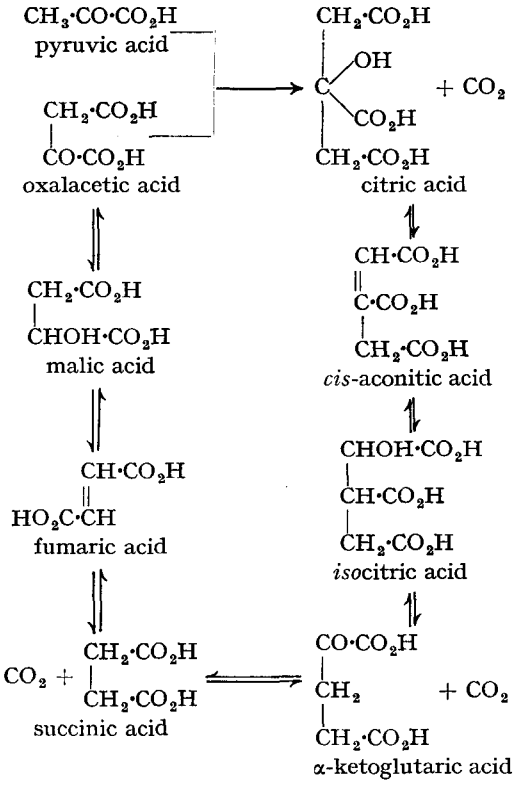
Many enzymes are inactive unless an *activator* is present. The inactive enzyme is known as a *zymogen*, and the activator as a *kinase* (if this is inorganic), *e.g.*, trypsinogen (the zymogen) together with enterokinase (the kinase) forms the enzyme trypsin. Some activators may be metallic or non-metallic, *e.g.*, salivary amylase requires chloride ions for activity. Activators, however, are not co-enzymes. Originally, co-enzymes were understood to include a small number of *organic* compounds of relatively low molecular weight which are required in catalytic amounts in enzyme reactions; the co-enzymes have no enzymic properties of their own. This description of a co-enzyme, however, is now losing this " definition "; most of the metallo-porphyrin catalysts [*i.e.*, the so-called prosthetic groups (§12)] are covered by the foregoing definition. On the other hand, nucleotide co-enzymes are only catalytic in enzyme reactions in which they can be regenerated continuously. From this point of view, it would seem that a co-enzyme behaves as a substrate for the " true " enzyme (*cf.* dehydrogenases above).

Many substances may behave as *inhibitors*, *i.e.*, in their presence the enzyme fails to act; *e.g.*, saccharase is inactivated by copper ions (*cf.* " poisons " in catalysis). Sometimes purely physical means may inactivate an enzyme, *e.g.*, crystalline pepsin is inactivated by sound waves with a frequency of 9 kilocycles per second.

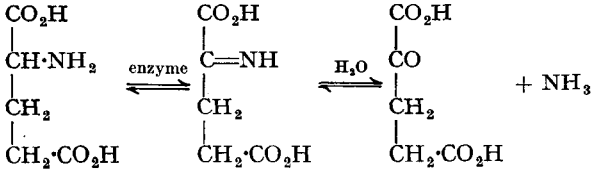
§18. Biosynthesis of amino-acids and proteins. First let us consider the *Krebs cycle* (1937). This is also known as the *citric acid cycle* and is the scheme proposed for the biological oxidation of hexoses to carbon dioxide and water. The first step is the conversion of a hexose molecule into two molecules of pyruvic acid; this occurs *via* the formation of phosphoglyceraldehyde (*cf.* §23a. VII). The pyruvic acid combines with carbon dioxide to form oxalacetic acid:

$$CH_3{\cdot}CO{\cdot}CO_2H + CO_2 \rightleftharpoons CO_2H{\cdot}CH_2{\cdot}CO{\cdot}CO_2H$$

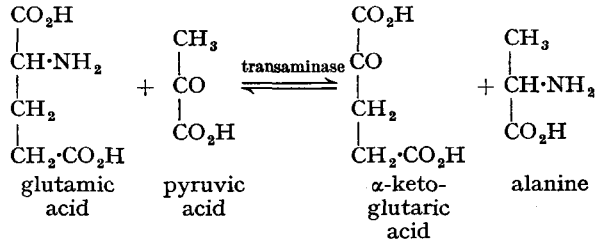
The Krebs cycle may then be written as follows (the various enzymes involved and mechanisms are not shown):



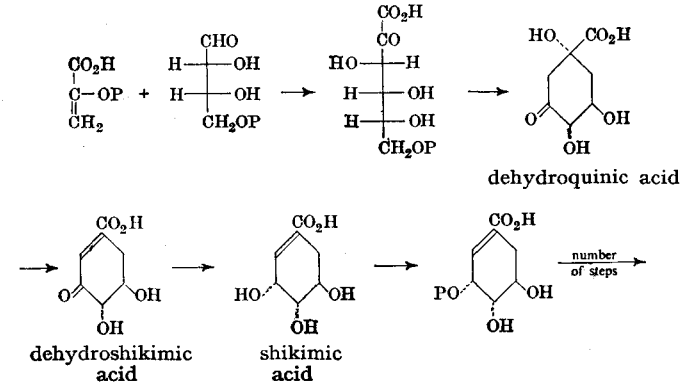
Amino-acids can be deaminated to keto-acids, and in addition to the general (amino-acid) dehydrogenases, there is a specific *glycine dehydrogenase* and a specific *glutamic dehydrogenase*. The case of glutamic acid is extremely important, since there is much evidence to show that this acid plays a vital part in the metabolism of amino-acids. Furthermore, it appears that the conversion of glutamic acid into α-ketoglutaric acid is the only *reversible* reaction in the oxidative deamination of amino-acids.



Keto-acids produced by deamination of amino-acids may undergo further transformations, one being their conversion into amino-acids. This, however, occurs by the process of transamination under the influence of *transaminases, e.g.,*



We have already seen (§32a. VIII) how various keto-acids could be synthesised in the organism. Thus, with the formation of α-ketoglutaric acid from the break-down of carbohydrates, its *direct* amination to glutamic acid, and the latter now capable of aminating other keto-acids by transamination, the cycle of events is set up for the biosynthesis of amino-acids in general. A point to be noted in this connection is that some amino-acids are essential (§1), *e.g.*, man cannot synthesise the benzene ring. Since, however, plants and bacteria synthesise aromatic compounds, a great deal of work has been carried out to elucidate the possible pathways. Two distinct routes have been recognised: (i) from acetate; (ii) from carbohydrates. The latter is believed to be the more important, and Davis *et al.* (1955, 1958), from their work with bacteria, have proposed the following route for the biosynthesis of phenylalanine and tyrosine; the two starting materials are phosphoenol pyruvate and D-erythrose 4-phosphate (P = ortho-phosphate residue):



3-Deoxy-D-arabinoheptulosonic acid has been isolated (Srinivasan *et al.*, 1959); in one stage (labelled *no. of steps*), the nature of the intermediates is not certain.

The shikimic acid pathway is also believed to operate in higher plants (Higuchi, 1958); some alkaloids are believed to be products of this pathway (see §28. XIV). Flavonoids are believed to be derived from both the acetate and shikimic acid pathways (see §14b. XV).

A very interesting problem related to the biosynthesis of amino-acids is the work of Miller (1953, 1955). This author subjected a mixture of methane, ammonia, hydrogen and water vapour (which possibly made up the atmosphere of the Earth in its early stages) to spark and silent discharges. Analysis of the gases showed that the initial gases were present and, in addition, carbon monoxide, carbon dioxide and nitrogen. The solid product was analysed by means of paper chromatography, and the following amino-acids were identified: glycine, sarcosine (*N*-methylglycine), D- and L-alanine, *β*-alanine, D- and L-*α*-amino-*n*-butyric acid and *α*-amino-*iso*butyric acid. Many other amino-acids (unidentified) were also formed, as well as formic, acetic, propionic, glycollic and lactic acids.

Bahadur (1954), on the other hand, has synthesised amino-acids by exposing a solution of paraformaldehyde and potassium nitrate to bright sunlight. Oró *et al.* (1961) have prepared amino-acids from hydrogen cyanide (see also §11a. XVI).

Finally, let us consider the biosynthesis of the proteins from amino-acids. Many workers have concluded that there are no intermediates, *i.e.*, protein synthesis is an " all-at-once " assembly of amino-acids. On the other hand, other workers have concluded that intermediates are formed, but these are so poorly defined or are so transient that they cannot be characterised. Steinberg *et al.* (1951–), using amino-acids labelled with ^{14}C, have shown that their results are compatible with the step-wise mechanism through intermediates. On the other hand, it is generally accepted that nucleic acids serve as matrices for protein synthesis; the D.N.A. (§13. XVI) is considered to be the master pattern, whereas the R.N.A. acts as the working matrix.

READING REFERENCES

Schmidt, *The Chemistry of the Amino-Acids and Proteins*, Thomas (1943, 2nd ed.).
Sahyum (Ed.), *Outline of the Amino-Acids and Proteins*, Reinhold (1948, 2nd ed.).
Gilman (Ed.), *Advanced Organic Chemistry*, Wiley. Vol. II (1943, 2nd ed.). Ch. 14. Natural Amino-Acids.
Rodd (Ed.), *Chemistry of Carbon Compounds*, Elsevier. Vol. 1B (1952). Ch. 22. Proteins.
Greenberg (Ed.), *Amino-Acids and Proteins*, Thomas (1951).
Haurowitz, *Chemistry and Biology of Proteins*, Academic Press (1950).
Springall, *The Structural Chemistry of Proteins*, Butterworth (1954).
Advances in Protein Chemistry, Academic Press (1944–).
Synge, Naturally Occurring Peptides, *Quart. Reviews (Chem. Soc.)*, 1949, **3**, 245.
Khorana, Structural Investigation of Peptides and Proteins, *Quart. Reviews (Chem. Soc.)*, 1952, **6**, 340.
Asimov, Potentialities of Protein Isomerism, *J. Chem. Educ.*, 1954, **31**, 125.
Springall and Law, Peptides: Methods of Synthesis and Terminal-Residue Studies, *Quart. Reviews (Chem. Soc.)*, 1956, **10**, 230.
Progress in Organic Chemistry, Butterworths. Vol. 4 (1958). Degradation and Synthesis of Peptides and Proteins, p. 140.
Advances in Organic Chemistry, Interscience. Vol. I (1960). Thompson, The Selective Degradation of Proteins, p. 149.
Kenner, Recent Progress in the Chemistry of Peptides, *J.C.S.*, **1956**, 3689.
Steinberg *et al.*, Kinetic Aspects of Assembly and Degradation of Proteins, *Science*, 1956, **124**, 389.
Sumner and Somers, *Chemistry and Methods of Enzymes*, Academic Press (1947, 2nd ed.).
Sumner, Enzymes, The Basis of Life, *J. Chem. Educ.*, 1952, **29**, 114.

Avison and Hawkins, The Role of Phosphoric Esters in Biological Reactions, *Quart. Reviews* (*Chem. Soc.*), 1951, **5**, 171.

Klyne (Ed.), *Progress in Stereochemistry*, Butterworth (1954). (i) Ch. 7. The Stereochemistry of Compounds of High Molecular Weight. (ii) Ch. 8. Stereospecificity of Enzyme Reactions.

Newer Methods of Preparative Organic Chemistry, Interscience Publishers (1948). The Use of Biochemical Oxidations and Reductions for Preparative Purposes (pp. 159–196).

Challenger, Biological Methylation, *Quart. Reviews* (*Chem. Soc.*), 1955, **9**, 255.

Miller, Production of Some Organic Compounds under Possible Primitive Earth Conditions, *J. Amer. Chem. Soc.*, 1955, **77**, 2351.

Downes, *The Chemistry of Living Cells*, Longmans, Green (2nd ed., 1963).

Dixon and Webb, *Enzymes*, Longmans, Green (1958).

Baddiley and Buchanan, Recent Developments in the Biochemistry of Nucleotide Coenzymes, *Quart. Reviews* (*Chem. Soc.*), 1958, **12**, 152.

Roth, Ribonucleic Acid and Protein Synthesis, *J. Chem. Educ.*, 1961, **38**, 217.

Neilands and Rogers, Progress in Enzyme Chemistry, *J. Chem. Educ.*, 1962, **39**, 152.

Thompson, Classification and Nomenclature of Enzymes and Coenzymes, *Nature*, 1964, **193**, 1227.

CHAPTER XIV

ALKALOIDS

§1. **Definition of an alkaloid.** Originally the name **alkaloid** (which means alkali-like) was given to *all organic bases isolated from plants.* This definition covers an extraordinary wide variety of compounds, and as the study of " alkaloids " progressed, so the definition changed. Königs (1880) suggested that alkaloids should be defined as naturally occurring organic bases which contain a pyridine ring. This definition, however, embraces only a limited number of compounds, and so the definition was again modified a little later by Ladenburg, who proposed to define alkaloids as natural plant compounds having a basic character and containing at least one nitrogen atom in a heterocyclic ring. Ladenburg's definition excludes any synthetic compounds and any compounds obtained from animal sources. One must admit that even today it is still difficult to define an alkaloid. The term is generally limited to organic bases formed in plants. Not all authors do this, and so they specify those alkaloids obtained from plants as *plant alkaloids* (or *vegetable alkaloids*). On the whole, alkaloids are very poisonous, but are used medicinally in very small quantities. Thus we find that the basic properties, physiological action and plant origin are the main characters which define plant alkaloids. Even so, the class of compounds known as the *purines* (Ch. XVI), which possess the above characters, are not usually included under the heading of alkaloids (some purines are also obtained from animal sources).

It is interesting to note in this connection that Sertürner (1806) isolated a basic compound from opium. Up to that time it was believed that plants produced only acids or neutral compounds.

§2. **Extraction of alkaloids.** In general, the plant is finely powdered and extracted with ethanol. The solvent is then distilled off, and the residue treated with dilute inorganic acids, whereupon the bases are extracted as their soluble salts. The free bases are liberated by the addition of sodium carbonate and extracted with various solvents, *e.g.*, ether, chloroform, etc. The mixtures of bases thus obtained are then separated by various methods into the individual compounds. More recent methods of extraction involve the use of chromatography. Lee (1960) has converted plant alkaloids into their reineckates, dissolved these in acetone, and passed this solution through an ion-exchange column, and thereby obtained the alkaloids in a high state of purity. (Reinecke's solution is $H[Cr(NH_3)_2(SCN)_4]$.)

§3. **General properties.** The alkaloids are usually colourless, crystalline, non-volatile solids which are insoluble in water, but are soluble in ethanol, ether, chloroform, etc. Some alkaloids are liquids which are soluble in water, *e.g.*, coniine and nicotine, and a few are coloured, *e.g.*, berberine is yellow. Most alkaloids have a bitter taste and are optically active. They are generally tertiary nitrogen compounds and contain one or two nitrogen atoms usually in the tertiary state in a ring system; most of the alkaloids also contain oxygen. The optically active alkaloids are very useful for resolving racemic acids. The alkaloids form insoluble precipitates with solutions of phosphotungstic acid, phosphomolybdic acid, picric acid, potassium mercuri-iodide, etc. Many of these precipitates have definite crystalline shapes and so may be used to help in the identification of an alkaloid.

§4. General methods for determining structure.

(**i**) After a pure specimen has been obtained it is subjected to qualitative analysis (invariably the alkaloid contains (carbon), hydrogen and nitrogen; most alkaloids also contain oxygen). This is then followed by quantitative analysis and thus the empirical formula is obtained; determination of the molecular weight finally leads to the molecular formula. If the alkaloid is optically active, its specific rotation is also measured.

(**ii**) When an alkaloid contains oxygen, the functional nature of this element is determined:

(*a*) *Hydroxyl group.* The presence of this group may be ascertained by the action of acetic anhydride, acetyl chloride or benzoyl chloride on the alkaloid (acylation must usually be considered in conjunction with the nature of the nitrogen also present in the molecule; see **iii**). When it has been ascertained that hydroxyl groups are present, then their number is also estimated (by acetylation, etc.). The next problem is to decide whether the hydroxyl group is alcoholic or phenolic. It is phenolic if the alkaloid is soluble in sodium hydroxide and reprecipitated by carbon dioxide; also a coloration with ferric chloride will indicate the presence of a phenolic group. If the compound does not behave as a phenol, then the hydroxyl group may be assumed to be alcoholic, and this assumption may be verified by the action of dehydrating agents (most alkaloids containing an alcoholic group are readily dehydrated by sulphuric acid or phosphorus pentoxide). The behaviour of the compound towards oxidising agents will also disclose the presence of an alcoholic group.

(*b*) *Carboxyl group.* The solubility of the alkaloid in aqueous sodium carbonate or ammonia indicates the presence of a carboxyl group. The formation of esters also shows the presence of a carboxyl group.

(*c*) *Oxo group.* The presence of an oxo group is readily ascertained by the formation of an oxime, semicarbazone and phenylhydrazone.

(*d*) Hydrolysis of the alkaloid and an examination of the products lead to information that the compound is an ester, lactone, amide, lactam or a betaine.

(*e*) The *Zerewitinoff active hydrogen determination* may be applied to the alkaloid (see Vol. I).

(*f*) *Methoxyl group.* The presence of methoxyl groups and their number may be determined by the *Zeisel method.* The alkaloid is heated with concentrated hydriodic acid at its boiling point (126°); the methoxyl groups are thereby converted into methyl iodide, which is then absorbed by ethanolic silver nitrate and the silver iodide is weighed. Only methoxyl groups have been found in natural alkaloids.

(*g*) *Methylenedioxyl group* (—O·CH_2·O—). The presence of this group is indicated by the formation of formaldehyde when the alkaloid is heated with hydrochloric or sulphuric acid.

(**iii**) *The functional nature of the nitrogen.*

(*a*) The general reactions of the alkaloid with acetic anhydride, methyl iodide and nitrous acid often show the nature of the nitrogen.

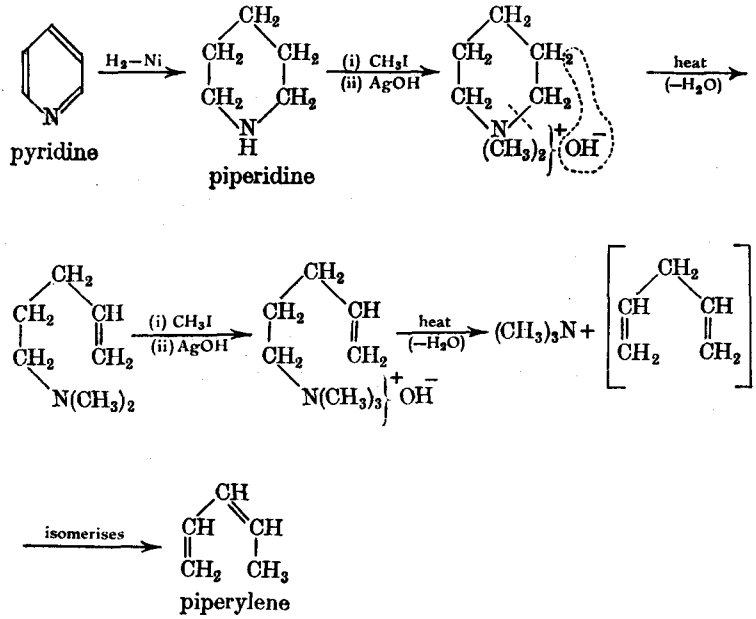
(*b*) Distillation of an alkaloid with aqueous potassium hydroxide usually leads to information regarding the nature and number of alkyl groups attached to nitrogen. The formation (in the volatile products) of methylamine, dimethylamine or trimethylamine indicates respectively the attachment of one, two or three methyl groups to a nitrogen atom; the formation of ammonia shows the presence of an amino group. Only *N*-methyl groups have been shown to be present in alkaloids with one exception, *viz.*, aconitine, which contains an *N*-ethyl group.

(*c*) The presence of *N*-methyl groups and their number may be determined by means of the *Herzig–Meyer method.* When the alkaloid is heated

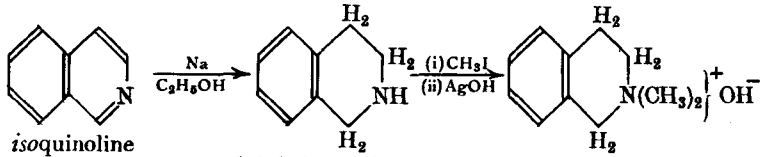
with hydriodic acid at 150–300° under pressure, *N*-methyl groups are converted into methyl iodide (*cf.* the Zeisel method, **ii***f*).

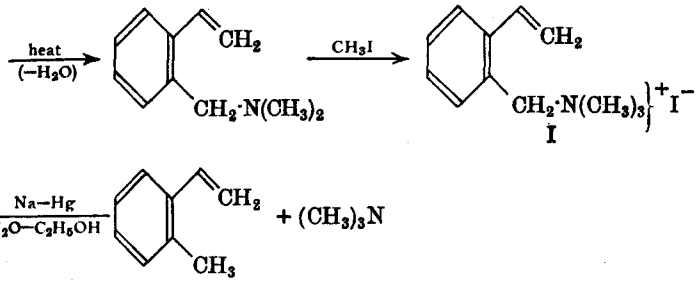
(*d*) The results of hydrolysis will show the presence of an amide, lactam or betaine (*cf.* **ii***d*).

(*e*) *Hofmann's exhaustive methylation method* (1881) is a very important process in alkaloid chemistry, since by its means heterocyclic rings are opened with the elimination of nitrogen, and the nature of the carbon skeleton is thereby obtained. The general procedure is to hydrogenate the heterocyclic ring (if this is unsaturated), then convert this compound to the quaternary methylammonium hydroxide which is then heated. In this last stage a molecule of water is eliminated, a hydrogen atom *in the β-position* with respect to the nitrogen atom combining with the hydroxyl group, and the ring is opened at the nitrogen atom on the *same* side as the β-hydrogen atom eliminated. The process is then repeated on the product; this results in the complete removal of the nitrogen atom from the molecule, leaving an unsaturated hydrocarbon which, in general, isomerises to a *conjugated* diene (see also Vol. I); *e.g.*,



Hofmann's method fails if there is no β-hydrogen atom available for elimination as water; in such cases the Emde modification (1909, 1912) may be used. In this method the quaternary ammonium halide is reduced with sodium amalgam in aqueous ethanol or catalytically hydrogenated, *e.g.*,

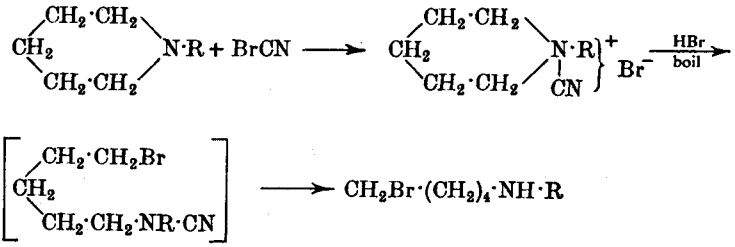




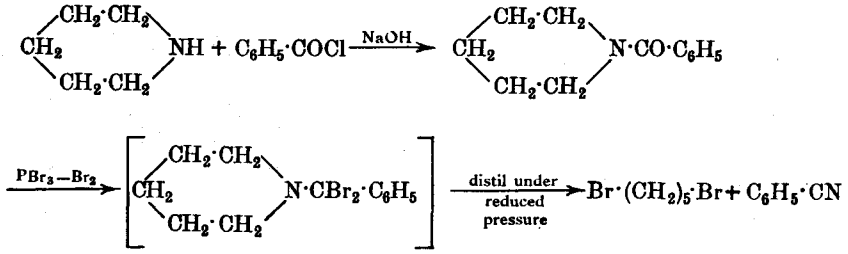
Examination of I shows that β-hydrogen is absent; hence Hofmann's method cannot be used.

Other methods for opening heterocyclic rings containing nitrogen are:

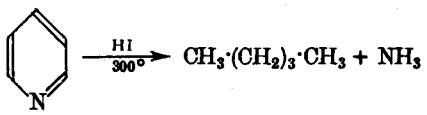
(i) *Von Braun's method* for *tertiary* cyclic amines (see also Vol. I); *e.g.*,



(ii) *Von Braun's method* for *secondary* cyclic amines (see also Vol. I); *e.g.*,



(iii) In a number of cases the ring may be opened by heating with hydriodic acid at 300°, *e.g.*,



(iv) The presence of unsaturation in an alkaloid may be ascertained by the addition of bromine and halogen acids, or by the ability to be hydroxylated with dilute alkaline permanganate. Reduction by means of sodium amalgam, sodium and ethanol, tin and hydrochloric acid, hydriodic acid, etc., also may be used to show the presence of unsaturation. In some cases, reduction may decompose the molecule. This often happens when catalytic reduction is used (ring cleavage occurs), and hence milder methods of reduction are desirable. Two particularly mild reducing reagents are lithium aluminium hydride and sodium borohydride. Sodium in liquid ammonia gives the Emde type of degradations (see **iii**).

(v) *Oxidation.* This is one of the most valuable means of determining

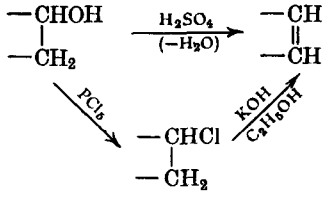
the structure of alkaloids (*cf.* terpenes, §3. VIII). By varying the " strength " of the oxidising agent, it is possible to obtain a variety of products:

(*a*) Mild oxidation is usually effected with hydrogen peroxide, ozone, iodine in ethanolic solution, or alkaline potassium ferricyanide.

(*b*) Moderate oxidation may be carried out by means of acid or alkaline potassium permanganate, or chromium trioxide in acetic acid.

(*c*) Vigorous oxidation is usually effected by potassium dichromate–sulphuric acid, chromium trioxide–sulphuric acid, concentrated nitric acid, or manganese dioxide–sulphuric acid.

This classification is by no means rigid; the " strength " of an oxidising agent depends to some extent on the nature of the compound being oxidised. In those cases where it can be done, better results are sometimes achieved by first dehydrating the compound and then oxidising the unsaturated compound thus obtained; oxidation is readily effected at a double bond.



More recently, mercuric acetate has been used to dehydrogenate certain alkaloids, thereby introducing olefinic bonds.

(**vi**) Fusion of an alkaloid with solid potassium hydroxide often produces relatively simple fragments, the nature of which will give information on the type of nuclei present in the molecule (*cf.* **iii***b*).

(**vii**) *Zinc dust distillation.* This usually gives the same products as (**vi**), except that when the alkaloid contains oxygen the oxygen is removed.

(**viii**) Physical methods are also now being used, in conjunction with chemical methods, to elucidate structure, *e.g.*, infra-red spectra studies are used to identify many functional groups; ultraviolet spectra are used to indicate the likely type of structure present; and X-ray analysis has offered a means of distinguishing between alternative structures that appear to fit equally well the alkaloid in question.

(**ix**) *Synthesis.* The foregoing analytical work will ultimately lead to the proposal of a tentative structure (or structures) for the alkaloid under consideration. The final proof of structure, however, depends on an unambiguous synthesis of the alkaloid.

§5. Classification of the alkaloids. Long before the constitutions of the alkaloids were known, the source of the alkaloid was considered the most important characteristic of the compound. Thus there could not be a rational classification. Even today, with the structures of so many known, the classification of the alkaloids is still somewhat arbitrary owing to the difficulty of classifying into distinct groups. Even so, it is probably most satisfactory (chemically) to classify the alkaloids according to the nature of the nucleus present in the molecule. Members of the following groups are described in this book:

(i) Phenylethylamine group.
(ii) Pyrrolidine group.
(iii) Pyridine group.
(iv) Pyrrolidine-pyridine group.
(v) Quinoline group.
(vi) *iso*Quinoline group.
(vii) Phenanthrene group.

It should be noted that in many cases different alkaloids obtained from the same plant often have similar chemical structures, and so sometimes the source of the alkaloids may indicate chemical similarity.

PHENYLETHYLAMINE GROUP

Many compounds of this group are known, some natural and others synthetic. Their outstanding physiological action is to increase the blood-pressure; hence they are often referred to as the *pressor drugs*.

§6. β-Phenylethylamine. This is the parent substance of this group of alkaloids, and occurs in putrid meat (it is formed by the decarboxylation of phenylalanine, an amino-acid). β-Phenylethylamine may be readily synthesised as follows:

$$C_6H_5{\cdot}CH_2Cl + KCN \longrightarrow C_6H_5{\cdot}CH_2{\cdot}CN \xrightarrow[C_2H_5OH]{Na} C_6H_5{\cdot}CH_2{\cdot}CH_2{\cdot}NH_2$$

β-Phenylethylamine is a colourless liquid, b.p. 197°.

§7. (−)-Ephedrine, m.p. 38·1°. (−)-Ephedrine occurs in the genus *Ephedra*; it is one of the most important drugs in *Ma Huang* (a Chinese drug). Physiologically, its action is similar to that of adrenaline (§12), and it can be taken orally.

The molecular formula of ephedrine is $C_{10}H_{15}ON$, and since on oxidation ephedrine forms benzoic acid, the structure therefore contains a benzene ring with only one side-chain. When treated with nitrous acid, ephedrine forms a nitroso-compound; therefore the compound is a secondary amine. Since ephedrine forms a dibenzoyl derivative, one hydroxyl group must be present (one benzoyl group is accounted for by the imino group). Finally, when heated with hydrochloric acid, ephedrine forms methylamine and propiophenone.

$$C_{10}H_{15}ON \xrightarrow{HCl} CH_3{\cdot}NH_2 + C_6H_5{\cdot}CO{\cdot}CH_2{\cdot}CH_3$$

The formation of these products can be explained if the structure of ephedrine is either I or II.

$$C_6H_5{\cdot}CHOH{\cdot}CH_2{\cdot}CH_2{\cdot}NH{\cdot}CH_3 \qquad \text{I}$$

$$C_6H_5{\cdot}CHOH{\cdot}CH(NH{\cdot}CH_3){\cdot}CH_3 \qquad \text{II}$$

It has been observed, however, that compounds of structure II undergo the *hydramine fission* to form propiophenone when heated with hydrochloric acid. Thus II is more likely than I. This is supported by the fact that when subjected to the Hofmann exhaustive methylation method, ephedrine forms *sym.*-methylphenylethylene oxide, III; this cannot be produced from I, but is to be expected from II.

$$C_6H_5{\cdot}CHOH{\cdot}CH(CH_3){\cdot}NHCH_3 \xrightarrow[(ii)\ AgOH]{(i)\ CH_3I}$$

II

$$C_6H_5{\cdot}CHOH{\cdot}CH(CH_3){\cdot}N(CH_3)_3\}^+OH^- \xrightarrow[(-H_2O)]{heat}$$

$$C_6H_5{\cdot}\underset{\diagdown O \diagup}{CH}{\cdot}CH{\cdot}CH_3 + (CH_3)_3N$$

III

Further support for II is afforded by the following evidence. Structure I contains one asymmetric carbon atom, and so replacement of the hydroxyl

group by hydrogen will result in the formation of an optically inactive compound. Structure II, however, contains two asymmetric carbon atoms, and so the replacement of the hydroxyl group by hydrogen should still give a compound that can be optically active. Experimentally it has been found that when this replacement is effected in (−)-ephedrine, the product, deoxyephedrine, is optically active. Thus II agrees with all the known facts, and this structure has been confirmed by synthesis, *e.g.*, Späth *et al.* (1920):

$$CH_3{\cdot}CH_2{\cdot}CHO \xrightarrow{Br_2} CH_3{\cdot}CHBr{\cdot}CHO \xrightarrow[HBr]{CH_3OH}$$

$$CH_3{\cdot}CHBr{\cdot}CH(OCH_3)(Br) \xrightarrow{C_6H_5MgBr} CH_3{\cdot}CHBr{\cdot}CH(OCH_3)(C_6H_5) \xrightarrow{CH_3{\cdot}NH_2}$$

$$CH_3{\cdot}CH(NH{\cdot}CH_3){\cdot}CH(OCH_3)(C_6H_5) \xrightarrow{HI} C_6H_5{\cdot}CHOH{\cdot}CH(CH_3){\cdot}NH{\cdot}CH_3$$
$$(\pm)\text{-}\psi\text{-ephedrine}$$

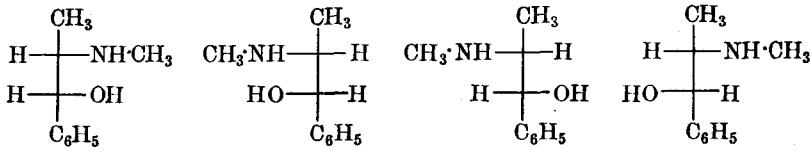
The racemic modification of ψ-ephedrine (see below) was resolved by means of tartaric acid.

(−)-Ephedrine itself has been synthesised by Manske *et al.* (1929) by the catalytic reduction of 1-phenylpropane-1 : 2-dione (benzoylacetyl) in the presence of methylamine in methanol solution.

$$C_6H_5{\cdot}CO{\cdot}CO{\cdot}CH_3 + CH_3{\cdot}NH_2 \longrightarrow C_6H_5{\cdot}CO{\cdot}C(=N{\cdot}CH_3){\cdot}CH_3 \xrightarrow{H_2-Pt}$$
$$C_6H_5{\cdot}CHOH{\cdot}CH(CH_3){\cdot}NH{\cdot}CH_3$$
$$(\pm)\text{-ephedrine}$$

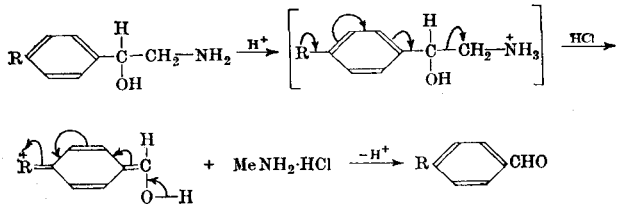
The racemic ephedrine was resolved by means of mandelic acid. Some $(\pm)$-ψ-ephedrine was also obtained in this synthesis.

Since the ephedrine molecule contains two dissimilar asymmetric carbon atoms, four optically active forms (two pairs of enantiomorphs) are theoretically possible. According to Freudenberg (1932), the configurations of ephedrine and ψ-ephedrine are:

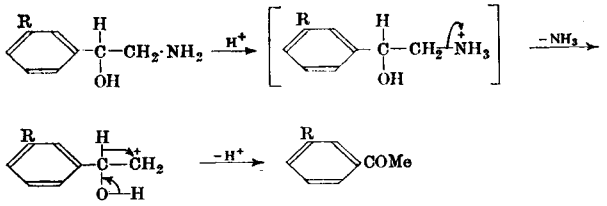


(−)-ephedrine (+)-ephedrine (−)-ψ-ephedrine (+)-ψ-ephedrine

Various mechanisms have been proposed for the hydramine fission. Chatterjee *et al.* (1961) have suggested two different mechanisms according to whether the aryl nucleus contains (i) an electron-releasing group in the *o* and/or *p*-position, *e.g.*, R = OMe, OH, Me:

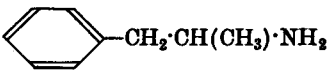


(ii) R in the *m*-position:



Thus hydramine fission gives an aldehyde or a ketone according to the nature and position of groups in the aryl nucleus. With a 4-nitro group the product is 4-nitroacetophenone (yield: very poor).

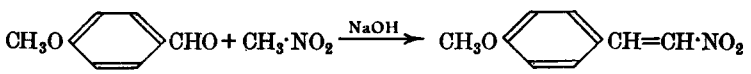
§8. **Benzedrine** (*Amphetamine*) was originally introduced as a substitute for ephedrine, but it is now used in its own right since it apparently produces a feeling of confidence.



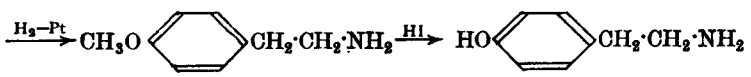
Benzedrine has been synthesised in many ways, *e.g.*, Mingoia (1940):

$$C_6H_5{\cdot}CH_2{\cdot}COCH_3 \xrightarrow[150-190^\circ]{HCONH_2} C_6H_5{\cdot}CH_2{\cdot}CH(CH_3){\cdot}NH{\cdot}CHO \xrightarrow{HCl} C_6H_5{\cdot}CH_2{\cdot}CH(CH_3){\cdot}NH$$

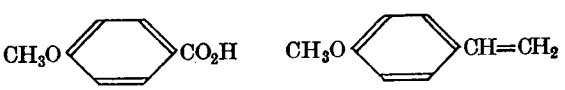
§9. **β-*p*-Hydroxyphenylethylamine** (*tyramine*), m.p. 160°, occurs in ergot, and is produced by the putrefaction of proteins (by the decarboxylation of tyrosine). Tyramine has been synthesised in various ways, *e.g.*,



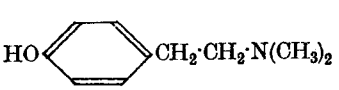
anisaldehyde



§10. **Hordenine** (*β-p*-hydroxyphenylethyldimethylamine, *Anhaline*), m.p. 117–118°, occurs naturally in germinating barley. The molecular formula of hordenine is $C_{10}H_{15}ON$; the routine tests show that hordenine is a tertiary base and that it contains a phenolic group. Since the methylation of hordenine, followed by oxidation (with alkaline permanganate), gives anisic acid, I, it therefore follows that the hydroxyl group is in the *para*-position with respect to the side-chain. Furthermore, since the methylated compound gives *p*-vinylanisole, II, after the Hofmann exhaustive methylation, the structure of hordenine is probably III.

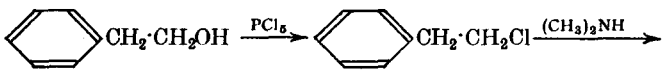


I II

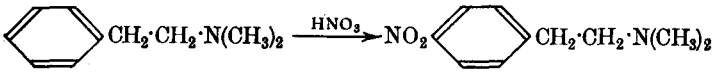


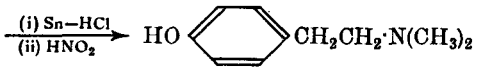
III

This has been confirmed by synthesis, *e.g.*, Barger (1909):

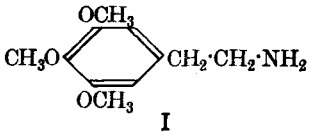


2-phenylethanol



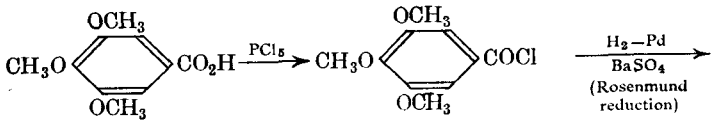


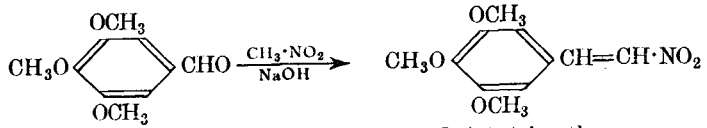
§11. **Mezcaline** (mescaline), $C_{11}H_{17}O_3N$, b.p. 180–180·5°/12 mm., occurs naturally in "mezcal buttons". The routine tests show that mezcaline contains a primary aliphatic amino-group and three methoxyl groups. On oxidation with alkaline permanganate, mezcaline gives 3 : 4 : 5-trimethoxybenzoic acid, and thus the probable structure of mezcaline is I.



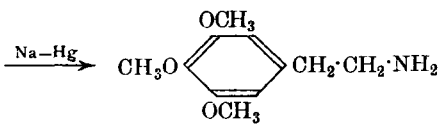
I

This has been confirmed by synthesis (Späth, 1919):



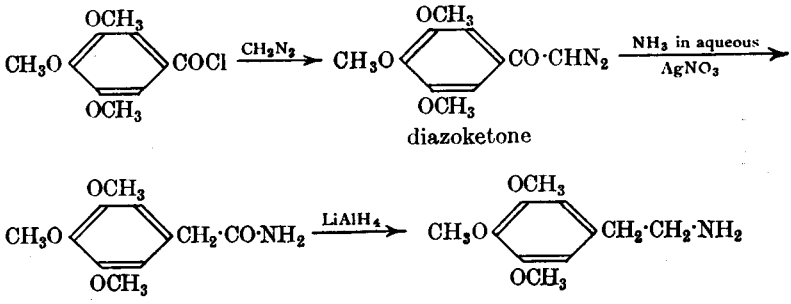


3:4:5-trimethoxy-ω-nitrostyrene



mezcaline

A more recent synthesis of mezcaline is that of Banholzer *et al.* (1952); this makes use of the Arndt–Eistert synthesis.

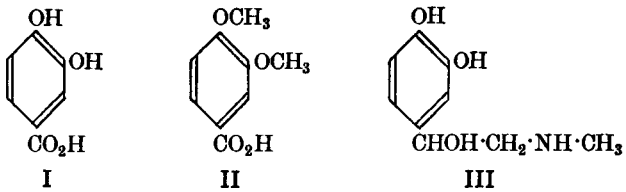


N-Methylmezcaline and *N*-acetylmezcaline also occur naturally in mezcal buttons.

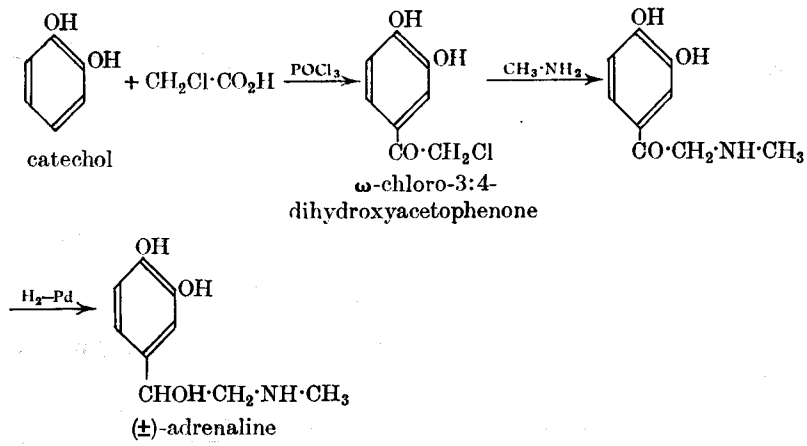
§12. Adrenaline (*Epinephrine*), $C_9H_{13}O_3N$, is a non-steroid hormone. The adrenal medulla is the source of the hormones adrenaline and noradrenaline. Adrenaline was the first hormone to be isolated in a crystalline form (Takamine, 1901; Aldrich, 1901). Adrenaline is active only when given by injection; it raises the blood-pressure, and is used locally to stop hæmorrhage.

Adrenaline is a colourless crystalline solid, m.p. 211°, and dissolves in acids and alkalis (it is insoluble in water); it is also optically active, having a lævorotation.

The phenolic character of adrenaline is indicated by its solubility in sodium hydroxide and its reprecipitation by carbon dioxide. Since it gives a green colour with ferric chloride, this led to the suggestion that adrenaline is a catechol derivative. When boiled with aqueous potassium hydroxide, adrenaline evolves methylamine; thus a methylamino group is probably present. On the other hand, when fused with potassium hydroxide, the product is protocatechuic acid, I (Takamine, 1901); methylation, followed

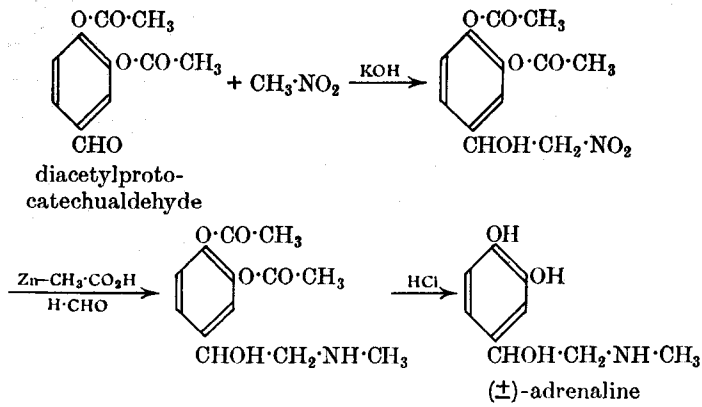


by fusion with potassium hydroxide, gives veratric acid, II, and trimethylamine (Jowett, 1904). The formation of trimethylamine indicates that the nitrogen atom must occur at the *end* of the side-chain. Since adrenaline is optically active, it must contain at least one asymmetric carbon atom. Now adrenaline contains three hydroxyl groups, two of which are phenolic (as shown by the formation of I and II). The third hydroxyl group was shown to be secondary alcoholic by the fact that when adrenaline is treated with benzenesulphonyl chloride, a tribenzenesulphonyl derivative is obtained which, on oxidation, gives a ketone (Friedmann, 1906). To account for the oxidation of adrenaline to the benzoic acid derivative, the —CHOH— group must be attached directly to the nucleus; had it been —CH_2·CHOH·, then a phenylacetic acid derivative would have been obtained. All the foregoing facts are in keeping with structure III for adrenaline, and this has been confirmed by synthesis by Stolz (1904) and Dakin (1905), with improvements by Ott (1926).

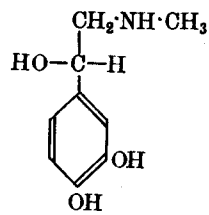


The racemic adrenaline has been resolved by means of (+)-tartaric acid.

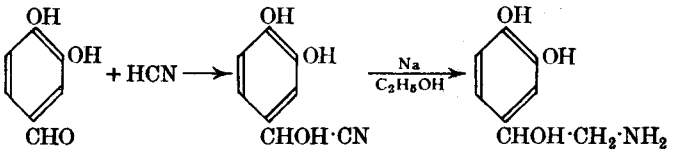
Nagai (1918) has also synthesised adrenaline as follows:



According to Dalgliesh (1953), the configuration of (−)-adrenaline is probably

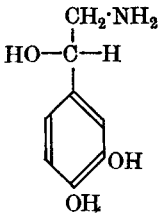


§12a. **Noradrenaline** (*Norepinephrine*), $C_8H_{11}O_3N$, is also present in the adrenal medulla. The natural compound is lævorŏtatory, and this (−)-isomer is the most powerful pressor-compound known. The structure of noradrenaline has been established by analytical work similar to that described for adrenaline, and has been confirmed by various syntheses, *e.g.*,



(±)-noradrenaline

According to Dalgliesh (1953), the configuration of (—)-noradrenaline is

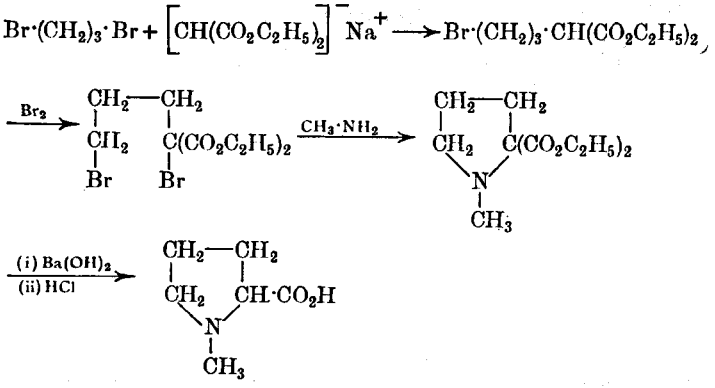


PYRROLIDINE GROUP

§13. Hygrine, $C_8H_{15}ON$, b.p. 193–195°, is one of the coca alkaloids. Its reactions show the presence of a keto group and a tertiary nitrogen atom, and when oxidised with chromic acid, hygrinic acid is formed.

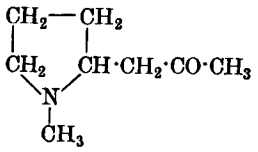
$$C_8H_{15}ON \xrightarrow{[O]} C_6H_{11}O_2N$$
hygrinic acid

Hygrinic acid was first believed to be a piperidinecarboxylic acid, but comparison with the three piperidine acids showed that this was incorrect. When subjected to dry distillation, hygrinic acid gives *N*-methylpyrrolidine; hence hygrinic acid is an *N*-methylpyrrolidinecarboxylic acid. Furthermore, since the decarboxylation occurs very readily, the carboxyl group was assumed to be in the 2-position (by analogy with the α-amino-acids). This structure, 1-methylpyrrolidine-2-carboxylic acid, for hygrinic acid was confirmed by synthesis (Willstätter, 1900).

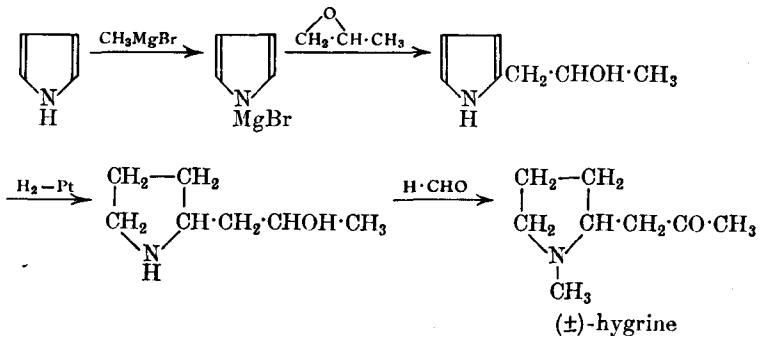


(±)-hygrinic acid

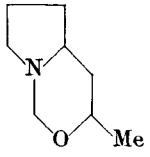
Thus a possible structure for hygrine is



Hess (1913) claimed to have confirmed this structure by synthesis; his synthesis starts with pyrrylmagnesium bromide and propylene oxide to form pyrrylpropanol (note the rearrangement that occurs). This compound is then catalytically hydrogenated and then treated with formaldehyde; the imino nitrogen is methylated and the secondary alcoholic is oxidised to a keto group.



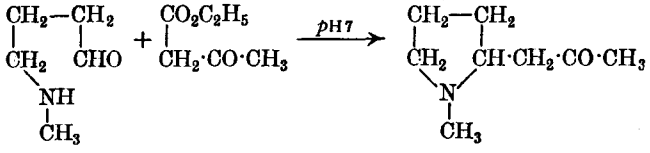
Lukeš *et al.* (1959) have repeated Hess's work and have shown that the



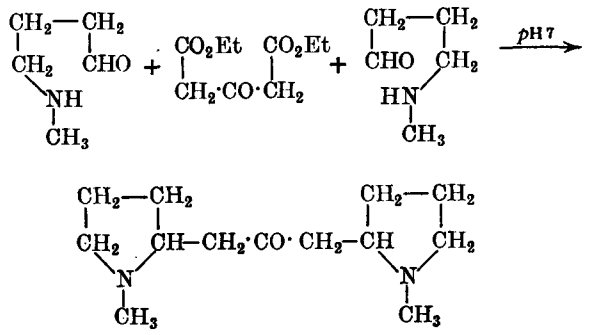
I

product is not hygrine but the tetrahydro-oxazine (I); it is the last stage of Hess's interpretation that has been shown to be incorrect.

Anet *et al.* (1949) have also synthesised (±)-hygrine by condensing γ-methylaminobutyraldehyde with ethyl acetoacetate in a buffered solution at a *p*H of 7 (physiological conditions).

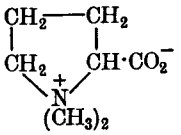


§13a. Cuscohygrine (*Cuskhygrine*), b.p. 169–170°/23 mm., occurs with hygrine. Its structure is established by the following synthesis (Anet *et al.*, 1949); γ-methylaminobutyraldehyde is condensed with acetonedicarboxylic ester:

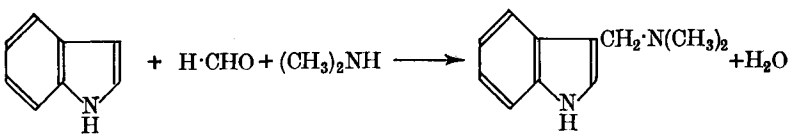


cuscohygrine

§13b. Stachydrine is obtained from the roots of *Stachys tuberifa*, from orange leaves, etc. It is the betaine (§4 C. XIII) of the quaternary ammonium compound of hygrinic acid.

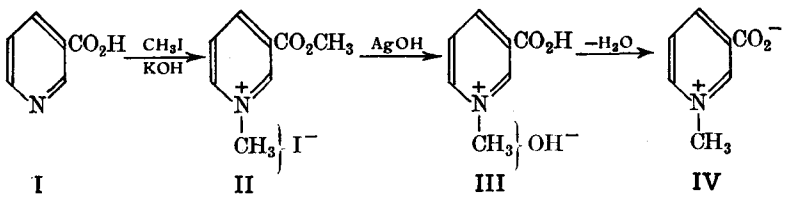


§14. Gramine has been found in barley mutants; it raises the blood-pressure in dogs when administered in small doses. Gramine has been synthesised by allowing indole to stand in an aqueous solution containing formaldehyde and dimethylamine (Snyder *et al.*, 1944).

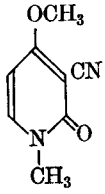


PYRIDINE GROUP

§15. Trigonelline, $C_7H_7O_2N$, m.p. 130°, is widely distributed in plants; the best source is the coffee bean. When boiled with barium hydroxide solution trigonelline produces methylamine; thus the molecule contains an *N*-methylamino group. On the other hand, when heated with hydrochloric acid at 250° under pressure, trigonelline forms methyl chloride and nicotinic acid; this suggests that the alkaloid is the methyl betaine of nicotinic acid. This structure for trigonelline has been confirmed by synthesis (Hantzsch, 1886). When heated with methyl iodide in the presence of potassium hydroxide, nicotinic acid, I, is converted into methyl nicotinate methiodide, II. II, on treatment with " silver hydroxide " solution, forms nicotinic acid methohydroxide, III, which then spontaneously loses a molecule of water to give trigonelline (a betaine), IV.

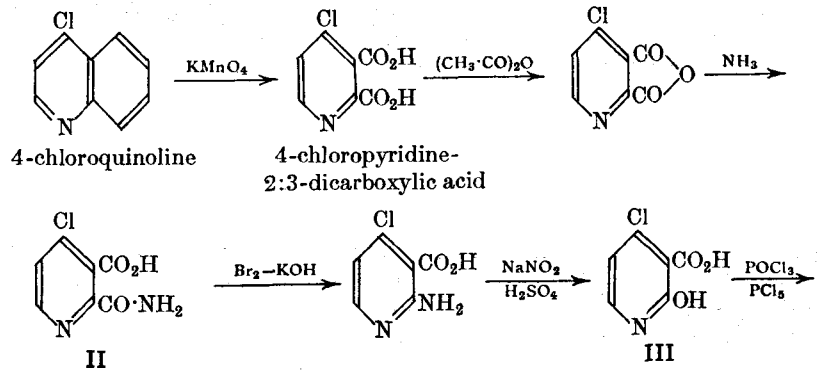


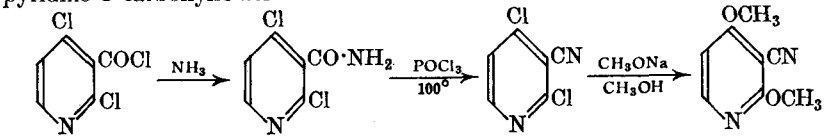
§16. Ricinine, $C_8H_8O_2N_2$, m.p. 201·5°, has been isolated from castor-oil seed; it is not a very toxic alkaloid. Degradative and synthetic work led to the suggestion that I is the structure of ricinine.

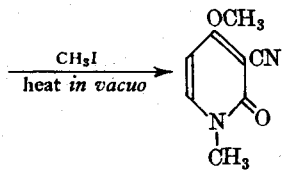


I

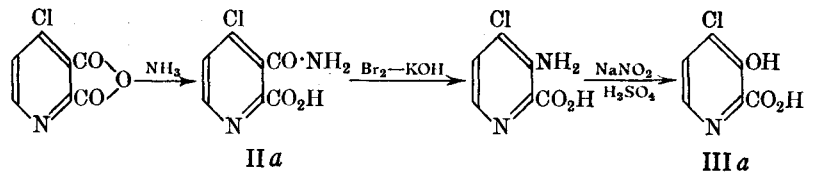
This has been confirmed by synthesis, *e.g.*, Späth *et al.* (1923);



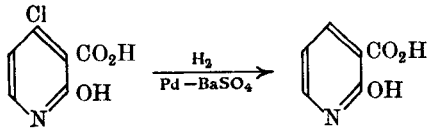




This is not an unambiguous synthesis, since II could have been 3-carbonamido-4-chloropyridine-2-carboxylic acid, II*a*, and consequently III would have been III*a*.

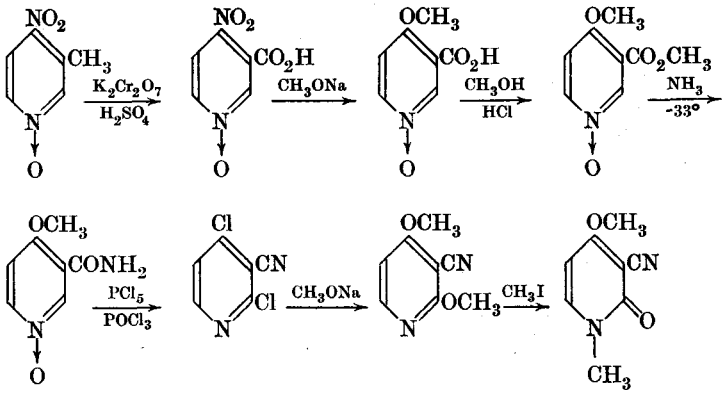


The structure of III was proved by the fact that on hydrogenation in the presence of Pd—$BaSO_4$, it gave 2-hydroxypyridine-3-carboxylic acid, IV.

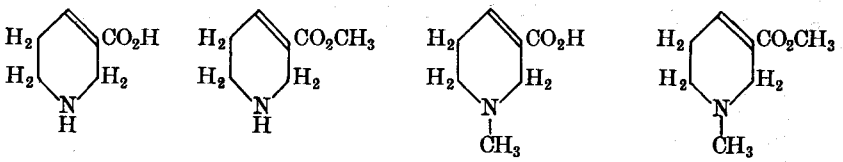


III IV

A more recent synthesis of ricinine is that of Taylor *et al.* (1956).

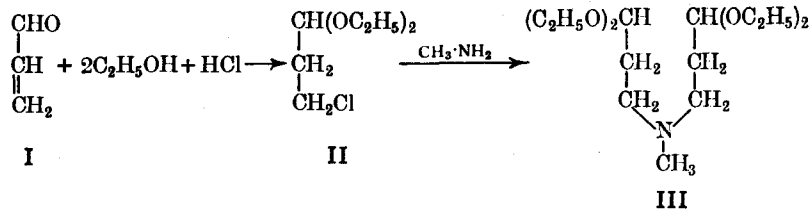


§17. Areca (or Betel) nut alkaloids. The betel nut is the source of a number of alkaloids which are all partially hydrogenated derivatives of nicotinic acid, *e.g.*,

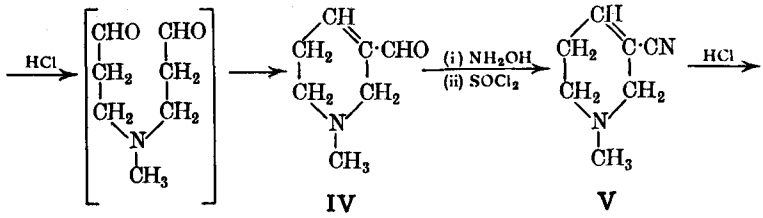


guvacine, m.p. 271-272° | guvacoline, b.p. 114°/13 mm. | arecaidine, m.p. 223° | arecoline, b.p. 209°

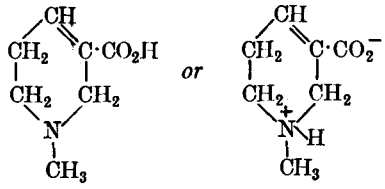
Let us consider arecaidine; its molecular formula is $C_7H_{11}O_2N$. When distilled with zinc dust, guvacine gives 3-methylpyridine; therefore this alkaloid is a pyridine derivative. Now guvacine is converted into arecaidine on heating with potassium methyl sulphate and sodium methoxide (Jahns, 1888, 1890); thus arecaidine is a methyl derivative of guvacine, and consequently is also a pyridine derivative. The usual tests show that arecaidine contains one carboxyl group, an *N*-methyl group and one double bond; hence the formula for arecaidine may be written as $C_5H_7N(CH_3)\cdot CO_2H$. Since the alkaloid is a pyridine derivative, the fragment C_5H_7N could be tetrahydropyridine. This was proved to be so by synthesis, and at the same time the positions of the double bond and carboxyl group were also established (Wohl *et al.*, 1907). Acraldehyde, I, on treatment with ethanol in the presence of hydrogen chloride, forms 3-chloropropionaldehyde acetal, II. II reacts with methylamine to form β-methyliminodipropionaldehyde tetra-acetal, III, which, on treatment with concentrated hydrochloric acid, ring closes to form 1 : 2 : 5 : 6-tetrahydro-1-methylpyridine-3-aldehyde, IV. This gives the cyano compound V on treatment with hydroxylamine, followed by dehydration of the oxime with thionyl chloride, and V is then converted into



I II III



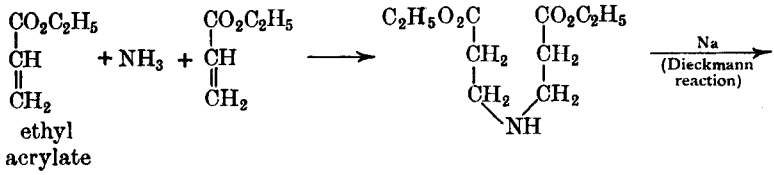
IV V



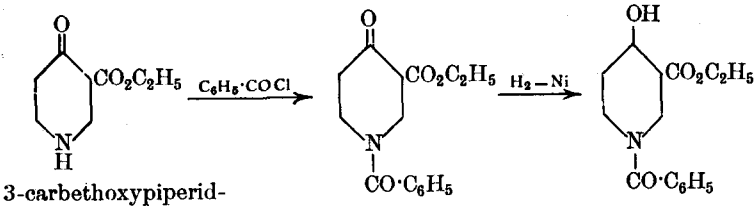
VI VI *a*

arecaidine by hydrolysis. Arecaidine is VI, or possibly VI*a*, the dipolar ion structure (*cf.* amino-acids and betaines).

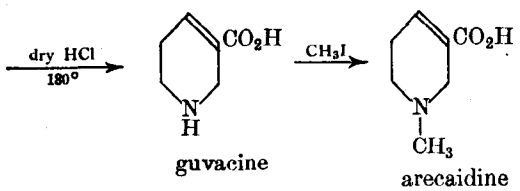
A more recent synthesis of arecaidine (and guvacine) is that of McElvain *et al.* (1946).



ethyl acrylate



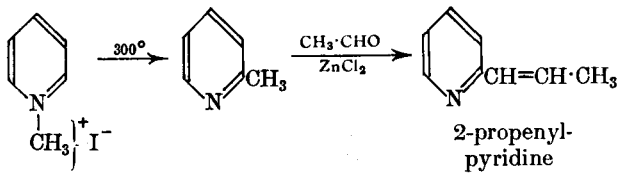
3-carbethoxypiperid-4-one



guvacine arecaidine

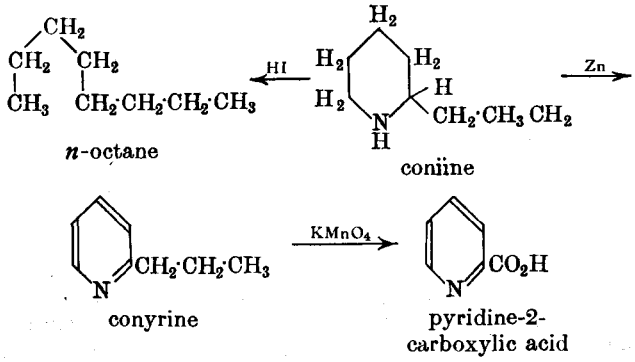
§18. Hemlock alkaloids. The most important alkaloid of this group is **coniine;** it was the first alkaloid to be synthesised. Oil of hemlock was drunk by Socrates when he was condemned to death in 399 B.C.

(+)-Coniine, $C_8H_{17}N$, b.p. 166–167°, is the form that occurs in oil of hemlock. When distilled with zinc dust, coniine is converted into conyrine, $C_8H_{11}N$ (Hofmann, 1884). Since the oxidation of conyrine with permanganate gives pyridine-2-carboxylic acid (α-picolinic acid), it follows that a pyridine nucleus is present with a side-chain in the 2-position. Thus coniine is probably a piperidine derivative with a side-chain in the 2-position. This side-chain must contain three carbon atoms, since two are lost when conyrine is oxidised. This side-chain is therefore either *n*-propyl or *iso*propyl, and it was actually shown to be *n*-propyl by the fact that when heated with

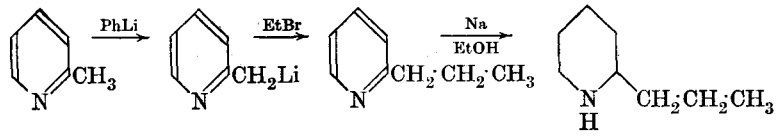


hydriodic acid at 300° under pressure, coniine forms *n*-octane. Had the side-chain been *iso*propyl, then the expected product would be *iso*-octane. From this evidence it therefore follows that coniine is 2-*n*-propylpiperidine, and this has been confirmed by synthesis (Ladenburg, 1885). The racemic coniine was resolved by means of (+)-tartaric acid, and the (+)-coniine so obtained was found to be identical with the natural compound.

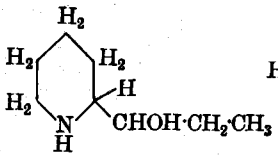
The reactions of coniine described above can therefore be formulated as follows:



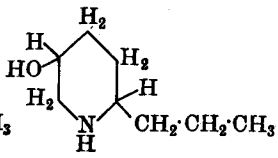
Coniine has also been synthesised from 2-methylpyridine and phenyllithium as follows (Bergmann *et al.*, 1932):



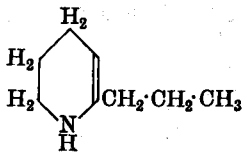
Other hemlock alkaloids are:



conhydrine

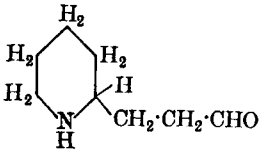


ψ-conhydrine

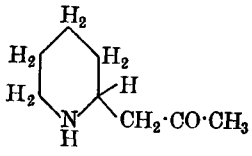


γ-coniceine

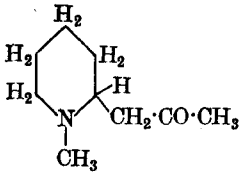
§19. Pomegranate alkaloids. The root bark of the pomegranate tree contains a number of alkaloids, the most important of which is pelletierine; three others are *iso*pelletierine, methyl*iso*pelletierine and pseudo-pelletierine. The last of these is related to atropine (§22).



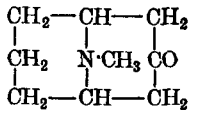
pelletierine



*iso*pelletierine

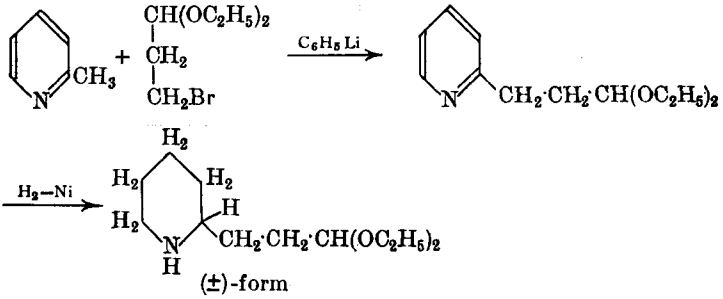


methyl*iso*pelletierine



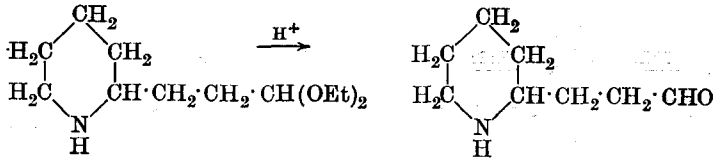
pseudo-pelletierine

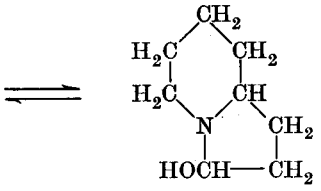
Pelletierine acetal has been synthesised by Spielman *et al.* (1941) by the action of 3-bromopropionaldehyde acetal on 2-methylpyridine (α-picoline) in the presence of phenyl-lithium, followed by catalytic reduction.



(±)-form

Pelletierine acetal was also prepared by Wibaut *et al.* (1940) who attempted to hydrolyse it to the free aldehyde; they obtained only viscous oils. Spielman *et al.* also failed to obtain the free aldehyde. Beets (1943) has therefore suggested that pelletierine can, and probably does, exist as some bicyclic structure such as I.



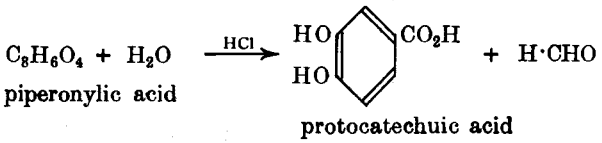


I

§20. Piperine, $C_{17}H_{19}O_3N$, m.p. 128–129·5°, occurs in pepper, especially black pepper (*Piper nigrum*). Hydrolysis of piperine with alkali gives

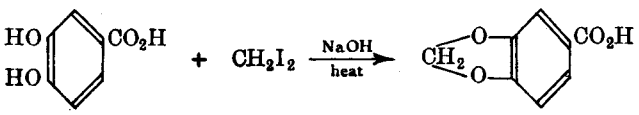
$$C_{17}H_{19}O_3N + N_2O \xrightarrow{KOH} \underset{\text{piperic acid}}{C_{12}H_{10}O_4} + \underset{\text{piperidine}}{C_5H_{11}N}$$

piperic acid and piperidine; thus the alkaloid is the piperidine amide of piperic acid (Babo *et al.*, 1857). Since piperidine is hexahydropyridine, the structure of piperine rests on the elucidation of that of piperic acid. The routine tests show that piperic acid contains one carboxyl group and two double bonds. When oxidised with permanganate, piperic acid gives first piperonal and then piperonylic acid. The structure of the latter is deduced from the fact that when heated with hydrochloric acid at 200° under pressure, piperonylic acid forms protocatechuic acid (3 : 4-dihydroxybenzoic acid) and formaldehyde.



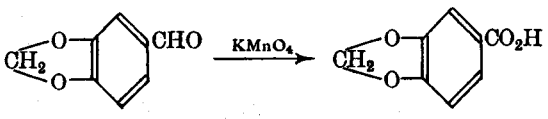
piperonylic acid — protocatechuic acid

Since one atom of carbon is eliminated, and there are no free hydroxyl groups in piperonylic acid, the structure of this acid is probably the methylene ether of protocatechuic acid, *i.e.*, piperonylic acid is 3 : 4-methylenedioxybenzoic acid; this has been confirmed by synthesis:



piperonylic acid

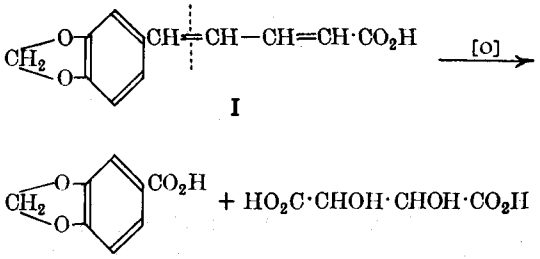
Furthermore, since piperonal (an aldehyde) gives piperonylic acid on oxidation, piperonal is therefore 3 : 4-methylenedioxybenzaldehyde.



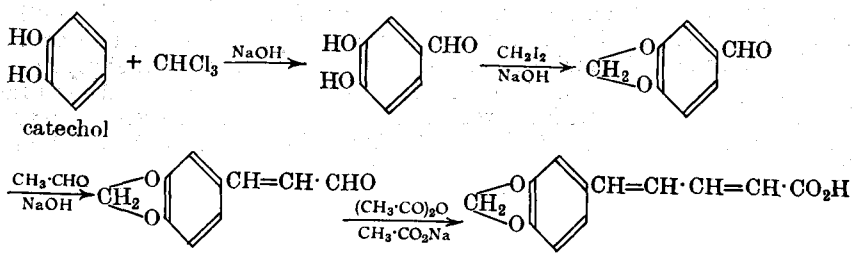
piperonal

From these results of oxidative degradation, it therefore follows that piperic acid is a benzene derivative containing only one side-chain. It is this side-chain that contains the two double bonds (the ready addition of four bromine atoms shows the presence of two *ethylenic* bonds), and since the careful oxidation of piperic acid gives tartaric acid in addition to piperonal and piperonylic acid, the side-chain is a " straight " chain. If we assume I as

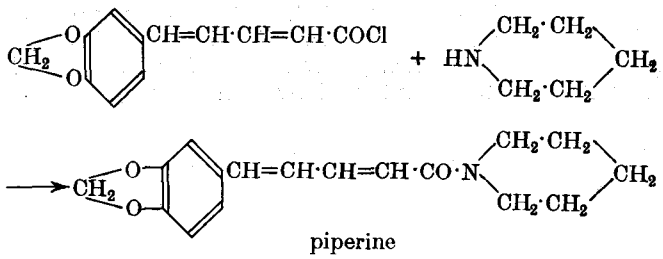
the structure of piperic acid, then all of the foregoing products of oxidation may be accounted for.



This has been confirmed by synthesis (Ladenburg *et al.*, 1894); piperonal (prepared *via* the Reimer–Tiemann reaction) is condensed with acetaldehyde in the presence of sodium hydroxide (Claisen–Schmidt reaction), and the product (a cinnamaldehyde derivative) is then heated with acetic anhydride in the presence of sodium acetate (Perkin reaction).



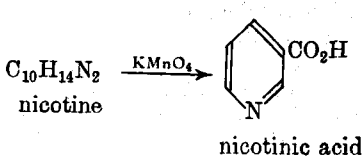
When the acid chloride of piperic acid (prepared by the action of phosphorus pentachloride on the acid) is heated with piperidine in benzene solution, piperine is formed; thus piperine is the piperidine amide of piperic acid.



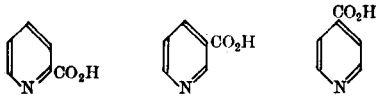
PYRROLIDINE-PYRIDINE GROUP

§21. Tobacco alkaloids. Many alkaloids have been isolated from the tobacco leaf, *e.g.*, nicotine, nicotimine (anabasine), nornicotine, etc.

Nicotine, $C_{10}H_{14}N_2$, b.p. 247°, is the best known and most widely distributed of the tobacco alkaloids; it occurs naturally as the (—)-form. When oxidised with dichromate–sulphuric acid (or permanganate or nitric acid), nicotine forms nicotinic acid (Huber, 1867).

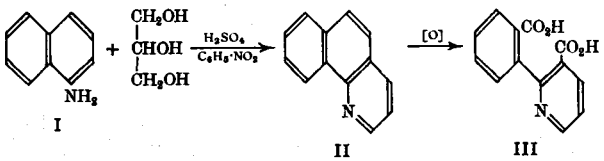


It is instructive, at this point, to see how the orientations of the three isomeric pyridinecarboxylic acids have been elucidated.

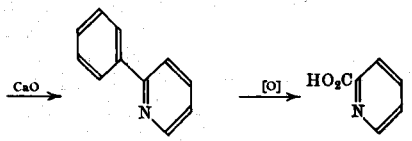


picolinic acid, m.p. 137° nicotinic acid, m.p. 234-237° *iso*nicotinic acid, m.p. 317°

Picolinic acid. 1-Naphthylamine, I, when subjected to the Skraup synthesis (see Vol. I), is converted into 7 : 8-benzoquinoline, II (this structure is established by its synthesis). II, on vigorous oxidation with alkaline permanganate, gives the dicarboxylic acid III which, when decarboxylated by heating with calcium oxide, is converted into 2-phenylpyridine, IV. This, on further oxidation with permanganate, gives a pyridinecarboxylic acid which must, from the structure of IV, be the 2-acid, *i.e.*, picolinic acid, V.

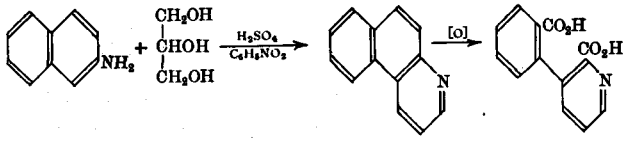


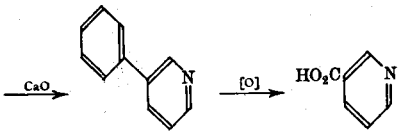
I II III



IV V

Nicotinic acid. This has been shown to be pyridine-3-carboxylic acid by a similar set of reactions, except that in this case the starting material is 2-naphthylamine.

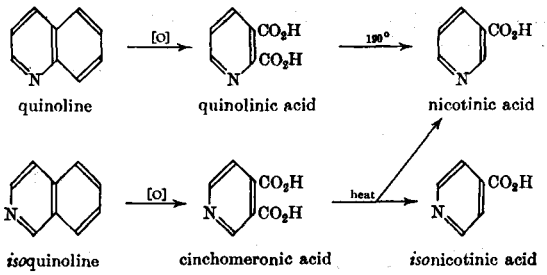




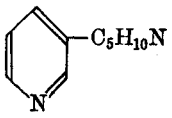
nicotinic acid

iso*Nicotinic acid.* This third isomer is therefore pyridine-4-carboxylic acid.

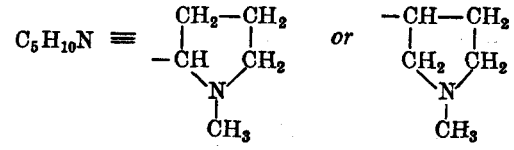
An alternative proof for the orientations of these three acids is based on the structures of quinoline and *iso*quinoline (which have been established by synthesis). Oxidation of quinoline with alkaline permanganate gives quinolinic acid which, by its method of preparation, must be pyridine-2 : 3-dicarboxylic acid. When quinolinic acid is heated to 190°, one carboxyl group is lost to produce nicotinic acid; thus nicotinic acid must be either pyridine-2- or -3-carboxylic acid. *iso*Quinoline, on oxidation with alkaline permanganate, produces cinchomeronic acid, which must therefore be pyridine-3 : 4-dicarboxylic acid. This, on gentle heating, gives a mixture of nicotinic and *iso*nicotinic acids; thus nicotinic acid must be the 3-acid, and *iso*nicotinic acid the 4-acid. Hence picolinic acid is pyridine-2-carboxylic acid.



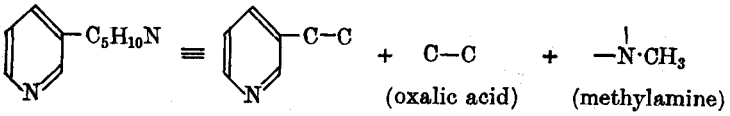
Returning to the structure of nicotine, since nicotinic acid is a product of oxidation, the alkaloid therefore contains a pyridine nucleus with a complex side-chain in the 3-position. Thus we may write the formula of nicotine as



Because of its formula, this side-chain was originally believed to be piperidine, but further work showed that this was incorrect. When nicotine zinci-chloride is distilled, the products are pyridine, pyrrole and methylamine (Laiblin, 1879). This suggests that the side-chain $C_5H_{10}N$ is a pyrrole derivative. Furthermore, when nicotine is heated with concentrated hydriodic acid at 150° (Herzig–Meyer method), methyl iodide is formed. Thus the side-chain contains an *N*-methyl group. It therefore appears that the side-chain could be *N*-methylpyrrolidine, but its point of attachment to the pyridine nucleus could be either 2 or 3 on the evidence obtained so far:

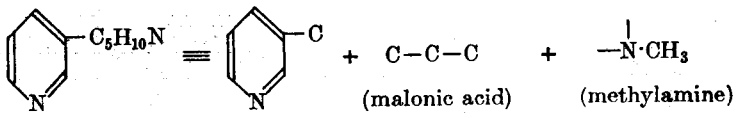


The correct structure of nicotine was obtained by Pinner (1892, 1893). Treatment of nicotine with bromine in acetic acid gives, among other products, the hydrobromide perbromide, $C_{10}H_{10}ON_2Br_2 \cdot HBr \cdot Br_2$, which, when treated with aqueous sulphurous acid, is converted into dibromocotinine, $C_{10}H_{10}ON_2Br_2$. This, on heating with a mixture of sulphurous and sulphuric acids at 130–140°, forms 3-acetylpyridine, oxalic acid and methylamine. Thus the structure of nicotine must account for the following skeleton structures:

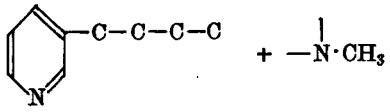


(3-acetylpyridine)

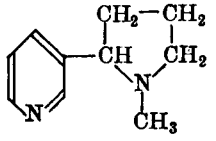
Now bromine, in the presence of hydrobromic acid, converts nicotine into dibromoticonine, $C_{10}H_8O_2N_2Br_2$, which, on heating with barium hydroxide solution at 100°, forms nicotinic acid, malonic acid and methylamine. Hence the structure of nicotine must also account for the following skeleton structures:



These two sets of reactions, taken in conjunction with one another, are satisfied by the following skeleton for nicotine:

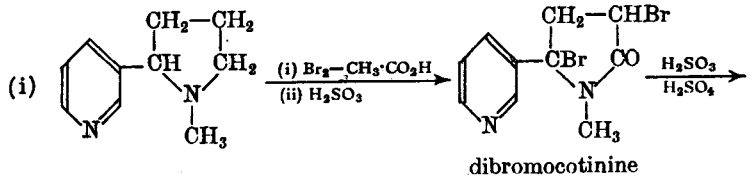


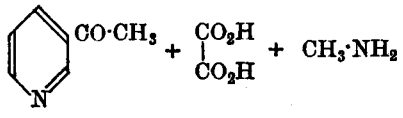
The problem now is: Where is the position of the *N*-methyl group? Nicotine behaves as a ***di-tertiary base***, and forms two isomeric " methyl iodide addition products ". Thus the nitrogen atom in the side-chain must be of the type —C—N(CH$_3$)—C—. Furthermore, it is extremely difficult to reduce nicotine beyond hexahydronicotine (the pyridine part is reduced to piperidine). Hence the side-chain must be saturated, and this can only be so if the side-chain is cyclic, *i.e.*, *N*-methylpyrrolidine ($C_5H_{11}N \equiv C_4H_8 \cdot NCH_3 \equiv C_4H_8$). The presence of this pyrrolidine nucleus also accounts for the formation of pyrrole when nicotine zincichloride is distilled (see above). All the foregoing facts are satisfied by the following structure for nicotine.

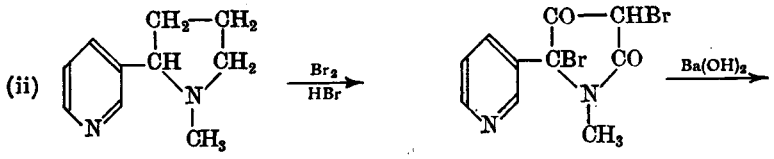


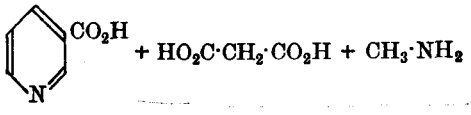
nicotine

On this basis, Pinner's work may be formulated:

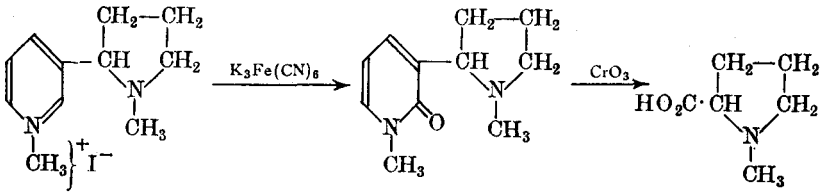






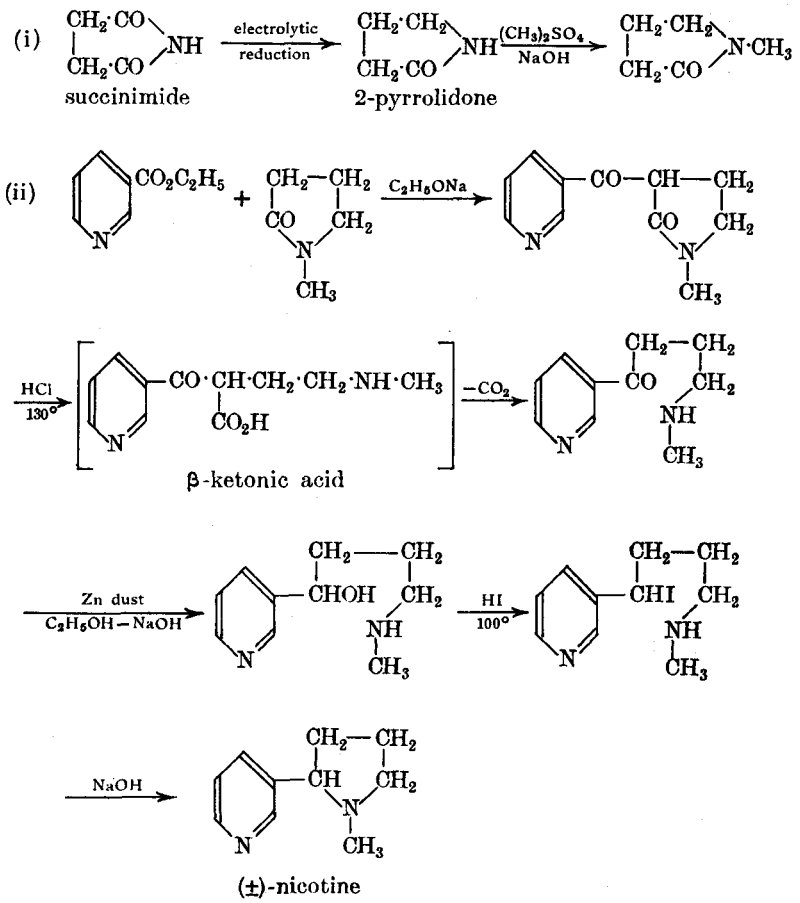


The most direct analytical evidence for the presence of the pyrrolidine nucleus has been given by Karrer (1925, 1926); nicotine hydriodide forms nicotine *iso*methiodide when warmed with methyl iodide and this, on oxidation with potassium ferricyanide, is converted into nicotone which, on oxidation with chromium trioxide, gives hygrinic acid (§13).



nicotine *iso*methiodide — nicotone — hygrinic acid

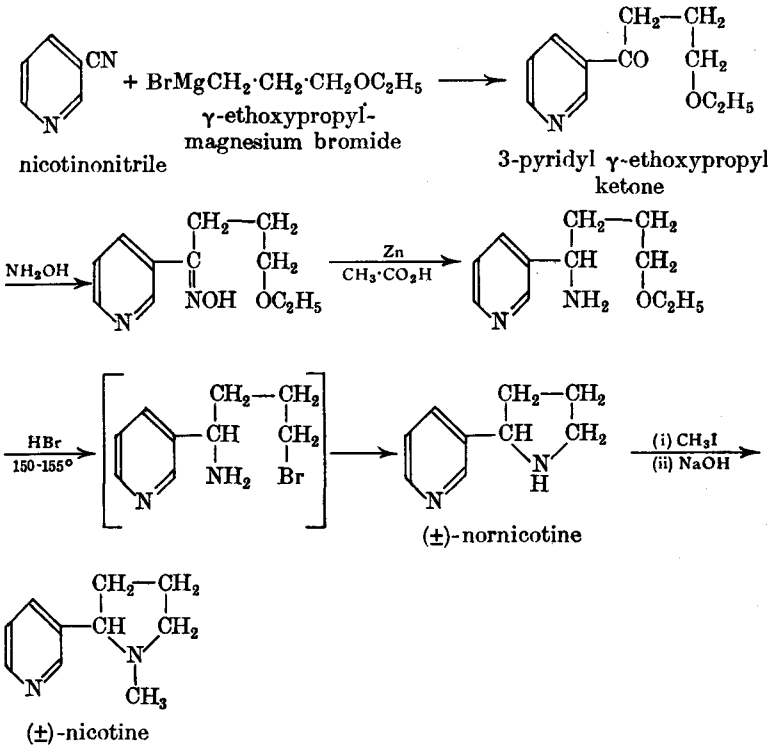
Pinner's formula for nicotine has been confirmed by synthesis, *e.g.*, *Späth and Bretschneider* (1928).



(±)-nicotine

This was resolved by means of (+)-tartaric acid; the synthetic (−)-nicotine is identical with the natural compound.

Craig (1933).



Späth *et al.* (1936) have resolved (±)-nornicotine; methylation of the (−)-form with formaldehyde and formic acid gave (−)-nicotine, identical with the natural product.

§22. Solanaceous alkaloids. This group includes atropine, hyoscyamine and scopolamine (hyoscine).

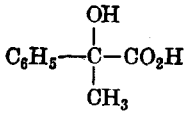
Atropine, $C_{17}H_{23}O_3N$, m.p. 118°, occurs in deadly nightshade (*Atropa belladonna*) together with hyoscyamine. Hyoscyamine is optically active (lævorotatory), but readily racemises to atropine when warmed in an ethanolic alkaline solution; thus atropine is (±)-hyoscyamine.

When warmed with barium hydroxide solution, atropine is hydrolysed to (±)-tropic acid and tropine (an alcohol); thus atropine is the tropine ester of tropic acid.

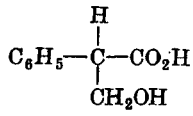
(±)-Tropic acid, $C_9H_{10}O_3$, m.p. **117°**, is a saturated compound (it does not add on bromine); the usual tests show that it contains one carboxyl group and one alcoholic group. When heated strongly, tropic acid loses a molecule of water to form atropic acid, $C_9H_8O_2$, and this, on oxidation,

$C_6H_5{\cdot}CH{=}CH{\cdot}CO_2H$ (I) $\qquad$ $C_6H_5{\cdot}C(=CH_2){\cdot}CO_2H$ (II)

gives benzoic acid. Thus tropic and atropic acids contain a benzene ring with one side-chain. It therefore follows that atropic acid could be either I or II. Since, however, I is known to be cinnamic acid, II must be atropic acid. Addition of a molecule of water to II would therefore give tropic

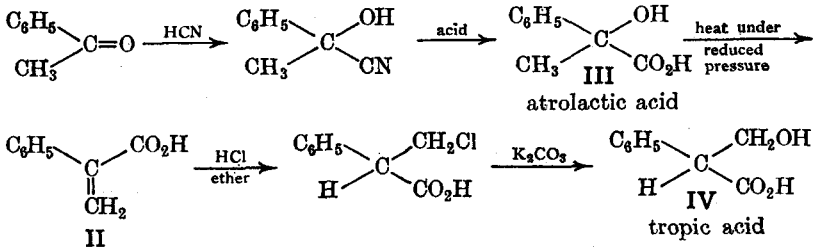


III



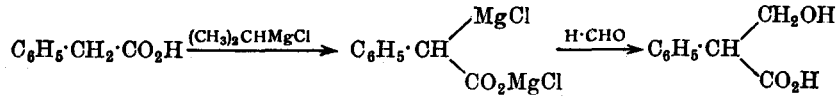
IV

acid which, consequently, must be either III or IV. Tropic acid has been shown to be IV by synthesis, *e.g.*, Mackenzie and Wood (1919), starting from acetophenone.

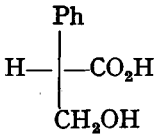


III is atrolactic acid, and its dehydration to II confirms the structure of atropic acid. It should also be noted that the addition of hydrogen chloride takes place contrary to Markownikoff's rule (see unsaturated acids, Vol. I); had the addition been in accordance with the rule, then atrolactic acid would have again been obtained. It is tropic acid that contains the asymmetric carbon atom which gives rise to the optically active hyoscyamine. The above synthesis results in (±)-tropic acid, and this has been resolved by means of quinine.

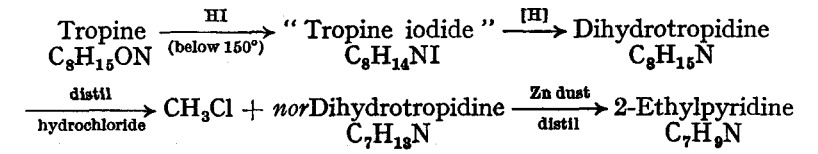
Blicke *et al.* (1952) have synthesised tropic acid by boiling phenylacetic acid with *iso*propylmagnesium chloride in ethereal solution, and then treating the product, a Grignard reagent, with formaldehyde.



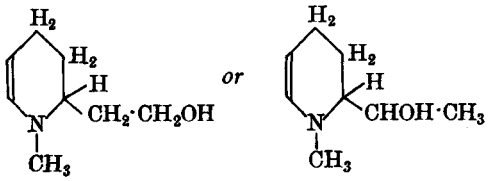
Fodor *et al.* (1961) have established the absolute configuration of (—)-tropic acid by its correlation with (—)-alanine. According to the Cahn–Ingold–Prelog convention (§5c. II), natural tropic acid is (*S*)-(—)-tropic acid.



Tropine (tropanol), $C_8H_{15}ON$, m.p. 63°, behaves as a saturated compound which contains an alcoholic group. The structure of tropine was investigated by Ladenburg (1883, 1887), who showed that the molecule contained a reduced pyridine nucleus:



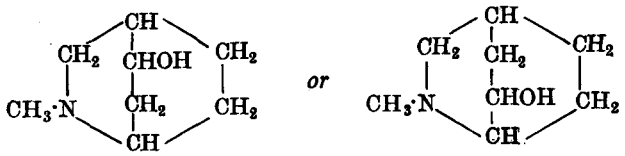
"Tropine iodide" is formed by the replacement of the alcoholic group in tropine by an iodine atom, which is then replaced by hydrogen to form dihydrotropidine (tropane). The formation of methyl chloride indicates the presence of an *N*-methyl group, and the isolation of 2-ethylpyridine shows the presence of this nucleus (in a reduced form). Largely on this evidence, Ladenburg was led to suggest the following alternative formulæ for tropine:



Merling (1891), by the oxidation of tropine with chromium trioxide, obtained (±)-tropinic acid.

$$\underset{\text{tropine}}{C_8H_{15}ON} \xrightarrow{CrO_3} \underset{(\pm)\text{-tropinic acid}}{C_8H_{13}O_4N}$$

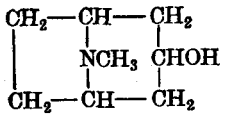
Tropinic acid is a dicarboxylic acid, and since there is no loss of carbon in its formation, the hydroxyl group in tropine must therefore be in a ring system. Thus Ladenburg's formula is untenable, and so Merling proposed the following structures for tropine:



Willstätter (1895–1901) then examined the oxidation products of tropine obtained as follows:

$$\underset{C_8H_{15}ON}{\text{Tropine}} \xrightarrow{CrO_3} \underset{C_8H_{13}ON}{\text{Tropinone}} \xrightarrow{CrO_3} \underset{C_8H_{13}O_4N}{(\pm)\text{-Tropinic acid}}$$

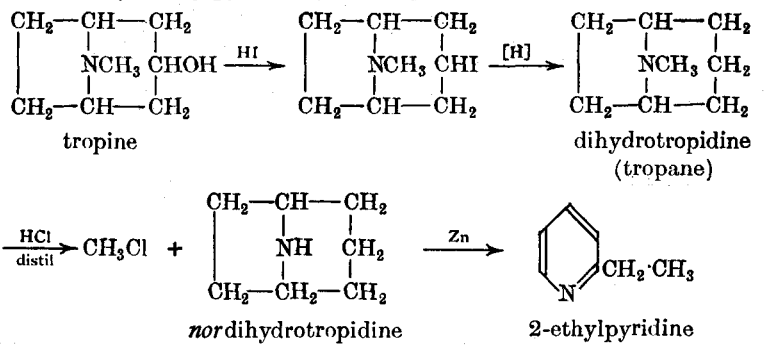
Tropinone behaved as a ketone; thus tropine is a secondary alcohol (*cf.* Merling's formula). Willstätter (1897) also showed that tropinone forms a dibenzylidene derivative with benzaldehyde, and a di-oximino derivative when treated with amyl nitrite and hydrochloric acid. Thus tropinone contains the $\cdot CH_2 \cdot CO \cdot CH_2 \cdot$ grouping, and so it follows that Merling's formula is also untenable. Willstätter therefore proposed three possible structures for tropine, but eliminated two by the consideration of various reactions of tropine, and was left with the following (which contains a pyridine and a pyrrole nucleus with the nitrogen atom common to both):



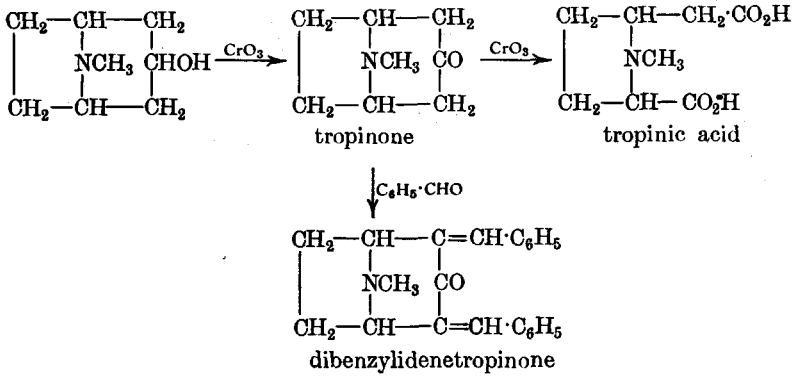
Not only did this fit the facts best, but it was also supported by the following evidence: (i) Exhaustive methylation of tropine gives tropilidene (*cyclo*-heptatriene), C_7H_8. (ii) Exhaustive methylation of tropinic acid gives an unsaturated dicarboxylic acid which, on reduction, forms pimelic acid.

All the foregoing reactions of tropine can be readily explained on the Willstätter formula.

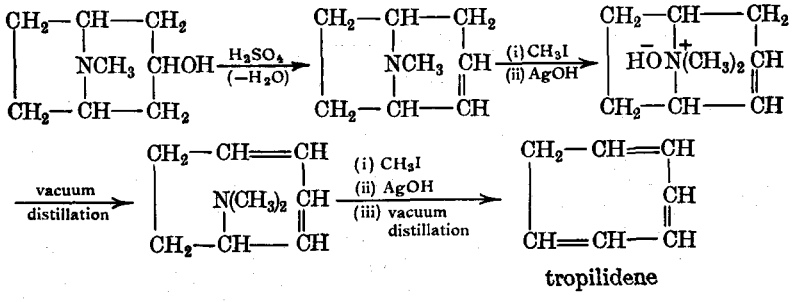
Formation of 2-ethylpyridine from tropine.



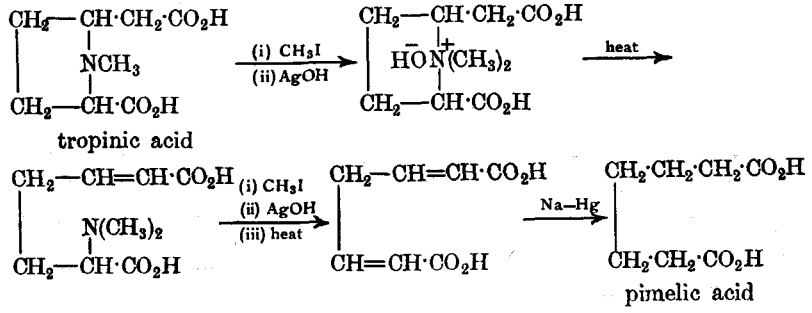
Formation of tropinone and tropinic acid from tropine.



Formation of tropilidene from tropine.

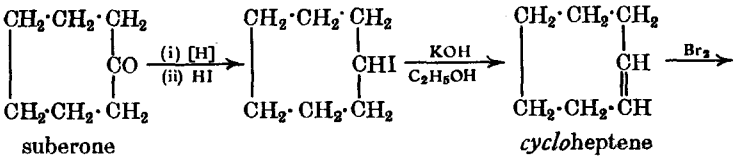


Formation of pimelic acid from tropinic acid.

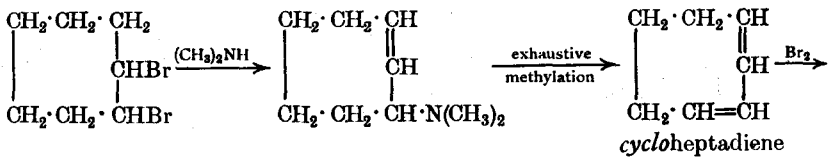


The structure of tropine has been confirmed by synthesis, one by Willstätter (1900–1903), and the other by Robinson (1917).

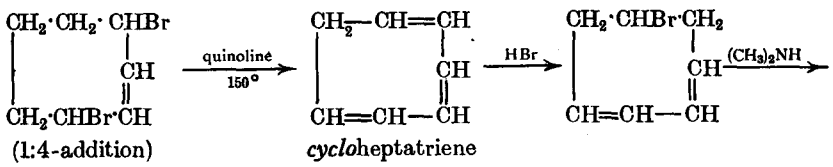
Willstätter's synthesis.



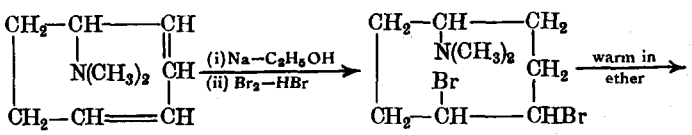
suberone *cyclo*heptene

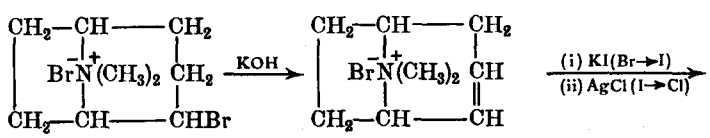


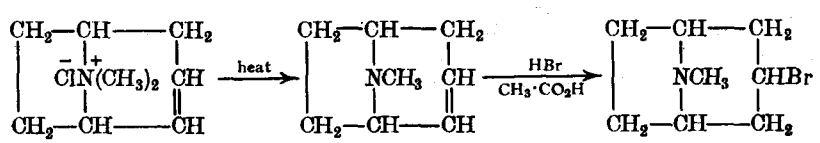
*cyclo*heptadiene



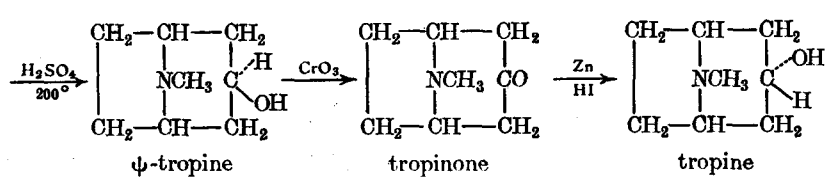
(1:4-addition) *cyclo*heptatriene (tropilidene)







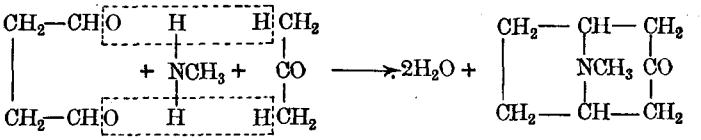
tropidine



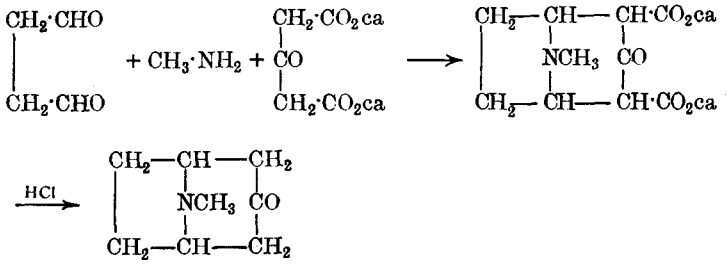
ψ-tropine tropinone tropine

Robinson's synthesis.

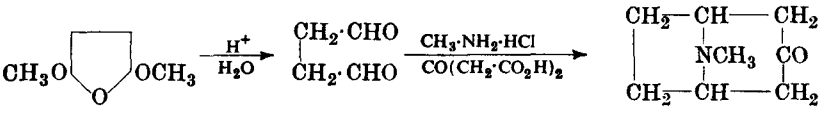
When a mixture of succinaldehyde, methylamine and acetone is allowed to stand in water for thirty minutes, tropinone is produced in very small yield.



A much better yield (40 per cent.) is obtained by using calcium acetonedicarboxylate or ethyl acetonedicarboxylate instead of acetone; the calcium salt or ester so produced is converted into tropinone by warming with hydrochloric acid, *e.g.* (ca = Ca/2):

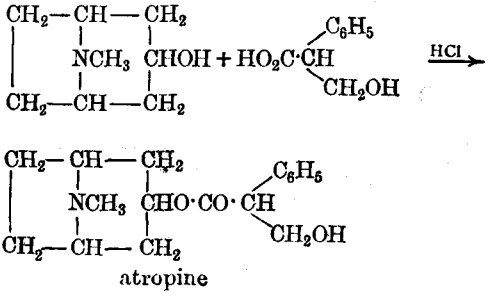


Schöpf *et al.* (1935) have obtained a yield of 70–85 per cent. by carrying out Robinson's synthesis at a *p*H of 7. Elming *et al.* (1958) have also synthesised tropinone using methylamine hydrochloride, acetonedicarboxylic acid and generating succinaldehyde *in situ* by the action of acid on 2,5-dimethoxytetrahydrofuran:



The yield was 81 per cent., but in this case "physiological conditions" were not necessary (see §28).

The final problem is to combine tropine with tropic acid; this has been done by heating the two together in the presence of hydrogen chloride (Fischer–Speier esterification; see Vol. I).



atropine

Stereochemistry of the tropines. Tropinone can be reduced to tropine, together with a small amount of *ψ*-tropine, by means of a metal and

acid, the best combination being zinc dust and hydriodic acid; or by means of electrolytic reduction. On the other hand, reduction with sodium amalgam converts tropinone into ψ-tropine. According to Mirza (1952), lithium aluminium hydride reduces tropinone quantitatively to ψ-tropine, but according to Beckett *et al.* (1957), 54 per cent. of ψ-tropine and 45 per cent. of tropine are obtained. A larger yield of the former (69 per cent.) is obtained with sodium borohydride, and reduction with sodium and *iso*butanol (in toluene) gives the maximum yield of ψ-tropine (88 per cent.).

Tropine and ψ-tropine are geometrical isomers, one isomer having the hydrogen atom on C_3 on the same side as the nitrogen bridge, and the other isomer has this hydrogen atom on the opposite side (*cf.* the borneols, §23b. VIII); Fig. 1 shows the two possible forms. Neither of these forms is optically

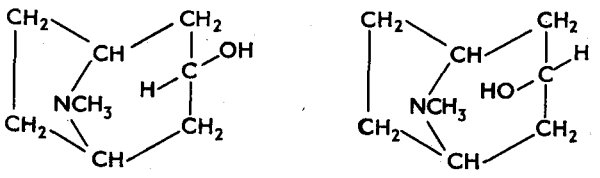


FIG. 14.1.

active, since the molecule has a plane of symmetry. C_1 and C_5 are asymmetric, but the molecule is optically inactive by internal compensation (see §7b. II), and so each isomer is a *meso*-form; C_3 is pseudo-asymmetric (see §8. IV). It should also be noted that another pair of *optically active forms* would exist if the fusion of the nitrogen bridge were *trans*; this, however, is not possible (*cf.* camphor, §23a. VIII; also cocaine, §23).

The problem now is to decide which geometrical isomer (of the two forms shown in Fig. 1) is tropine and which is ψ-tropine. Fodor (1953) has given evidence to show that ψ-tropine is the *syn*-compound (nitrogen bridge and hydroxyl group are in the *cis*-position; Fig. 1*b*), and that tropine is the *anti*-compound (nitrogen bridge and hydroxyl group are in the *trans*-position; Fig. 1*a*). The problem, however, is more involved than this, since the conformation of the piperidine ring has also to be considered. Fodor gives the configuration of the piperidine ring as the boat form in both isomers (Fig. 2).

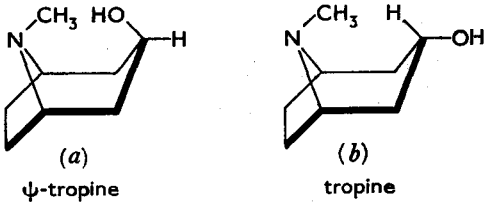


FIG. 14.2.

Zenitz *et al.* (1952) and Clemo *et al.* (1953) support these configurations from evidence obtained by measurements of the dipole moments of these two isomers; ψ-tropine has been shown to have a higher dipole moment than tropine. Zenitz *et al.* have also shown from infra-red absorption spectra measurements that ψ-tropine has intramolecular hydrogen bonding; this is only possible in the *syn*-form. Bose *et al.* (1953), however, have assumed the chair form for the piperidine ring by analogy with the chair conformation of *cyclo*hexane compounds and pyranosides (see §11. IV). Thus these authors have suggested that ψ-tropine is Fig. 3 (*a*), in which the hydroxyl

group is equatorial, and that tropine is Fig. 3 (*b*), in which the hydroxyl group is axial.

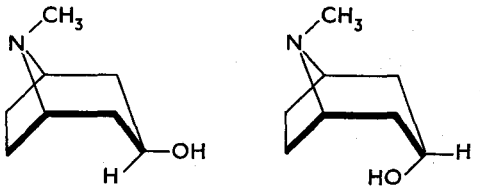
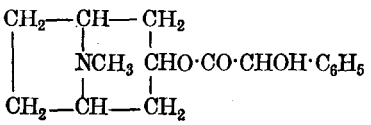


FIG. 14.3.

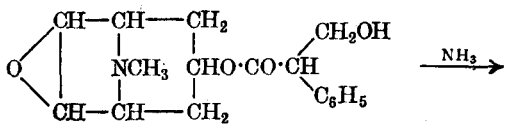
If these be the configurations, then it is difficult to explain Fodor's work (which involves rearrangements), and also the fact that there is intramolecular hydrogen bonding in ψ-tropine. Sparke (1953) has suggested that the chair form can readily change into the boat form; this would then explain the intramolecular hydrogen bonding. Archer and Lewis (1954) also adopt this explanation, but make the assumption that the bond energy involved in the hydrogen bond is sufficient to transform, at least partially, the more stable chair form into the less stable boat form; in ψ-tropine the chair and boat forms are in mobile equilibrium, the latter being the predominant form.

§22a. Tropeines and pseudotropeines. These are synthetic esters formed respectively from tropine and ψ-tropine with various organic acids. The tropeines (including atropine itself) are powerful mydriatics (pupil dilators) and feeble anæsthetics; the ψ-tropeines are the reverse. One of the most important tropeines is *homatropine* (*mandelyltropeine*), which is prepared by combining tropine with mandelic acid.

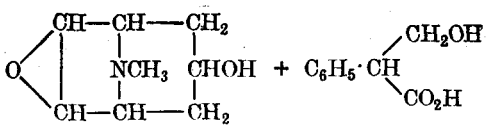


homatropine

§22b. Hyoscine (scopolamine), $C_{17}H_{21}O_4N$, is a syrup and is lævorotatory; it is obtained from various sources, *e.g.*, *Datura Metel.* Hyoscine is a constituent of travel sickness tablets, and when administered with morphine, produces " twilight sleep ". Hyoscine is the (—)-tropic ester of the aminoalcohol *scopine*; these two compounds are produced by the hydrolysis of hyoscine with ammonia.

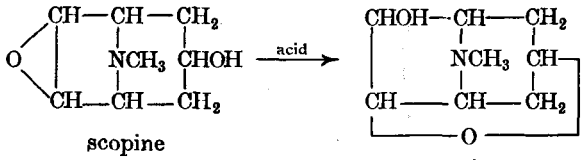


hyoscine



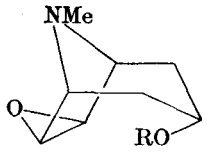
scopine tropic acid

More vigorous hydrolysis of hyoscine with acids or alkalis produces *oscine* (*scopoline*), which is formed by the isomerisation of scopine.



scopine oscine

It is interesting to note, in this connection, that the action of *ethanolic* sodium hydroxide on (—)-hyoscine at room temperature causes the latter



I

to racemise to (±)-hyoscine. Fodor *et al.* (1959) have carried out a total synthesis of (±)-hyoscine and shown its conformation to be

(I; R = CO·CHPh·CH_2OH).

§23. Coca alkaloids. In this group occur cocaine, benzoylecgonine, tropacocaine, hygrine (§13), cuscohygrine (§13a), etc.

(—)-Cocaine, $C_{17}H_{21}O_4N$, m.p. 98°, occurs in coca leaves; it is sparingly soluble in water, but its hydrochloride is quite soluble and is used as a local anæsthetic. When heated with water, cocaine is hydrolysed to methanol and benzoylecgonine.

$$\underset{\text{cocaine}}{C_{17}H_{21}O_4N} + H_2O \longrightarrow \underset{\text{benzoylecgonine}}{C_{16}H_{19}O_4N} + CH_3OH$$

Thus cocaine contains a carbomethoxyl group, and benzoylecgonine a carboxyl group. When benzoylecgonine is heated with barium hydroxide solution, further hydrolysis occurs, the products obtained being benzoic acid and ecgonine.

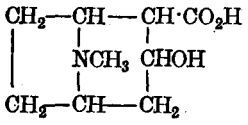
$$\underset{\text{benzoylecgonine}}{C_{16}H_{19}O_4N} + H_2O \xrightarrow{Ba(OH)_2} \underset{\text{ecgonine}}{C_9H_{15}O_3N} + C_6H_5{\cdot}CO_2H$$

Ecgonine shows the reactions of an alcohol, and so benzoylecgonine is the benzoyl derivative of a hydroxycarboxylic acid. The structure of ecgonine has been deduced from the nature of the products obtained by oxidation, *viz.*,

$$\underset{C_9H_{15}O_3N}{\text{Ecgonine}} \xrightarrow{CrO_3} \underset{C_8H_{13}ON}{\text{Tropinone}} \xrightarrow{CrO_3} \underset{C_8H_{13}O_4N}{\text{Tropinic acid}} + \underset{C_7H_{11}O_3N}{\text{Ecgoninic acid}}$$

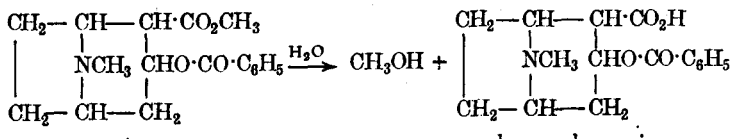
From these results, it follows that ecgonine contains the tropane structure and that the alcoholic group must be in the same position as in tropine (§22). Now in the formation of tropinone from ecgonine, a carboxyl group is lost (as we have seen, ecgonine contains a carboxyl group). Thus the carboxyl group is in a position such that the oxidation of the secondary alcoholic group in ecgonine to a keto group is accompanied by the elimination of the carboxyl group. This type of elimination is characteristic of β-ketonic acids, and this interpretation of the results is confirmed by the

fact that Willstätter *et al.* (1898) actually observed the formation of an unstable β-ketonic acid which lost carbon dioxide to give tropinone. Thus ecgonine is:

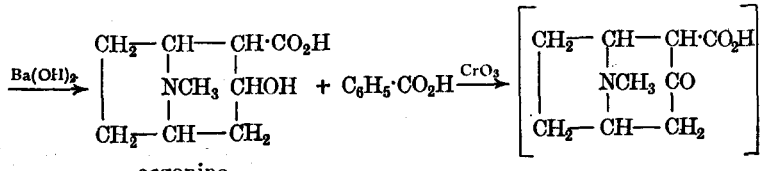


ecgonine

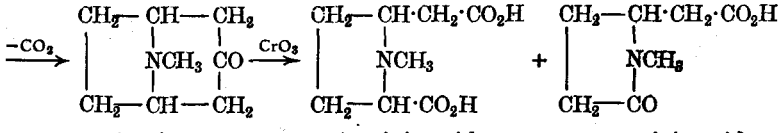
On this basis, the foregoing reactions may therefore be written:



cocaine benzoylecgonine

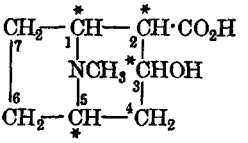


ecgonine



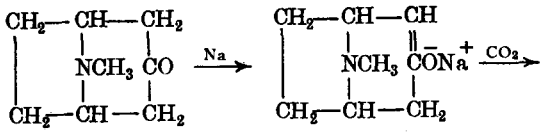
tropinone tropinic acid ecgoninic acid

The structure of ecgonine has been confirmed by synthesis (Willstätter *et al.*, 1901); the starting point is tropinone (see §22 for its synthesis). Before describing this synthesis, let us first examine the structure of ecgonine from the stereochemical point of view; it will be seen that there are four dissimilar

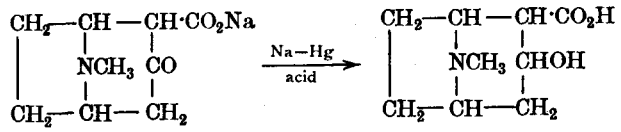


asymmetric carbon atoms present (*), and so there are $2^4 = 16$ optically active forms (eight pairs of enantiomorphs) possible (*cf.* tropine, §22). Since, however, only the *cis* fusion of the nitrogen bridge is possible in practice, C_1 and C_5 therefore have only one configuration (the *cis*-form), and so there are only eight optically active forms (four pairs of enantiomorphs) actually possible (*cf.* camphor, §23a. VIII); three pairs of enantiomorphs have been prepared synthetically.

In the original synthesis of Willstätter, the racemic ecgonine obtained was not identical with the (—)-ecgonine from (—)-cocaine, but its chemical properties were the same.

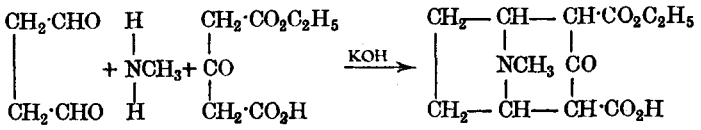


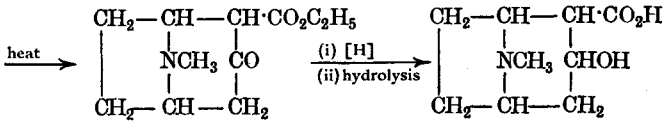
tropinone



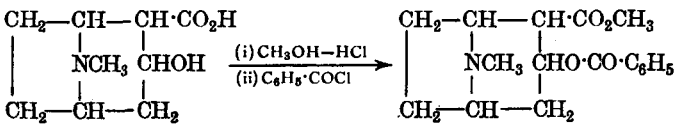
sodium tropinonecarboxylate — a (±)-ecgonine

Later, Willstätter *et al.* (1921) synthesised ecgonine by means of the Robinson method (see §22):





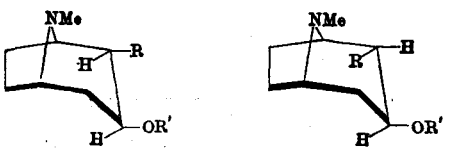
The final product was shown to be a mixture of three racemates, (±)-ecgonine, (±)-ψ-ecgonine and a third pair of enantiomorphs (Willstätter *et al.*, 1923). The racemic ecgonine was resolved, and the (−)-form esterified with methanol and then benzoylated; the product was (−)-cocaine.



(−)-ecgonine — (−)-cocaine

In a similar way, the (+)- and (−)-ψ-cocaines were obtained from the corresponding ψ-ecgonines. An interesting point in this connection is that Einhorn *et al.* (1890) showed that the prolonged action of 33 per cent. aqueous potassium hydroxide converts ecgonine into ψ-ecgonine, and Findlay (1953) has found that cocaine gives ψ-ecgonine methyl ester by the action of sodium methoxide in hot methanol.

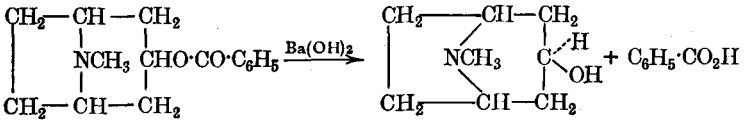
Fodor *et al.* (1953, 1954) and Findlay (1953, 1954) have established the conformations of ecgonine and ψ-ecgonine (R = CO_2H; R′ = H) and the corresponding cocaines (R = CO_2Me; R′ = COPh) (*cf.* §22):



cocaine (and ecgonine) — ψ-cocaine (and ψ-ecgonine)

Hardegger *et al.* (1955) have correlated (−)-cocaine with L-glutamic acid and have shown that the formula represents the absolute configuration of L(−)-cocaine.

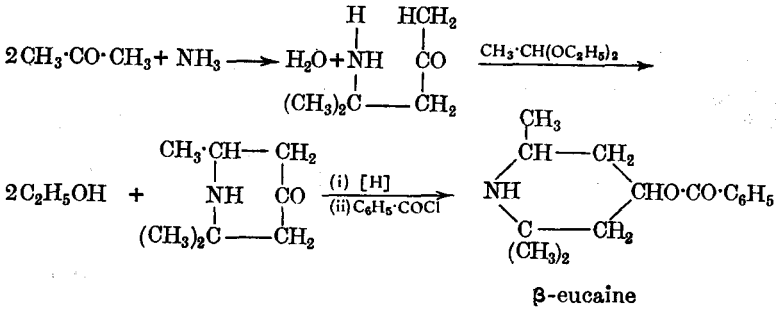
§23a. **Tropacocaine,** $C_{15}H_{19}O_2N$, m.p. 49°, occurs in Java coca leaves. When heated with barium hydroxide solution, tropacocaine is hydrolysed to ψ-tropine and benzoic acid; thus the alkaloid is benzoyl-ψ-tropine.



tropacocaine ψ-tropine

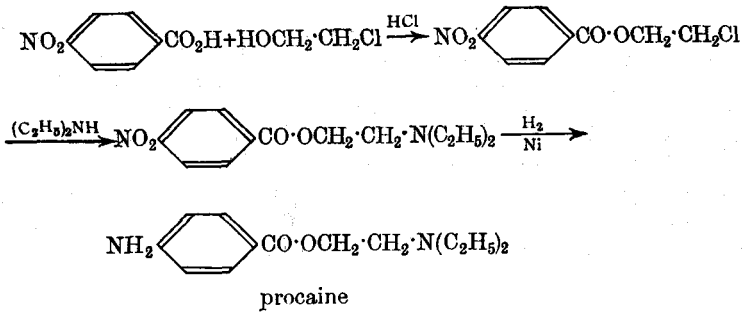
§23b. **Cocaine substitutes.** Cocaine is a very good local anæsthetic, but has certain disadvantages. The anæsthetic properties are lost if either the benzoyl group or the methyl ester group is removed; removal of the *N*-methyl group has no effect. A number of synthetic drugs have now been introduced to replace cocaine as a local anæsthetic; their anæsthetic properties are as good as those of cocaine, and they are less toxic. Two important substitutes are **β-eucaine** and **procaine (novocaine).**

β-Eucaine has been synthesised by treating acetone with ammonia and then treating the product, diacetonamine (see Vol. I), with diethyl acetal. The piperidone thereby produced is then reduced and finally benzoylated to give β-eucaine.



β-eucaine

Procaine has been synthesised from *p*-nitrobenzoic acid.

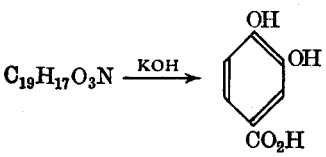


procaine

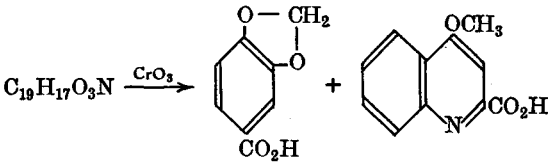
QUINOLINE GROUP

§24. **Angostura alkaloids.** A number of alkaloids have been isolated from angostura bark, *e.g.*, cusparine, galipine, galipoline, etc.

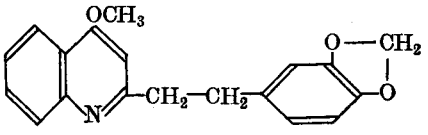
Cusparine, $C_{19}H_{17}O_3N$, m.p. 90–91°, has been shown to contain one methoxyl group (Zeisel method), and when fused with potassium hydroxide, protocatechuic acid is obtained.



On the other hand, controlled oxidation of cusparine gives piperonylic acid and 4-methoxyquinoline-2-carboxylic acid.

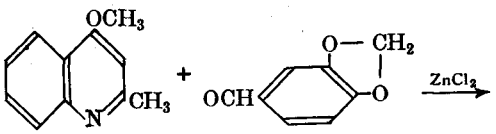


Consideration of this information led to the suggestion of the following structure for cusparine.



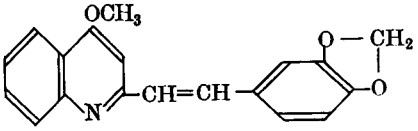
cusparine

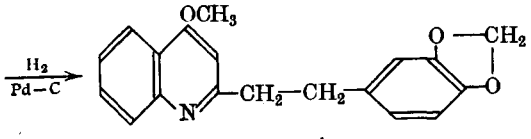
This has been confirmed by synthesis (Späth *et al.*, 1924).



4-methoxy-2-methylquinoline

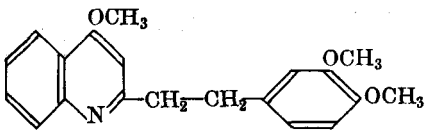
piperonal





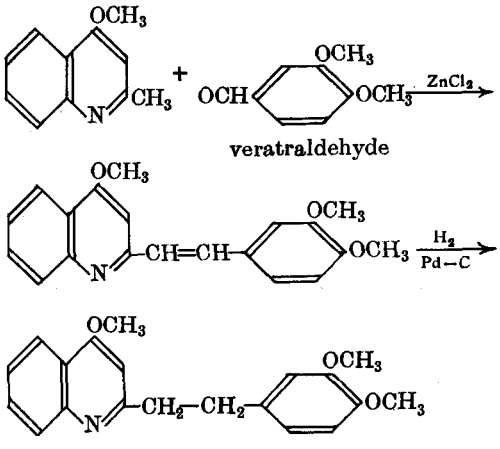
cusparine

Galipine, $C_{20}H_{21}O_3N$, m.p. 113°, contains three methoxyl groups (Zeisel method). When oxidised with chromic acid, galipine produces 4-methoxyquinoline-2-carboxylic acid and veratric acid. Thus the formula of the alkaloid is probably:



galipine

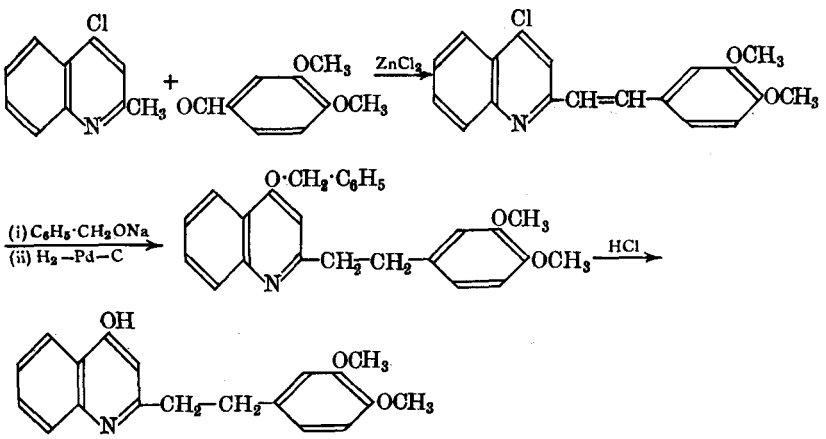
This has been confirmed by synthesis (Späth *et al.*, 1924).



veratraldehyde

galipine

Galipoline, $C_{19}H_{19}O_3N$, m.p. 193°, contains two methoxyl groups and one phenolic group. When methylated with diazomethane, galipoline is converted into galipine. Thus one of the methoxyl groups in the latter is a hydroxyl group in the former. The position of this phenolic hydroxyl was shown to be in the quinoline nucleus by synthesis (Späth *et al.*, 1924).



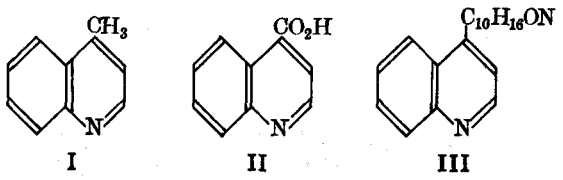
galipoline

§25. Cinchona alkaloids. Cinchonine and quinine, together with many other alkaloids, occur in the bark of various species of *Cinchona*. Cinchonine may be regarded as the parent substance of the cinchona alkaloids,

but quinine is the most important member of this group, its main use being in the treatment of malaria.

§25a. (+)-Cinchonine, $C_{19}H_{22}ON_2$, m.p. 264°, adds on two molecules of methyl iodide to form a di-quaternary compound; thus the alkaloid is a di-tertiary base. Since cinchonine forms a mono-acetate and a mono-benzoate, the molecule contains one hydroxyl group. Furthermore, this hydroxyl group is secondary alcoholic, since on oxidation, cinchonine forms the ketone *cinchoninone*. Cinchonine has been shown to contain one ethylenic double bond by the fact that it adds on one molecule of bromine or halogen acid, and that it is readily catalytically reduced, one molecule of hydrogen being added on.

Fusion of cinchonine with potassium hydroxide gives lepidine (4-methylquinoline), I, and on vigorous oxidation with chromic acid in sulphuric acid solution, cinchoninic acid, II, is obtained (Königs, 1894). Thus cinchonine



I II III

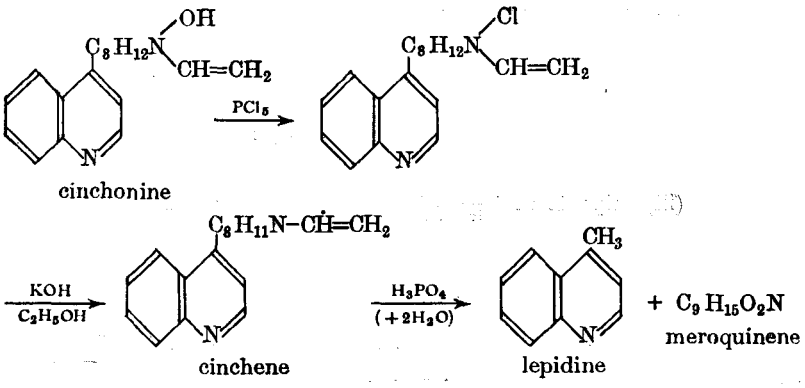
contains a quinoline nucleus with a side-chain in position 4 (III); this side-chain was referred to by Skraup as the " second-half " of the molecule. The hydroxyl group in cinchonine must be in this " second-half ", since if it were not, then a hydroxy derivative or a carboxy derivative (since the hydroxyl is alcoholic) of cinchoninic acid would have been obtained.

Oxidation of cinchonine with permanganate gives cinchotenine and formic acid (Königs, 1879).

$$C_{19}H_{22}ON_2 + 4[O] \xrightarrow{KMnO_4} \underset{\text{cinchotenine}}{C_{18}H_{20}O_3N_2} + H{\cdot}CO_2H$$

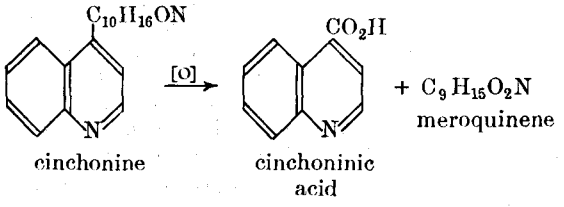
This suggests that there is a $-CH{=}CH_2$ group in the side-chain in the " second-half ".

When treated with phosphorus pentachloride, followed by ethanolic potassium hydroxide, cinchonine is converted into cinchene which, when heated with 25 per cent. phosphoric acid, forms lepidine and a compound Königs named meroquinene (Königs *et al.*, 1884). With the information obtained so far, we may formulate the work of Königs as follows:



cinchonine cinchene lepidine

Meroquinene (meroquinenine) is also obtained, together with cinchoninic acid, when cinchonine is oxidised with chromic acid (Königs, 1894).

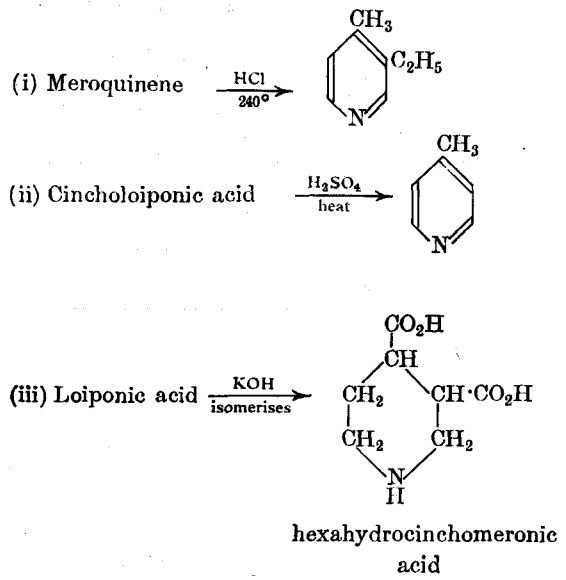


Thus the key to the structure of the " second-half " is the structure of meroquinene. The routine tests showed that meroquinene contains one carboxyl group and one double bond; the presence of the latter indicates that the $—CH=CH_2$ side-chain is still present in meroquinene. Oxidation of meroquinene with cold acid permanganate produces formic acid and cincholoiponic acid, the latter being a dicarboxylic acid (Königs, 1879). The formation of formic acid confirms the presence of the $—CH=CH_2$ side-

$$\underset{\text{meroquinene}}{C_9H_{15}O_2N} \xrightarrow[H_2SO_4]{KMnO_4} \underset{\text{cincholoiponic acid}}{C_8H_{13}O_4N} + H{\cdot}CO_2H$$

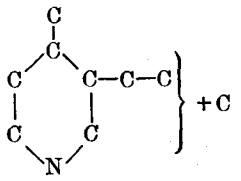
chain in meroquinene. The presence of this group has also been demonstrated by the ozonolysis of meroquinene; formaldehyde is produced (Seekles, 1923). Oxidation of cincholoiponic acid with acid permanganate produces loiponic acid, $C_7H_{11}O_4N$ (Königs, 1890). This is also a dicarboxylic acid, and since it contains one methylene group less than its precursor cincholoiponic acid, this suggests that the latter contains at least a side-chain $—CH_2{\cdot}CO_2H$.

The reactions of the above three acids indicated that they were all secondary bases; that they all contained a piperidine ring is shown by the following reactions.

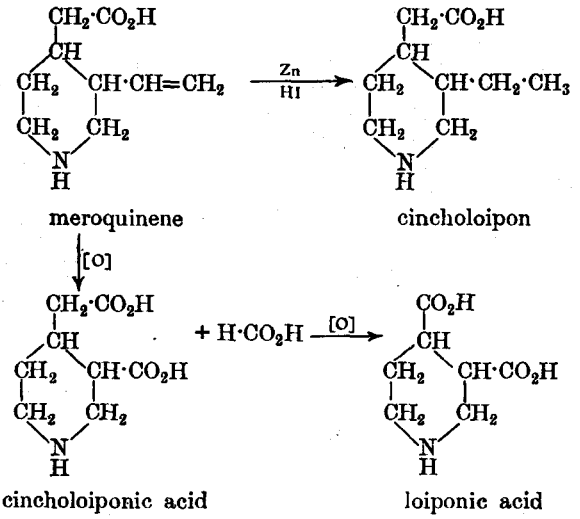


The structure of hexahydrocinchomeronic acid is known from its synthesis (*cf.* §21).

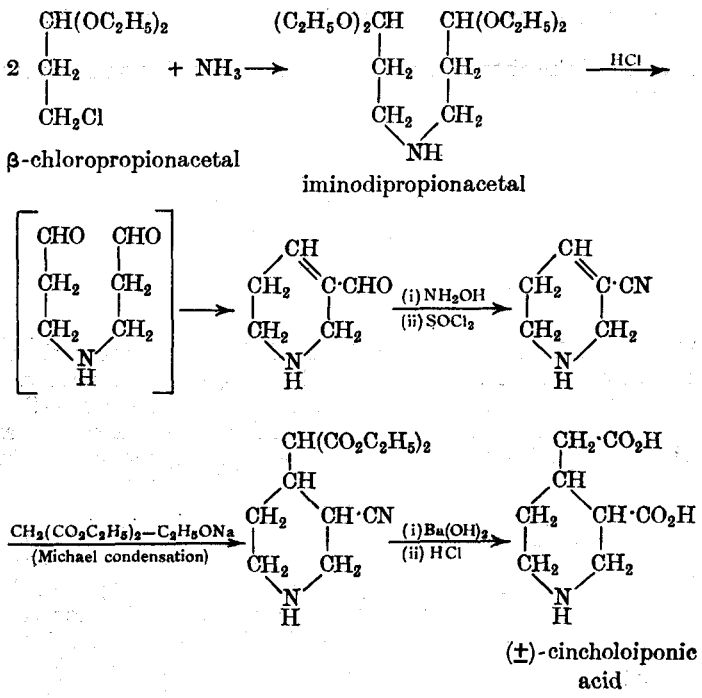
Consideration of the above results shows that a possible skeleton structure of meroquinene is:



The problem then is to find the position of the remaining carbon atom. This carbon atom cannot be an *N*-methyl group, since all three acids are secondary bases. As we have seen, meroquinene contains a $-CH{=}CH_2$ group in the side-chain. A possible position for the extra carbon atom is the side-chain containing this unsaturated group, *i.e.*, the side-chain is an allyl group, $-CH_2{\cdot}CH{=}CH_2$. All the foregoing facts can be explained on this basis, but the following fact cannot, *viz.*, that reduction of meroquinene gives cincholoipon, $C_9H_{17}O_2N$, a compound which contains one carboxyl group and one *ethyl group*. Thus the unsaturated side-chain cannot be allyl (this should have given a propyl group on reduction); the side-chain is therefore vinyl. This leaves only one possible position for the extra carbon atom, *viz.*, 4; this would give a $-CH_2{\cdot}CO_2H$ group at this position, and the presence of such a group has already been inferred (see above). All the reactions of meroquinene can therefore be explained on the following structures:

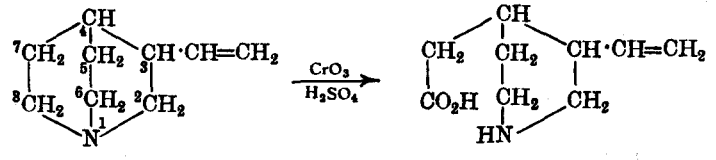


This formula for meroquinene is supported by the synthesis of cincholoiponic acid (Wohl *et al.*, 1907; *cf.* §17) (see next page).



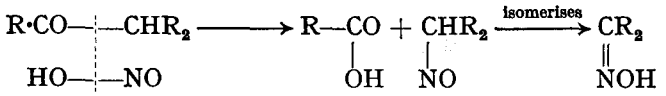
The racemic cincholoiponic acid was acetylated, and then this derivative was resolved by means of brucine; the (+)-form was identical with the acid obtained from meroquinene.

Since meroquinene is obtained from cinchonine by oxidation, the carbon atom of the carboxyl group in meroquinene will be the point of linkage to the " quinoline-half " at which scission of the " second-half " occurs. Since cinchonine is a di-tertiary base, the " second-half " therefore contains a tertiary nitrogen atom. But meroquinene is a *secondary* base, and it therefore follows that in its formation the tertiary nitrogen atom is converted into a secondary nitrogen atom, *a carboxyl group also being produced at the same time.* A possible explanation for this behaviour is that the tertiary nitrogen atom is a part of a bridged ring, one C—N bond being broken when cinchonine is oxidised:

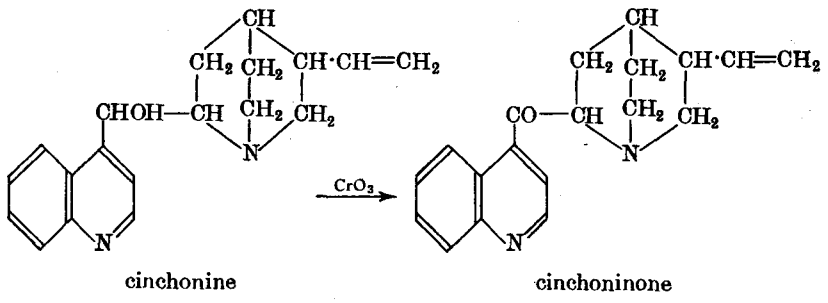


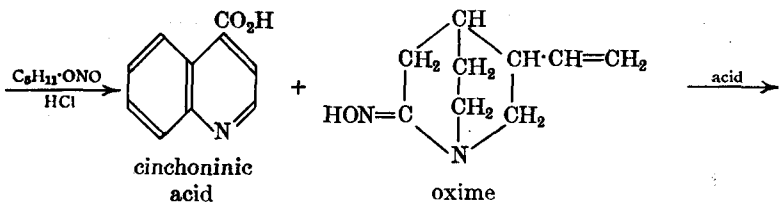
Thus, in cinchonine, the " quinoline-half " must be joined *via* its side-chain at position 4 to the " quinuclidine-half " at position 8. The remaining problem is to ascertain the position of the secondary alcoholic group in the " second-half ". Rabe *et al.* (1906, 1908) converted cinchonine into the ketone cinchoninone by gentle oxidation (chromium trioxide). This ketone, in which both nitrogen atoms are still tertiary, on treatment with amyl

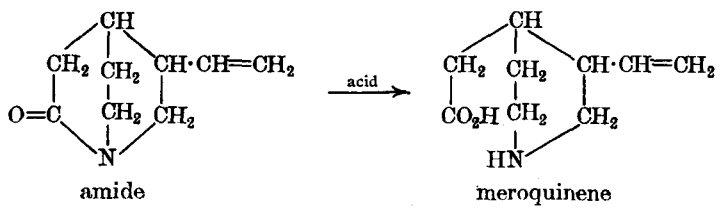
nitrite and hydrogen chloride, gives cinchoninic acid and an oxime. The formation of an acid and an oxime indicates the presence of the group —CO·CH(NO)—, *i.e.*, a methyne group adjacent to a carbonyl group:



The structure of the oxime obtained from cinchoninone was shown to be 8-oximino-3-vinylquinuclidine by its hydrolysis to hydroxylamine and meroquinene. If we assume that the secondary alcoholic group connects the "quinoline-half" to the quinuclidine nucleus, then the foregoing reactions may be written as follows, on the assumption that the structure of cinchonine is as given.

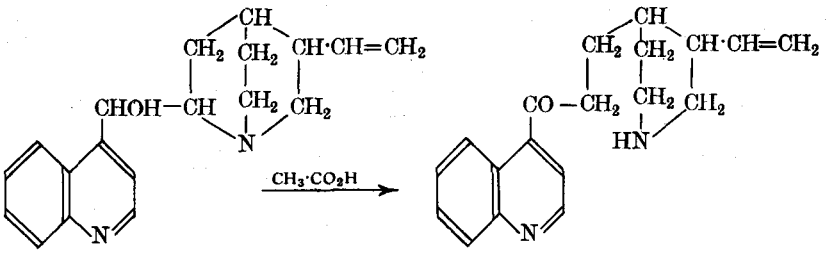






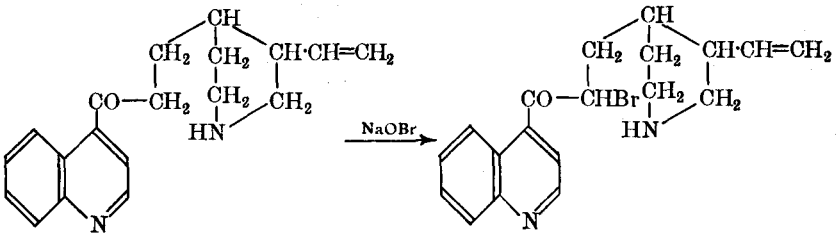
The above structure of cinchonine contains four dissimilar asymmetric carbon atoms, *viz.*, 3, 4, 8, and the carbon atom of the CHOH group (see 3-vinylquinuclidine for numbering). One pair of enantiomorphs is (±)-cinchonine, and another pair is (±)-*cinchonidine*; the configurations of C_3 and C_4 are the same in both, since both give the *same* 8-oximino-3-vinylquinuclidine (see §25b).

A partial synthesis of cinchonine has been carried out by Rabe (1911, 1913). This starts from cinchotoxine, which is prepared by the prolonged action of acetic acid on cinchonine; the latter isomerises (Rabe *et al.*, 1909).

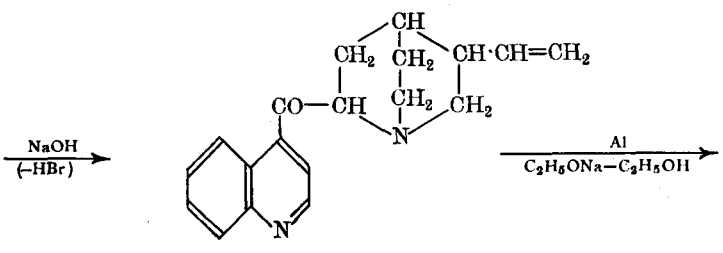


cinchonine cinchotoxine

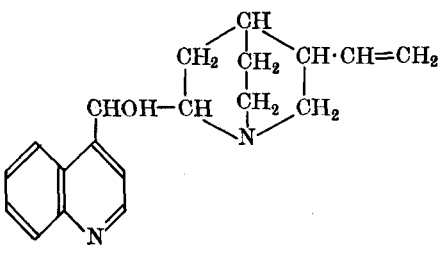
This isomerisation is an example of the *hydramine fission* (see §7). The conversion of cinchotoxine into cinchonine was carried out as follows:



cinchotoxine



cinchoninone



(±)-cinchonine

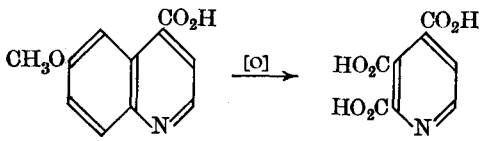
§25b. (—)-Quinine, $C_{20}H_{24}O_2N_2$, m.p. 177°, is used as a febrifuge and as an antimalarial. Since quinine adds on two molecules of methyl iodide to form a di-quaternary salt, it is therefore a di-tertiary base. When heated with hydrochloric acid, quinine eliminates one carbon atom as methyl chloride; therefore there is one methoxyl group present in the molecule. Since quinine forms a mono-acetate and a mono-benzoate, one hydroxyl group must be present, and that this is secondary alcoholic is shown by the fact that oxidation of quinine with chromium trioxide produces quininone, a ketone.

$$\underset{\text{quinine}}{C_{20}H_{24}O_2N_2} \xrightarrow{CrO_3} \underset{\text{quininone}}{C_{20}H_{22}O_2N_2}$$

Quinine also contains one ethylenic double bond, as is shown by the fact that it adds on one molecule of bromine, etc. (*cf.* cinchonine). Oxidation of quinine with chromic acid produces, among other products, quininic acid.

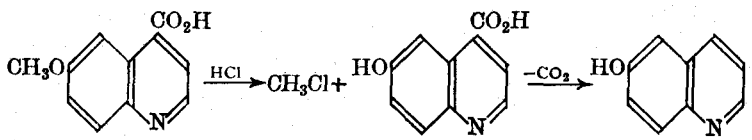
$$\underset{\text{quinine}}{C_{20}H_{24}O_2N_2} \xrightarrow[H_2SO_4]{CrO_3} \underset{\text{quininic acid}}{C_{11}H_9O_3N}$$

On the other hand, controlled oxidation of quinine with chromic acid gives quininic acid and meroquinene. Thus the " second-half " in both quinine and cinchonine is the same, and so the problem is to elucidate the structure of quininic acid. When heated with soda-lime, quininic acid is decarboxylated to a methoxyquinoline, and since, on oxidation with chromic acid, quininic acid forms pyridine-2 : 3 : 4-tricarboxylic acid, the methoxyl group must be a substituent in the benzene ring (of quinoline), and the carboxyl group at position 4 (Skraup, 1881). The position of the methoxyl group was ascertained by heating quininic acid with hydrochloric acid and then



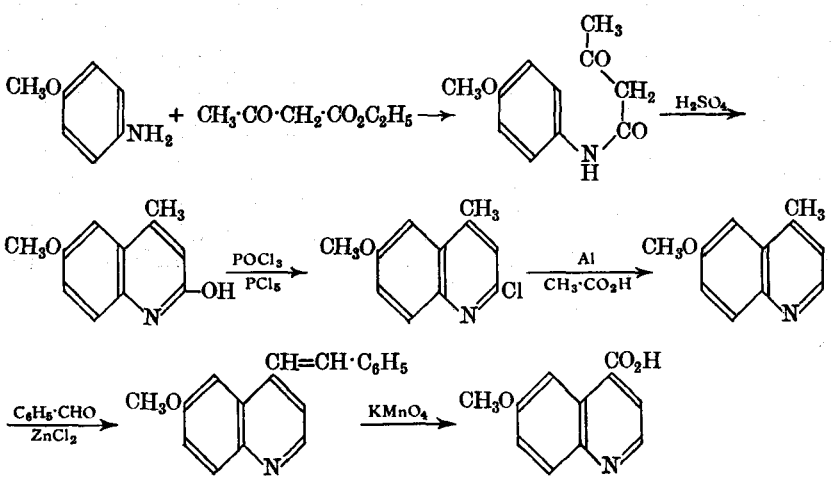
quininic acid — pyridine-2:3:4-tricarboxylic acid

decarboxylating the demethylated product; 6-hydroxyquinoline (a known compound) was obtained. Thus quininic acid is 6-methoxycinchoninic acid.



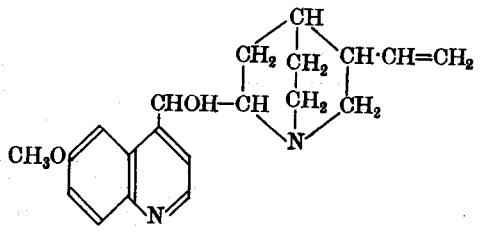
quininic acid — 6-hydroxyquinoline

This structure for quininic acid has been confirmed by synthesis (Rabe *et al.*, 1931).



The direct oxidation of 6-methoxy-4-methylquinoline to quininic acid is extremely difficult; oxidation of the methyl group is accompanied by the oxidation of the benzene ring, the final product being pyridine-2 : 3 : 4-tricarboxylic acid (see §26).

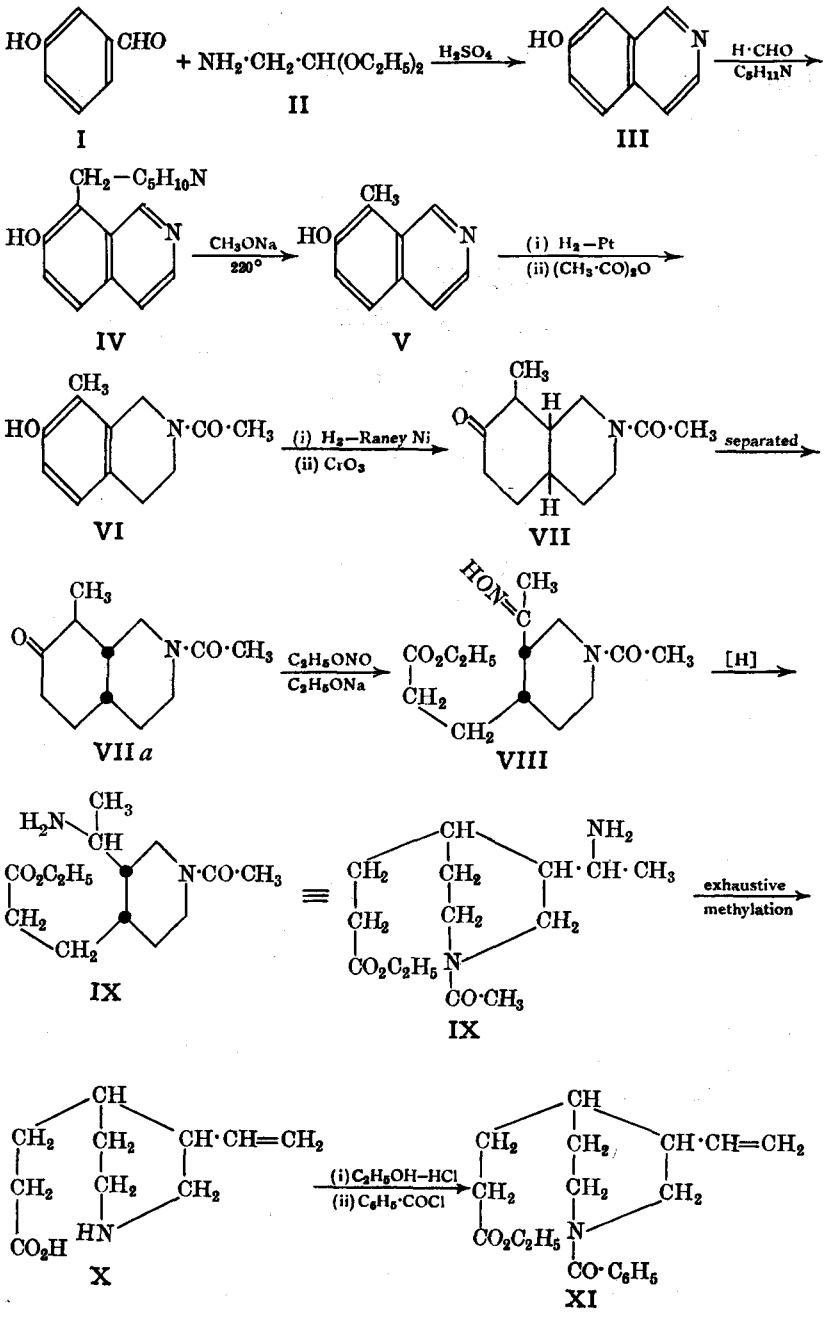
Thus, on the basis of the foregoing evidence, the structure of quinine is:

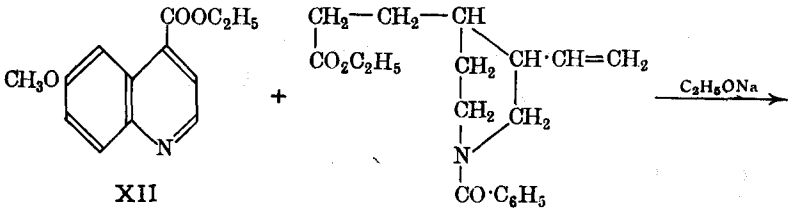


quinine

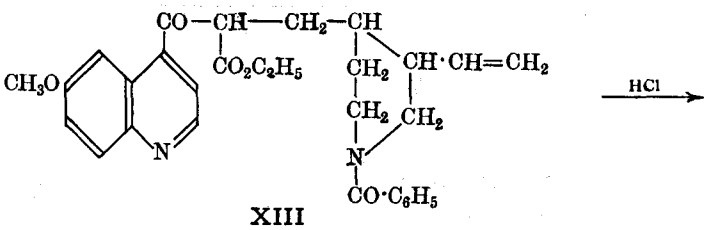
This formula contains the same four asymmetric carbon atoms as cinchonine; thus the same number of pairs of enantiomorphs is possible. One pair is (±)-quinine, and another pair is (±)-*quinidine*; the configurations of C_3 and C_4 are the same in quinine, quinidine, cinchonine and cinchonidine, since all four give the *same* 8-oximino-3-vinylquinuclidine (see §25a).

Rabe *et al.* (1918) carried out a partial synthesis of quinine starting from quinotoxine, which is prepared by heating quinine in acetic acid (*cf.* cinchotoxine). Woodward and Doering (1944) have synthesised (+)-quinotoxine, and so we now have a *total* synthesis of quinine. The following is Woodward and Doering's work up to (+)-quinotoxine, and from this to quinine is Rabe's work. *m*-Hydroxybenzaldehyde (I) is condensed with aminoacetal (II) and the product, 7-hydroxy*iso*quinoline (III), is treated with formaldehyde in methanol solution containing piperidine. The complex formed (IV) is converted into 7-hydroxy-8-methyl*iso*quinoline (V) by heating with methanolic sodium methoxide at 220°. V, on catalytic reduction (platinum) followed by acetylation, gives *N*-acetyl-7-hydroxy-8-methyl-1 : 2 : 3 : 4-tetrahydro*iso*quinoline (VI), which, on further catalytic reduction by heating with a Raney nickel catalyst under pressure and then followed by oxidation with chromium trioxide, is converted into *N*-acetyl-7-keto-8-methyldecahydro*iso*quinoline (VII). VII is a mixture of *cis*- and *trans*-isomers; these were separated and the *cis*-isomer (VII*a*; see §11 vii. IV for conventions) then treated with ethyl nitrite in the presence of sodium ethoxide to give the homomeroquinene derivative VIII. This, on reduction, gives IX, which may now be written more conveniently as shown. Exhaustive methylation of IX, followed by hydrolysis, gives *cis*-(±)-homomeroquinene (X). X, after esterification and benzoylation, gives XI which, on condensation with ethyl quininate (XII), produces XIII. This, on heating with 16 per cent. hydrochloric acid, is hydrolysed and decarboxylated to (±)-quinotoxine (XIV). This was resolved *via* its dibenzoyltartrate (tartaric acid proved unsuccessful for resolution). The conversion of (±)-quinotoxine into quinine had already been accomplished by Rabe *et al.* (the equations for this conversion are also given below).

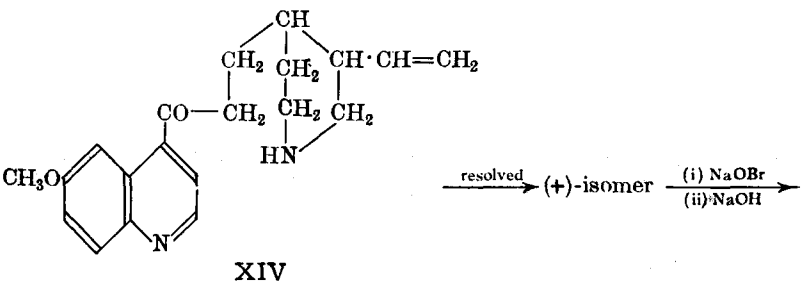

HO
CHO
+ NH2·CH2·CH(OC2H5)2
H2SO4
II
I
HO
N
III
H·CHO
C5H11N
CH2—C5H10N
HO
N
CH3ONa
220°
IV
CH3
HO
N
V
(i) H2—Pt
(ii) (CH3·CO)2O
CH3
HO
N·CO·CH3
VI
(i) H2—Raney Ni
(ii) CrO3
CH3
H
O
N·CO·CH3
H
VII
separated
CH3
O
N·CO·CH3
VII a
C2H5ONO
C2H5ONa
HON=C
CH3
CO2C2H5
N·CO·CH3
CH2
CH2
VIII
[H]
H2N
CH3
CH
CO2C2H5
N·CO·CH3
CH2
CH2
IX
≡
CH
NH2
CH2
CH2
CH·CH·CH3
CH2
CH2
CH2
CO2C2H5
N
CO·CH3
IX
exhaustive
methylation
CH
CH2
CH2
CH·CH=CH2
CH2
CH2
CH2
CO2H
HN
X
(i) C2H5OH–HCl
(ii) C6H5·COCl
CH
CH2
CH2
CH·CH=CH2
CH2
CH2
CH2
CO2C2H5
N
CO·C6H5
XI



XII XI

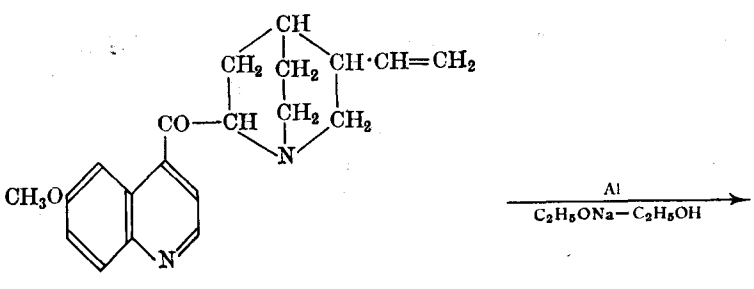


XIII

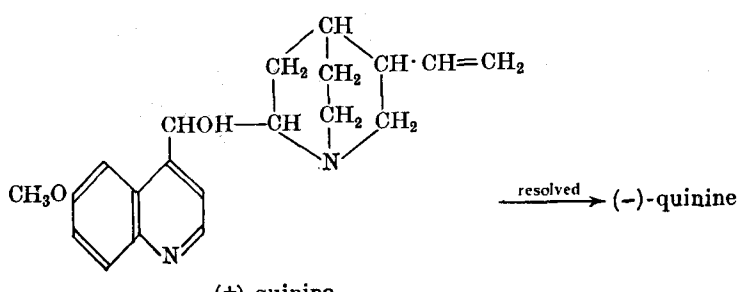


XIV

(±)-quinotoxine



(+)-quininone



(±)-quinine

ISOQUINOLINE GROUP

Opium alkaloids. Many alkaloids have been isolated from opium, and they are divided into two groups according to the nature of their structure:

(i) iso*Quinoline group, e.g.*, papaverine, laudanosine, etc.
(ii) *Phenanthrene group, e.g.*, morphine (see §27).

§26. Papaverine, $C_{20}H_{21}O_4N$, m.p. 147°, is one of the optically inactive alkaloids; it does not contain any asymmetric carbon atom. The structure of papaverine was established by Goldschmiedt and his co-workers (1883–1888). Since papaverine adds on one molecule of methyl iodide to form a quaternary iodide, the nitrogen atom in the molecule is in the tertiary state. The application of the Zeisel method shows the presence of four methoxyl groups; the demethylated product is known as papaveroline.

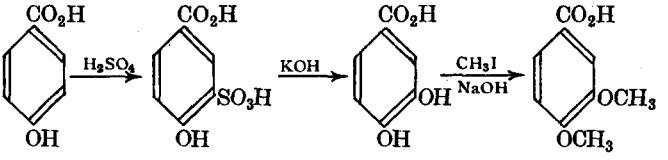
$$\underset{\text{papaverine}}{C_{20}H_{21}O_4N} + 4HI \longrightarrow 4CH_3I + \underset{\text{papaveroline}}{C_{16}H_{13}O_4N}$$

When oxidised with cold dilute permanganate, papaverine is converted into the secondary alcohol papaverinol, $C_{20}H_{21}O_5N$. This, on more vigorous oxidation with hot dilute permanganate, is oxidised to the ketone papaveraldine, $C_{20}H_{19}O_5N$ (it is the formation of this *ketone* that shows that papaverinol is a *secondary* alcohol). Finally, the prolonged action of hot permanganate oxidises papaveraldine to papaverinic acid, $C_{16}H_{13}O_7N$. This acid is a dibasic acid and still contains the keto group present in its precursor—it forms an oxime, etc.; papaverinic acid also contains two methoxyl groups. The foregoing reactions lead to the conclusion that papaverine contains a methylene group.

$$\underset{\text{papaverine}}{(C_{19}H_{19}O_4N)CH_2} \xrightarrow{[O]} \underset{\text{papaverinol}}{(C_{19}H_{19}O_4N)CHOH} \xrightarrow{[O]} \underset{\text{papaveraldine}}{(C_{19}H_{19}O_4N)CO}$$

When oxidised with hot concentrated permanganate, papaverine (or the oxidised products mentioned above) is broken down into smaller fragments, *viz.*, veratric acid, metahemipinic acid, pyridine-2 : 3 : 4-tricarboxylic acid and 6 : 7-dimethoxy*iso*quinoline-1-carboxylic acid. Let us now consider the evidence for the structures of these compounds.

Veratric acid. When decarboxylated, veratric acid forms veratrole. Since this is *o*-dimethoxybenzene, veratric acid is therefore a dimethoxybenzoic acid. The position of the carboxyl group with respect to the two methoxyl groups (in the *ortho*-position) is established by the following synthesis.

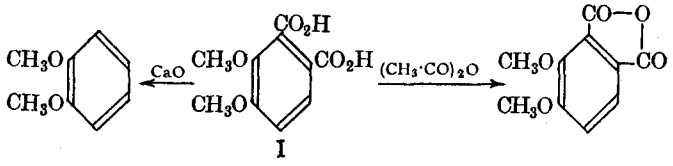


veratric acid

Thus veratric acid is 3 : 4-dimethoxybenzoic acid.

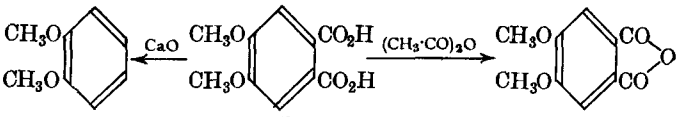
Metahemipinic acid. This is a dicarboxylic acid, and when decarboxylated by heating with calcium oxide, veratrole is formed; thus metahemipinic acid contains two methoxyl groups in the *ortho*-position. Furthermore, since the acid forms an anhydride when heated with acetic anhydride, the two carboxyl groups must be in the *ortho*-position. Thus metahemipinic acid is either I or II. Now metahemipinic acid forms only *one* monoester; II permits the formation of only one monoester, but I can give rise to two

different monoesters. Thus II is metahemipinic acid; I is actually hemipinic acid (this isomer was known before metahemipinic acid).



I

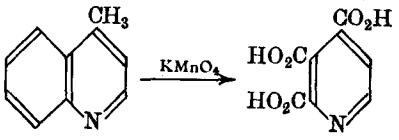
hemipinic acid



II

metahemipinic acid

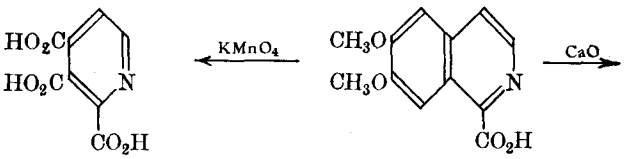
Pyridine-2 : 3 : 4-tricarboxylic acid. The routine tests showed that this contains three carboxyl groups, and since decarboxylation gives pyridine, the acid must be a pyridinetricarboxylic acid. The positions of the three carboxyl groups are established by the fact that this pyridinetricarboxylic acid is produced when lepidine (4-methylquinoline) is oxidised.



lepidine

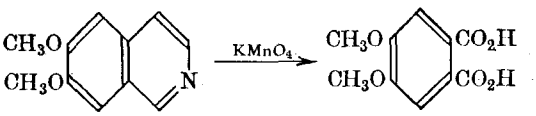
pyridine-2:3:4-tricarboxylic acid

6 : 7-Dimethoxy*iso*quinoline-1-carboxylic acid. The usual tests showed that this compound contains one carboxyl group and two methoxyl groups. On oxidation, this acid forms pyridine-2 : 3 : 4-tricarboxylic acid; when decarboxylated, the acid forms a dimethoxy*iso*quinoline which, on oxidation, gives metahemipinic acid; thus the structure is established.



pyridine-2:3:4-tricarboxylic acid

6:7-dimethoxy-*iso*quinoline-1-carboxylic acid

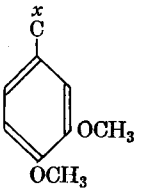


metahemipinic acid

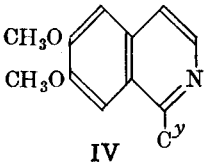
We may now deduce the structure of papaverine as follows:

(i) The isolation of veratric acid indicates the presence of group III in papaverine.

(ii) The isolation of 6 : 7-dimethoxy*iso*quinoline-1-carboxylic acid indicates the presence of group IV in the molecule.



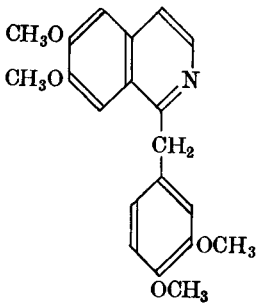
III



IV

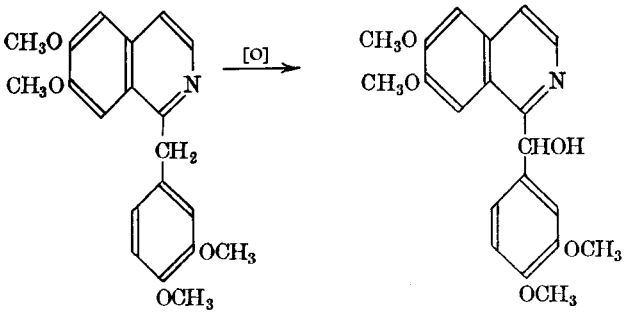
The presence of these two groups also accounts for the isolation of the other two fragments.

(iii) The total carbon content of III (9 carbon atoms) and IV (12 carbon atoms) is 21 carbon atoms. But papaverine contains only 20. There is, however, a $-CH_2-$ group present, and if we assume that C^x and C^y are one and the same carbon atom, *viz.*, the carbon atom of the CH_2 group, then the following structure of papaverine accounts for all the facts:

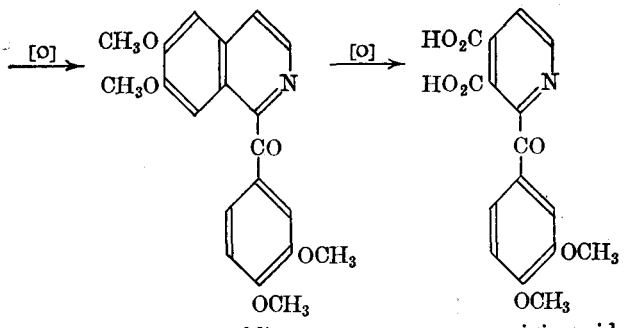


papaverine

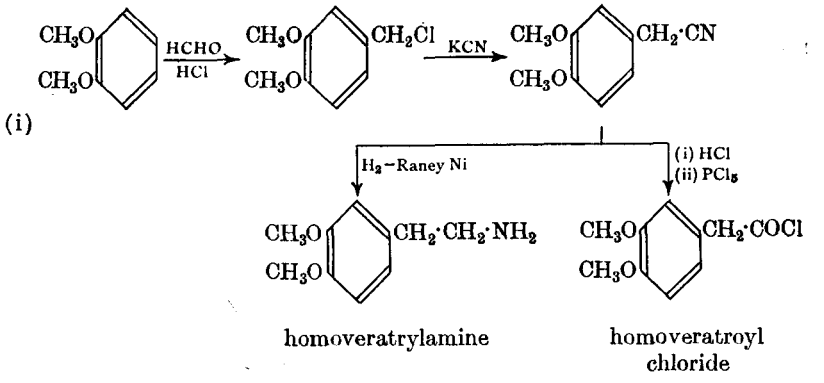
Thus, with this formula, we can formulate the oxidation of papaverine as follows:

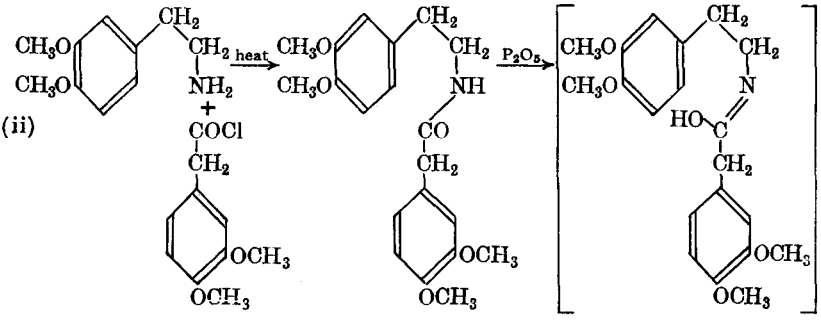


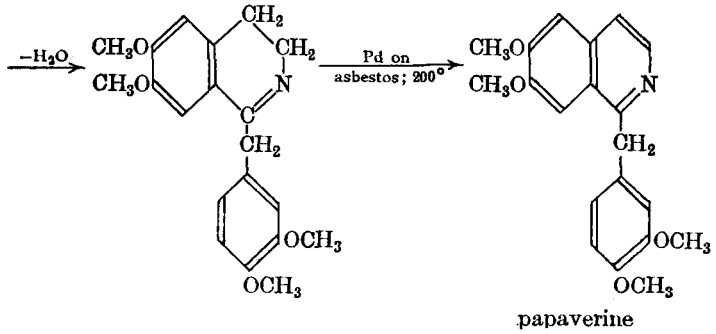
papaverine papaverinol



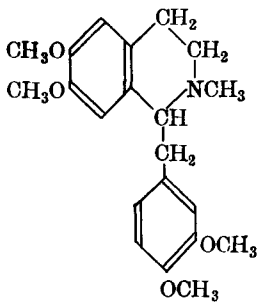
This structure for papaverine has been confirmed by synthesis. The first synthesis was by Pictet and Gams (1909), but Bide and Wilkinson (1945) carried out a simpler one, and it is this that is described here.



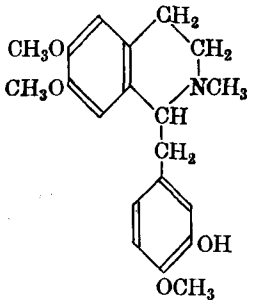




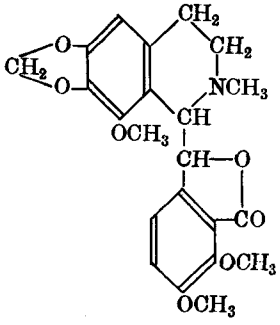
§26a. Some other alkaloids of the *iso*quinoline group are:



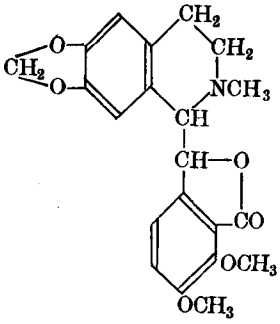
laudanosine



laudanine



narcotine



hydrastine

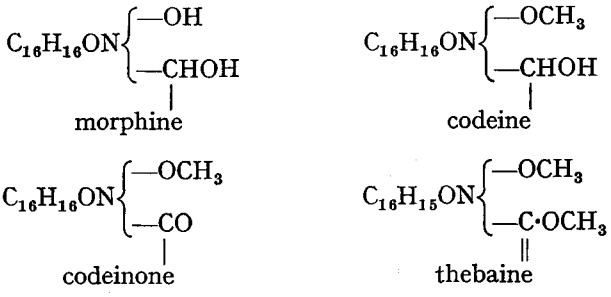
PHENANTHRENE GROUP

§27. Morphine, codeine and thebaine. These are three important opium alkaloids which contain the phenanthrene nucleus.

(−)-*Morphine*, $C_{17}H_{19}O_3N$, m.p. 254°, is the chief alkaloid in opium, and was the first alkaloid to be isolated (Sertürner, 1806). The usual tests show that the nitrogen atom is in the tertiary state, and since morphine forms a diacetate and a dibenzoate, two hydroxyl groups are therefore present in the molecule. Morphine gives the ferric chloride test for phenols, and dissolves in aqueous sodium hydroxide to form a *monosodium* salt, and this is reconverted into morphine by the action of carbon dioxide; thus *one* of the hydroxyl groups is phenolic (Matthiessen *et al.*, 1869). The second hydroxyl group is secondary alcoholic, as is shown by the following reactions. Halogen acids convert morphine into a monohalogeno derivative, one hydroxyl group being replaced by a halogen atom. When heated with methyl iodide in the presence of aqueous potassium hydroxide, morphine is methylated to give (−)-*codeine*, $C_{18}H_{21}O_3N$, m.p. 155° (Grimaux, 1881). Since codeine is no longer soluble in alkalis, it therefore follows that it is only the *phenolic* hydroxyl group in morphine that has been methylated. Furthermore, codeine can be oxidised by chromic acid to *codeinone*, a ketone (Hesse, 1884). Thus the hydroxyl group in codeine (and this one in morphine) is secondary alcoholic, and so codeine is the monomethyl (phenolic) ether of morphine.

(−)-*Thebaine*, $C_{19}H_{21}O_3N$, m.p. 193°, produces two molecules of methyl iodide when heated with hydriodic acid (Zeisel method); hence thebaine is a dimethoxy derivative. When heated with sulphuric acid, thebaine

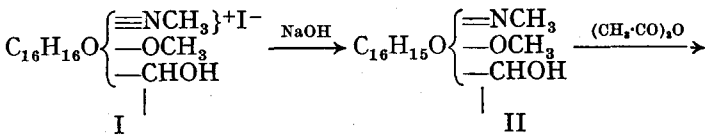
eliminates one methyl group as methyl hydrogen sulphate, and forms codeinone (Knorr, 1906). The formation of a *ketone* led Knorr to suggest that thebaine is the methyl ether of the *enolic* form of codeinone. The foregoing work can thus be summarised by assigning the following formulæ to the compounds described:



So far, we have accounted for the functional nature of two of the oxygen atoms; the unreactivity of the third oxygen atom suggests that it is probably of the ether type (Vongerichten, 1881).

All three alkaloids are tertiary bases (each combines with one molecule of methyl iodide to form a methiodide). When heated with hydrochloric acid at 140° under pressure morphine loses one molecule of water to form *apomorphine*, $C_{17}H_{17}O_2N$. Codeine, under the same conditions, also gives apomorphine (and some other products). Thebaine, however, when heated with dilute hydrochloric acid, forms *thebenine*, $C_{18}H_{19}O_3N$ (a secondary base), and with concentrated hydrochloric acid, morphothebaine, $C_{18}H_{19}O_3N$ (a tertiary base). Thus in the formation of thebenine from thebaine, a tertiary nitrogen atom is converted into a secondary one. For this change to occur, the tertiary nitrogen must be of the type $>N\cdot R$, where the nitrogen is in a ring system; had the nitrogen been in the group $—NR_2$, then the formation of a *primary* base could be expected.

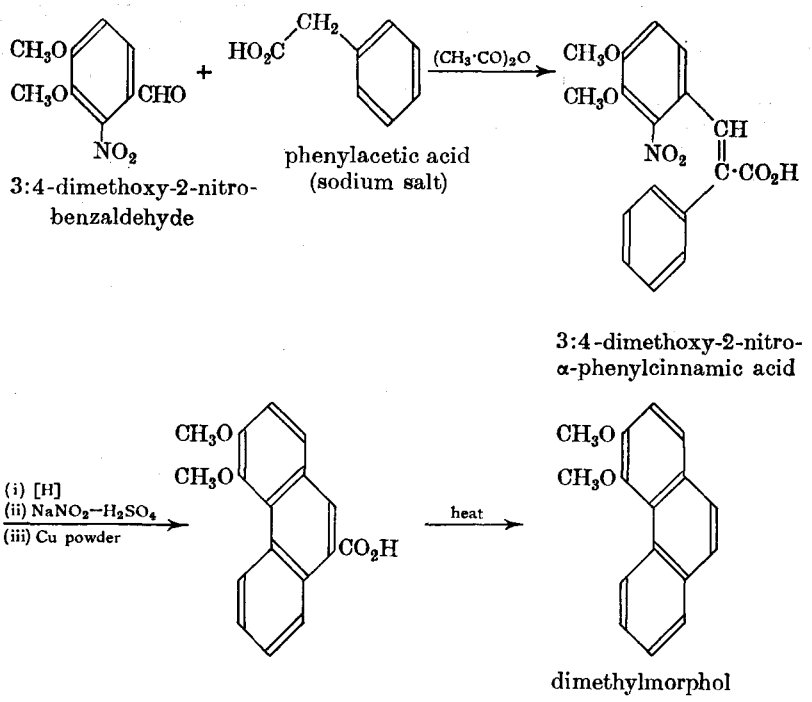
When morphine is distilled with zinc dust, phenanthrene and a number of bases are produced (Vongerichten *et al.*, 1869). This suggests that a phenanthrene nucleus is probably present, and this has been confirmed as follows. When codeine methiodide, I, is boiled with sodium hydroxide solution, α-methylmorphimethine, II, is obtained and this, on heating with acetic anhydride, forms methylmorphol, III, and ethanoldimethylamine, IV (some of II isomerises to β-methylmorphimethine).



$$\underset{\text{III}}{C_{15}H_{12}O_2} + \underset{\text{IV}}{(CH_3)_2N\cdot CH_2\cdot CH_2OH}$$

The structure of methylmorphol (III) was ascertained by heating it with hydrochloric acid at 180° under pressure; methyl chloride and a dihydroxyphenanthrene, *morphol*, were obtained. Oxidation of diacetylmorphol gives a diacetylphenanthraquinone; thus positions 9 and 10 are free. On further oxidation (permanganate), the quinone is converted into phthalic acid; therefore the two hydroxyl groups are in the same ring. Since the fusion of morphine with alkali gives protocatechuic acid, this shows that both

hydroxyl groups in morphol are in the *ortho*-position. Finally, Pschorr *et al.* (1900) showed by synthesis that dimethylmorphol is 3 : 4-dimethoxyphenanthrene (*cf.* Pschorr synthesis, §2 via. X).



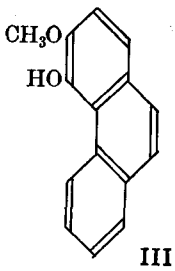
3:4-dimethoxy-2-nitrobenzaldehyde

phenylacetic acid (sodium salt)

3:4-dimethoxy-2-nitro-α-phenylcinnamic acid

dimethylmorphol

Then Pschorr *et al.* (1902) synthesised methylmorphol (III), and showed it to be 4-hydroxy-3-methoxyphenanthrene (in this synthesis Pschorr used 3-acetoxy-4-methoxy-2-nitrobenzaldehyde).



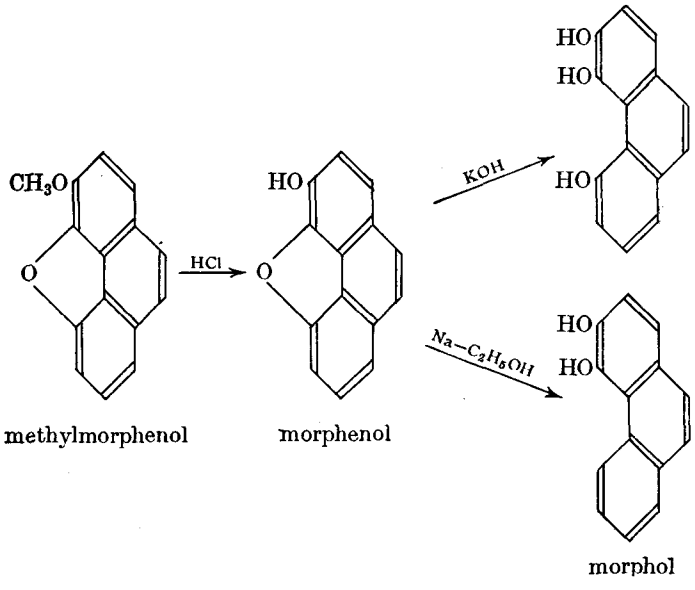
III

methylmorphol

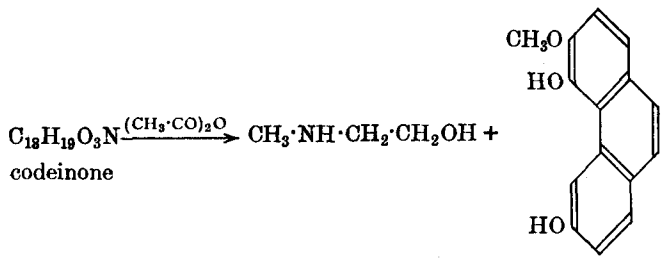
The formation of ethanoldimethylamine (IV) from α-methylmorphimethine indicates that there is a $>NCH_3$ group in codeine (only *one* methyl iodide molecule adds to codeine to form codeine methiodide; it has also been shown above that this nitrogen is in a heterocyclic ring).

When β-methylmorphimethine is heated with water, the products obtained are trimethylamine, ethylene and *methylmorphenol* (Vongerichten, 1896). Demethylation of this compound with hydrochloric acid produces *morphenol*, a compound which contains one phenolic hydroxyl group and an inert

oxygen atom. On fusion with potassium hydroxide, morphenol gives 3 : 4 : 5-trihydroxyphenanthrene (Vongerichten *et al.*, 1906). The structure of this compound was shown by the synthesis of 3 : 4 : 5-trimethoxyphenanthrene, which was found to be identical with the product obtained by methylating the trihydroxyphenanthrene obtained from morphenol (Pschorr *et al.*, 1912). Furthermore, the reduction of morphenol with sodium and ethanol gives morphol (Vongerichten, 1898). These results can be explained by assuming that morphenol has a structure containing an ether linkage in positions 4 : 5 (of the phenanthrene nucleus).

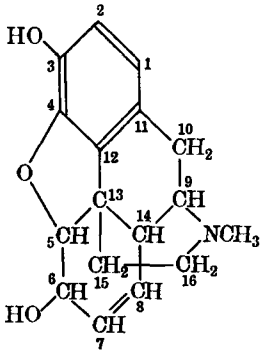


Codeinone, on heating with acetic anhydride, gives ethanolmethylamine and the diacetyl derivative of 4 : 6-dihydroxy-3-methoxyphenanthrene.

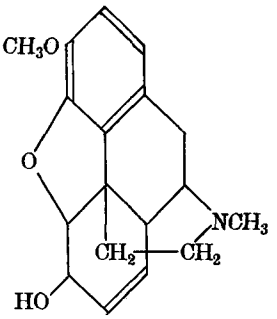


The position 3 of the methoxyl group and the position 4 of the hydroxy group have already been accounted for; the hydroxyl group in the 6-position must therefore be produced from the oxygen of the keto group in codeinone.

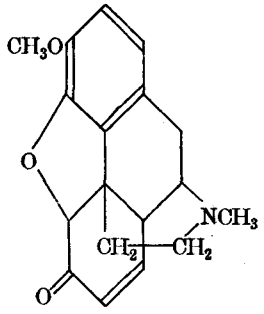
Based on the foregoing evidence, and a large amount of other experimental work, Gulland and Robinson (1923, 1925) have proposed the following structures.



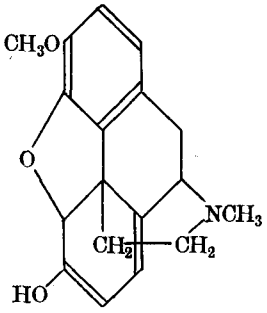
morphine



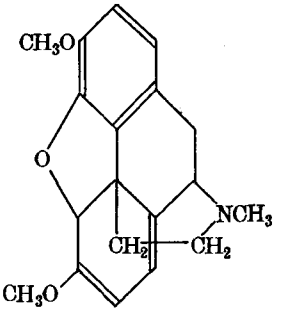
codeine



codeinone
(*keto form*)



codeinone
(*enol form*)



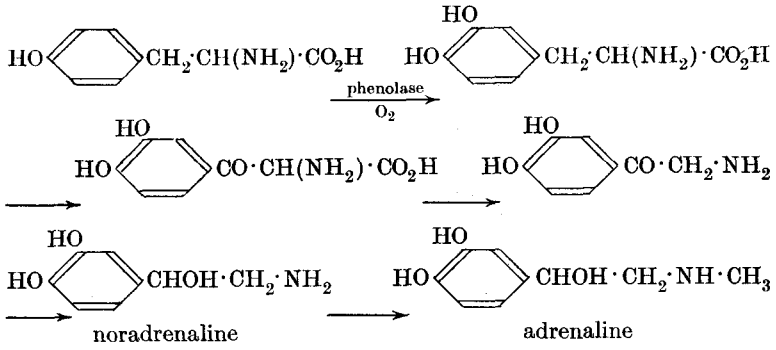
thebaine

Gates *et al.* (1956) have now synthesised morphine.

§28. Biosynthesis of alkaloids. As more and more structures of alkaloids were elucidated, it became increasingly probable that the precursors in the biosynthesis of alkaloids were amino-acids and amino-aldehydes and amines derived from them. A particularly interesting point is that the consideration of biosynthesis has led to deductions in structure, *e.g.*, Woodward (1948) proposed a biosynthesis of strychnine, and from this Robinson (1948) deduced the structure of emetine which was later confirmed by the synthetic work of Battersby *et al.* (1950).

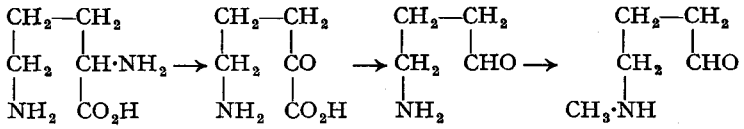
We have already seen (§18. XIII) how keto-acids may be converted into amino-acids, and *vice versa*. There are also enzymes which bring about the decarboxylation of amino-acids to amines and the decarboxylation of α-keto-acids to aldehydes. Thus amino-acids, amines and amino-aldehydes, together with formaldehyde (or its equivalent) are believed to be the units involved in the biosynthesis of alkaloids. The general technique has been to administer labelled precursors to plants and to isolate the alkaloid after some time has elapsed for the growth of the plant.

The following examples of biosynthesis illustrate the principles outlined above. Alkaloids containing a benzene ring are believed to be products of the shikimic acid route (§18. XIII); the amino-acids phenylalanine and tyrosine are the starting points for the biosynthesis of, *e.g.*, ephedrine, hordenine, mezcaline, etc. As an example, we may describe the biosynthesis of adrenaline (§12) from tyrosine; the route is possibly:

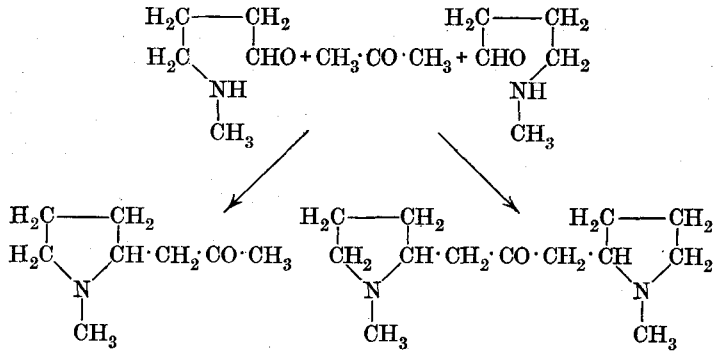


Leete *et al.* (1952–) have shown, using labelled compounds, that phenylalanine, tyrosine and 3,4-dihydroxyphenylalanine are precursors for the alkaloids of the phenylalanine and *iso*quinoline groups (see also later).

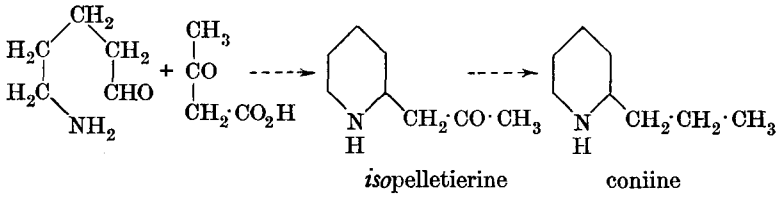
A study of the formulæ of hygrine (§13) and cuscohygrine (§13a) shows that the two most reasonable units are acetone and pyrrolidine. The biosynthesis of acetone occurs *via* acetoacetic acid (see §32a. VIII), but the precursor of the pyrrolidine fragment is less certain. The most likely amino-acid precursor appears to be ornithine, which could undergo the following reactions to give 4-methylaminobutanal (see also later):



This compound may then be imagined to condense with acetone (or acetoacetic acid) to form hygrine and cuscohygrine (*cf.* §§13, 13a).

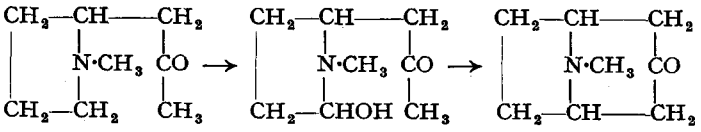


In the same way, the pelletierine group of alkaloids (§19) may all be imagined to be formed from 5-aminopentanal, *e.g.*, Anet *et al.* (1949) have condensed this aldehyde with acetoacetic acid at *p*H 11 to give *iso*pelletierine; and 5-methylaminopentanal with acetoacetic acid at *p*H 7 to give methyl*iso*pelletierine. The amino-acid precursor of 5-aminopentanal is most likely lysine (the homologue of ornithine). It should also be noted that conversion of the keto group in *iso*pelletierine into a methylene group gives coniine:



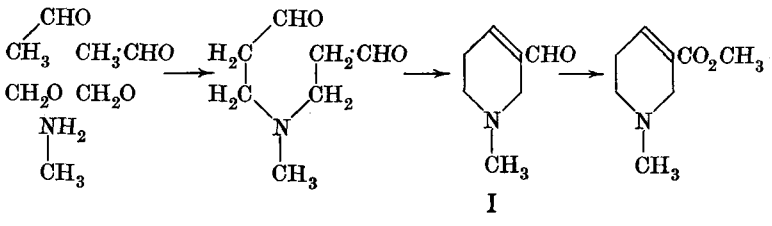
*iso*pelletierine coniine

Now let us consider tropinone. Since this compound contains the hygrine skeleton, one possible mode of biosynthesis of tropinone could be *via* hygrine as the precursor:



On the other hand, tropinone has been synthesised from succinaldehyde, methylamine and acetonedicarboxylic acid under physiological conditions (§22). In this case, the problem is the nature of the precursor of succinaldehyde. Glutamic acid is one possibility, and succinic acid is another. The biosynthesis of cocaine (§23) is similar to that of tropinone.

The biosynthesis of some alkaloids containing a piperidine ring has already been discussed. Mannich (1942) has suggested that arecoline (§17) is formed as follows:



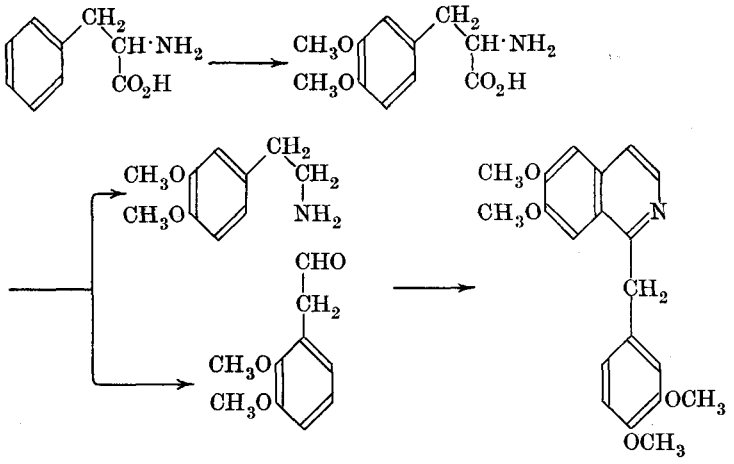
I

Mannich obtained I by carrying out the condensation with a mixture of acetaldehyde, formaldehyde and methylamine at room temperature at *p*H 3.

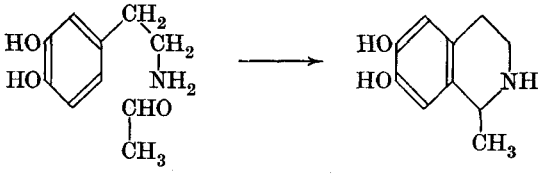
Leete (1955–1958) has shown, using labelled ornithine, that this amino-acid is a good precursor for the pyrrolidine ring in nicotine, and has also suggested that putrescine, glutamic acid and proline are incorporated into the pyrrolidine ring, but are less efficient precursors than ornithine. Marion *et al.* (1954) have also shown that labelled ornithine is incorporated into hyoscyamine (§22). Kaczkowski *et al.* (1960), using labelled compounds, have found that acetate is incorporated into the tropane ring in hyoscyamine, possibly *via* acetoacetate. Leete (1960) has shown that phenylalanine is a precursor of tropic acid.

The origin of the pyridine ring is still obscure. Some suggestions have been described above. It appears that alanine and aspartic acid are precursors of nicotinic acid, and experiments using tritium-labelled nicotinic acid support the hypothesis that it is converted into nicotine *via* a 6-pyridone derivative (Dawson *et al.*, 1958).

It has been pointed out above that phenylalanine, etc. are precursors for the *iso*quinoline alkaloids. Thus, *e.g.*, papaverine (§26) might possibly undergo biosynthesis as follows:



Support for the plausibility of this mechanism is given, *e.g.*, by the formation of the tetrahydro*iso*quinoline from the condensation between 3 : 4-dihydroxyphenylethylamine and acetaldehyde at *p*H 3–5 (Schöpf *et al.*, 1934).



Rapoport *et al.* (1960), using labelled carbon dioxide (^{14}C), have shown that the primary product of synthesis in the morphine alkaloids is apparently thebaine, which is later converted into codeine and morphine.

READING REFERENCES

Henry, *The Plant Alkaloids*, Churchill (1949, 4th ed.).

Gilman (Ed.), *Advanced Organic Chemistry*, Wiley (1943, 2nd ed.). Vol. II. Ch. 15. Alkaloids.

Cook (Ed.), *Progress in Organic Chemistry*, Butterworth. Vol. I (1952). Ch. I. Molecular Structure of Strychnine, Brucine and Vomicine. Vol. III (1955). Ch. 5. Indole Alkaloids.

Manske and Holmes (Ed.), *The Alkaloids*, Academic Press. (Vol. I, 1950; —.)

Bergel and Morrison, Synthetic Analgesics, *Quart. Reviews* (*Chem. Soc.*), 1948, **2**, 349.

Stern, Synthetic Approaches to the Morphine Structure, *Quart. Reviews* (*Chem. Soc.*), 1951, **5**, 405.

Gates and Tschudi, The Synthesis of Morphine, *J. Amer. Chem. Soc.*, 1956, **78**, 1380.

McKenna, Steroidal Alkaloids, *Quart. Reviews* (*Chem. Soc.*), 1953, **7**, 231.

Bentley, *The Chemistry of the Morphine Alkaloids*, Oxford Press (1954).

Bentley, *The Alkaloids*, Interscience Publishers (1957).

Glenn, The Structure of the Ergot Alkaloids, *Quart. Reviews* (*Chem. Soc.*), 1954, **8**, 192.

Saxton, The Indole Alkaloids Excluding Harmine and Strychnine, *Quart. Reviews* (*Chem. Soc.*), 1956, **10**, 108.

Sir Robert Robinson, *The Structural Relations of Natural Products*, Oxford Press (1955).

Morgan and Barltrop, Veratrum Alkaloids, *Quart. Reviews* (*Chem. Soc.*), 1958, **12**, 34.

Rodd (Ed.), *Chemistry of the Carbon Compounds*, Elsevier. Vol. IVC (1960). Alkaloids, Chh. XXIII–XXIX.

Sangster, Determination of Alkaloid Structures, *J. Chem. Educ.*, 1960, **37**, 454, 518.

Battersby, Alkaloid Biosynthesis, *Quart. Reviews* (*Chem. Soc.*), 1961, **15**, 259.

Huisgen, Richard Willstätter, *J. Chem. Educ.*, 1961, **38**, 259.

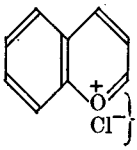
Ray, Alkaloids—the World's Pain Killers, *J. Chem. Educ.*, 1960, **37**, 451.

CHAPTER XV

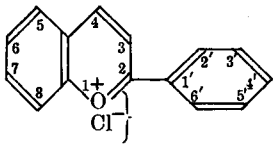
ANTHOCYANINS

§1. Introduction. Anthocyanins are natural plant pigments; they are glycosides and their aglycons, *i.e.*, the sugar-free pigments, are known as the *anthocyanidins*. The anthocyanins, which are water-soluble pigments, generally occur in the aqueous cell-sap, and are responsible for the large variety of colours in flowers; red—violet—blue. Willstätter *et al.* (1913–) showed that the various shades of colour exhibited by all flowers are due to a very small number of different compounds. Furthermore, these different compounds were shown to contain the same carbon skeleton, and differed only in the nature of the substituent groups. The anthocyanin pigments are amphoteric; their acid salts are usually red, their metallic salts usually blue and in neutral solution the anthocyanins are violet (see also §5).

§2. General nature of the anthocyanins. The fundamental nucleus in anthocyanidins is benzopyrylium chloride, but the parent compound is 2-phenylbenzopyrylium chloride or **flavylium chloride.** (The formulæ are now usually written with the oxygen atom at the top, *i.e.*, the formulæ



benzopyrylium chloride



flavylium chloride

shown are turned upside down; there is no change in numbering.) All anthocyanidins are derivatives of 3 : 5 : 7-trihydroxyflavylium chloride. The following table on page 546 shows some common anthocyanidins (as chlorides).

Various sugars have been found in anthocyanins; the most common are glucose, galactose and rhamnose, and the most important of these is glucose, which occurs as the diglucoside. Some pigments, as well as being glycosides, are also acylated derivatives, two common acids being *p*-hydroxybenzoic acid and malonic acid. The acid radical may be attached either to a phenolic hydroxyl group in the flavylium nucleus or to a hydroxyl group in the sugar residue.

A number of qualitative tests have been introduced to identify the various anthocyanins without actually isolating them (Robinson *et al.*, 1931–1933, 1938); *e.g.*,

(i) The pigment is extracted with amyl (pentyl) alcohol in the presence of sodium acetate containing a trace of ferric chloride; cyanidin gives a blue colour, delphinidin a less intense blue colour, and the others still less colour or no colour at all.

(ii) A dilute sodium hydroxide solution of the pigment is shaken with air; delphinidin (and petunidin) is decolorised and the others are not.

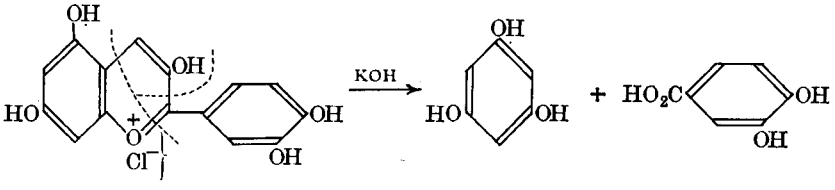
(iii) More recently chromatographic analysis has been used to identify anthocyanins (see also §5).

(iv) The spectra of the anthocyanins in the region 5000–5500 A are similar, but Geissman *et al.* (1953) have shown that the addition of aluminium chloride to solutions of certain anthocyanins shifts the absorption maximum. Only

Aglycon		Occurrence
Trivial name	Chemical name	
Pelargonidin .	3 : 4′ : 5 : 7-Tetrahydroxyflavylium chloride	Present in orange-red to scarlet flowers, *e.g.*, scarlet *Pelargonium*, orange-red dahlia.
Cyanidin . .	3 : 3′ : 4′ : 5 : 7-Pentahydroxyflavylium chloride	Present in crimson to bluish-red flowers, *e.g.*, deep red dahlia, red roses, blue cornflower.
Delphinidin .	3 : 3′ : 4 : 5 : 5′ : 7-Hexahydroxyflavylium chloride	Present in violet to blue flowers, *e.g.*, Delphinium.
Peonidin . .	3 : 4′ : 5 : 7-Tetrahydroxy-3′-methoxyflavylium chloride	Present in flowers less blue than the Cyanidin group, *e.g.*, red peony.
Malvidin (Syringidin)	3 : 4′ : 5 : 7-Tetrahydroxy-3′ : 5′-dimethoxyflavylium chloride	Present in flowers less blue than the Delphinidin group, *e.g.*, *Primula viscosa*.
Hirsutidin . .	3 : 4′ : 5-Trihydroxy-3′ : 5′ : 7-trimethoxyflavylium chloride	Present in *Primula hirsuta*.

anthocyanins with the 3′ : 4′-dihydroxyl groups *free* show this shift, and so this observation may offer a method for analysing anthocyanin mixtures.

§3. Structure of the anthocyanidins. The anthocyanin is first hydrolysed with hydrochloric acid and the anthocyanidin is then isolated as the chloride. The usual analytical methods are applied to determine the number of hydroxyl and methoxyl groups present in the molecule. The structure of the anthocyanidin is ascertained by the nature of the products obtained by fusing the anthocyanidin with potassium hydroxide (Willstätter *et al.*, 1915); phloroglucinol or a methylated phloroglucinol and a phenolic acid are always obtained, *e.g.*, cyanidin chloride gives phloroglucinol and protocatechuic acid.

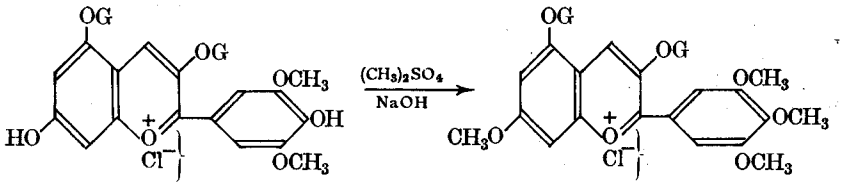


cyanidin chloride

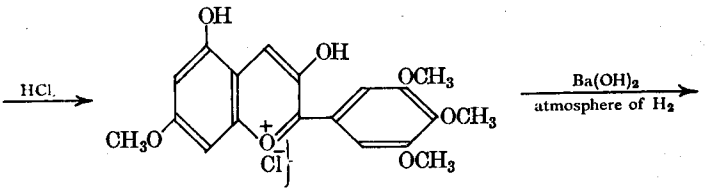
This method suffers from the disadvantage that the fusion (or boiling with concentrated potassium hydroxide solution) not only degrades the anthocyanidin, but also often demethylates it at the same time. Thus the positions of the methoxyl groups in the original compound are now rendered uncertain. This difficulty was overcome by Karrer *et al.* (1927), who degraded the anthocyanidin with a 10 per cent. solution of barium hydroxide or sodium hydroxide in an atmosphere of hydrogen; in this way, the methoxyl groups are left intact.

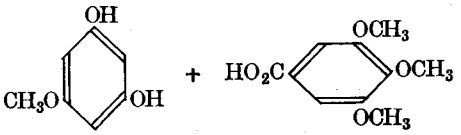
The next problem is to ascertain the positions of the sugar residues.

(i) Karrer *et al.* (1927) methylated the anthocyanin, then removed the sugar residues by hydrolysis (hydrochloric acid), and finally hydrolysed with barium hydroxide solution in an atmosphere of hydrogen; the positions of the *free* hydroxyl indicate the points of attachment of the sugar residues. In some cases, however, interpretation of the results is uncertain, *e.g.* (G represents a sugar residue):



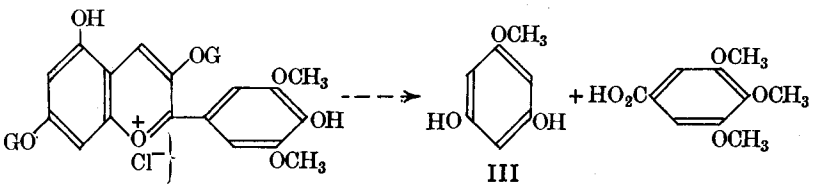
I II





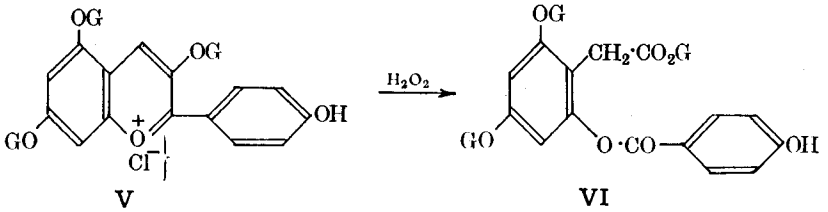
III

The problem is: Which of the two hydroxyl groups in monomethylphloroglucinol was originally attached to G? The above results do not lead to a definite answer, since had the structure of the anthocyanin been IV instead of I, III would still have been obtained:



IV III

(ii) Hydrogen peroxide (15 per cent.) attacks anthocyanins as follows (Karrer *et al.*, 1927):



V VI

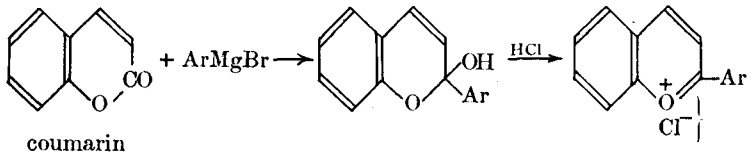
If the anthocyanin, V, has a glucose residue in the 3-position, then *this* glucose residue in VI is readily hydrolysed by dilute ammonia. If the glucose residue in V is in either the 5- or 7-position, then this glucose residue in VI is removed only by heating with dilute hydrochloric acid. Thus position 3 can be distinguished from positions 5 or 7, but the latter two cannot be distinguished from each other.

(iii) Anthocyanins with a free hydroxyl group in the 3-position are very readily oxidised by ferric chloride; the anthocyanins are rapidly decolorised in this oxidation (Robinson *et al.*, 1931).

Conclusive evidence for the positions of the sugar residues is afforded by the synthesis of the anthocyanins (see, *e.g.*, cyanin, §5). In general, it has been found that glucose residues are linked at positions 3 or 3 : 5.

§4. General methods of synthesising the anthocyanidins.

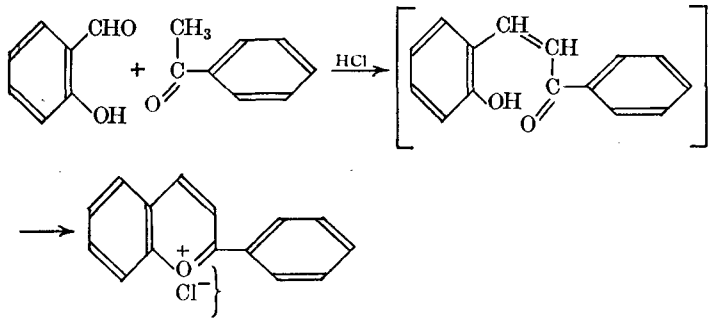
(i) Willstätter (1914) synthesised anthocyanidins starting from coumarin.



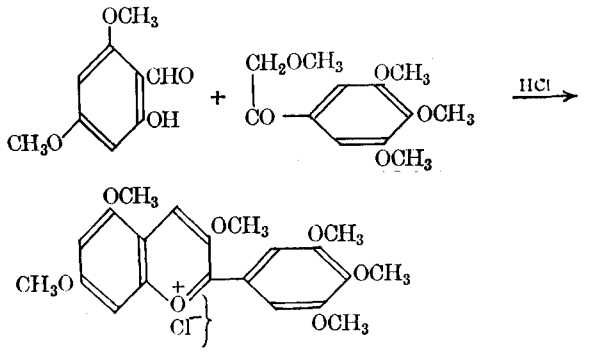
coumarin

This method has very limited application.

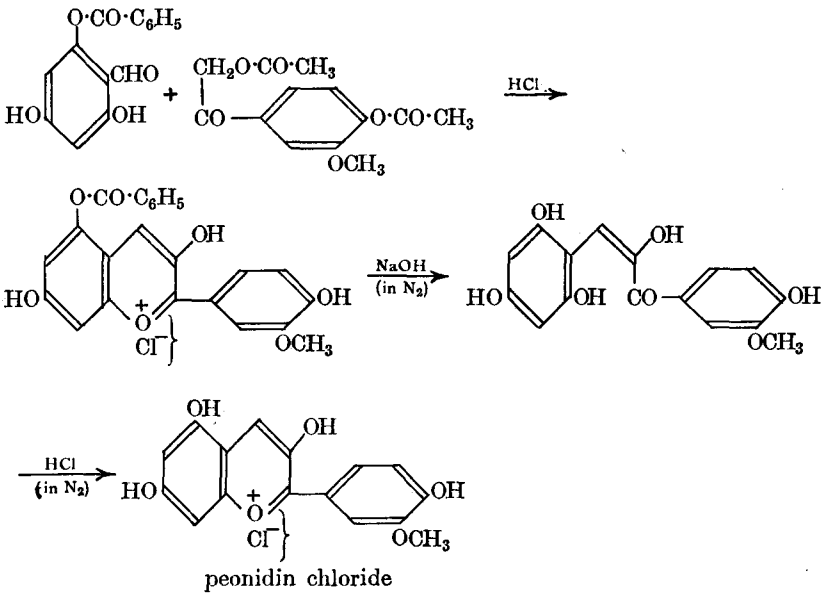
(ii) Robinson has introduced a number of methods whereby *all* anthocyanidins can be prepared. The basic reaction of these methods is the condensation between *o*-hydroxybenzaldehyde and acetophenone in ethyl acetate solution which is then saturated with hydrogen chloride.



The original method of Robinson (1924) resulted in the formation of a product in which the substituent groups were either all hydroxyl groups or all methoxyl groups, *e.g.*,



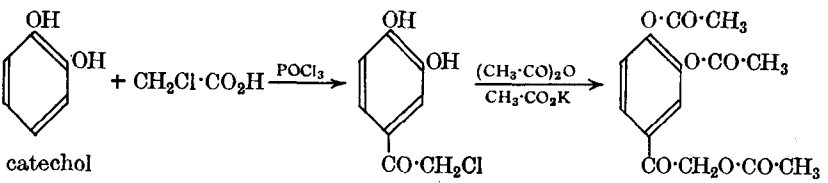
Robinson (1928, 1931) then modified this method so that the product could have both hydroxyl and methoxyl substituent groups, *e.g.*,



peonidin chloride

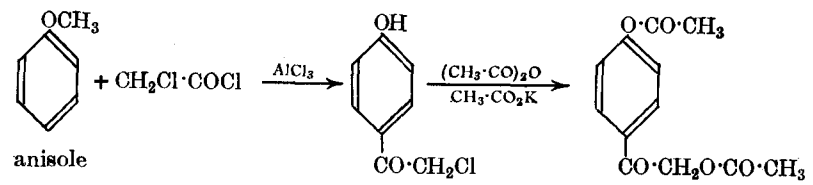
The following is a brief account of the methods used by Robinson and his co-workers for preparing the substituted acetophenones and substituted benzaldehydes.

ω : 3 : 4-Triacetoxyacetophenone.



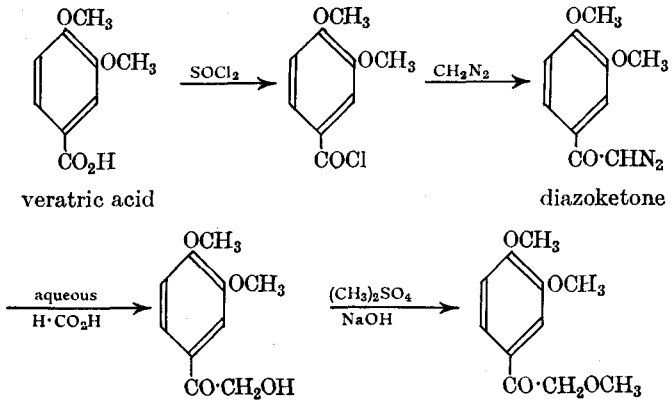
catechol

ω : 4-Diacetoxyacetophenone.



anisole

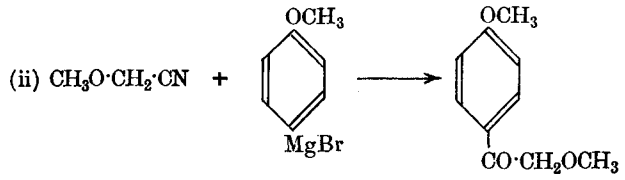
ω : 3 : 4-Trimethoxyacetophenone.



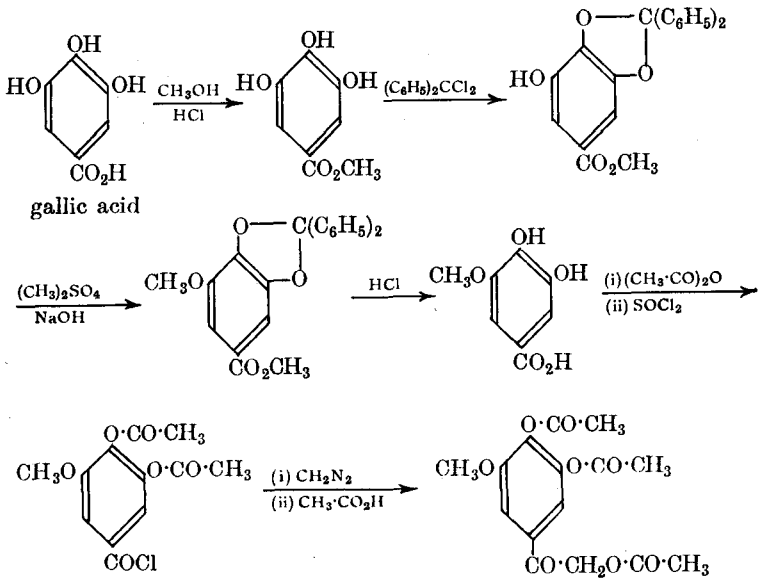
veratric acid diazoketone

ω : 4-Dimethoxyacetophenone.

(i) $CH_2O + (CH_3)_2SO_4 + KCN \longrightarrow CH_3O{\cdot}CH_2{\cdot}CN$ cyanodimethyl ether

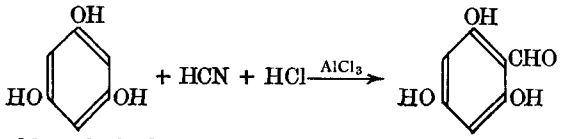


ω : 3 : 4-Triacetoxy-5-methoxyacetophenone.



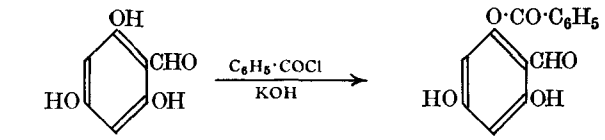
gallic acid

2 : 4 : 6-Trihydroxybenzaldehyde (phloroglucinaldehyde).



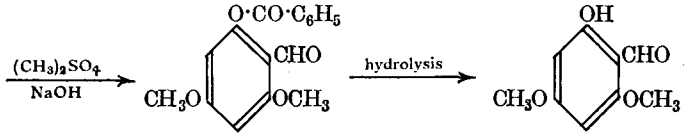
phloroglucinol

2-Hydroxy-4 : 6-dimethoxybenzaldehyde.



phloroglucinaldehyde

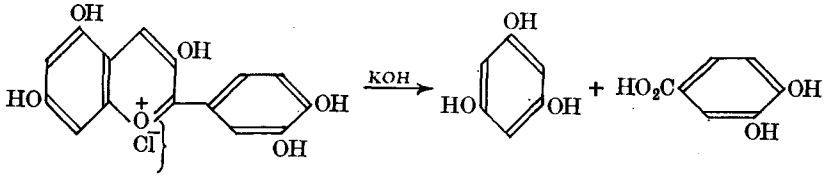
2-benzoylphloroglucin-aldehyde-(2-benzoyloxy-4:6-dihydroxybenzaldehyde)



§5. Cyanidin chloride, $C_{15}H_{11}O_6Cl$. Cyanin chloride, on hydrolysis with hydrochloric acid, gives cyanidin chloride and two molecules of D-glucose.

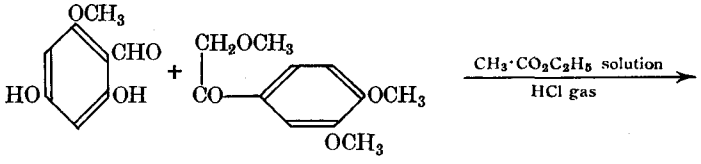
$$C_{27}H_{31}O_{16}Cl + 2H_2O \xrightarrow{HCl} C_{15}H_{11}O_6Cl + 2C_6H_{12}O_6$$

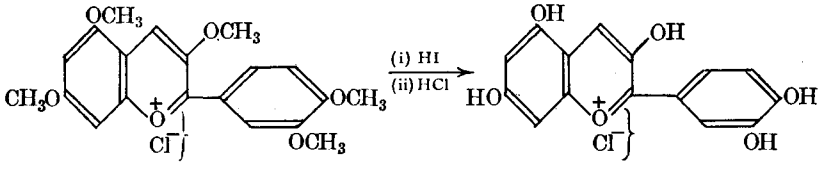
Since cyanidin chloride forms a penta-acetate, the molecule therefore contains five hydroxyl groups. No methoxyl groups are present, and so the potassium hydroxide fusion may be used to degrade this compound; this gives phloroglucinol and protocatechuic acid. Thus cyanidin chloride has the following structure:



cyanidin chloride

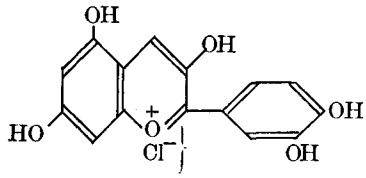
This structure has been confirmed by synthesis (Robinson *et al.*, 1928):



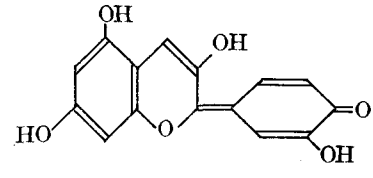


The formation of phloroglucinol and protocatechuic acid by the alkaline fusion of cyanidin chloride suggests a relationship to quercetin, since the latter also gives the same fusion products (see §14).

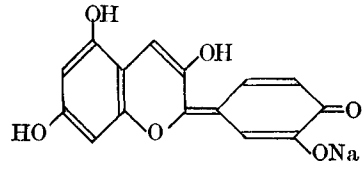
Cyanidin is insoluble in water, but is very soluble in ethanol. It is also soluble in aqueous sodium hydroxide, the solution being blue. The addition of hydrochloric acid changes the colour to purple when the solution is neutral, and when acid the solution becomes red. According to Everest (1914), the colours are due to the following structures (see also Ch. XXXI, Vol. I):



Oxonium salt
Red in acid solution



Colour base
Purple in neutral solution

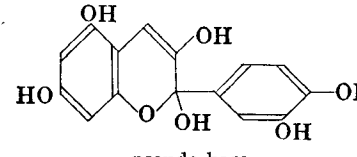


Salt of the colour base
Blue in alkaline solution

Thus a variation of the *p*H will produce a variation in the range of colour.

On the basis of these ionic structures (positive for oxonium salts and negative for salts of the colour bases), anthocyanins should migrate in an electric field. Markakis (1960) has shown that various anthocyanins, when placed within an electric field applied across filter paper, move to the anode or cathode according to the *p*H of the solution. The method of paper electrophoresis may prove to be a very good means of separating, purifying, characterising and preparing anthocyanins.

Markakis also showed that isoelectric point (§4c. XIII) and the *p*H of minimum colour display coincide. On the acidic side of the isoelectric point, the oxonium salt-form predominates; and when the *p*H is higher than that of the isoelectric point, the salt of the colour base predominates. Sondheimer (1953) proposed that a pseudo-base of the structure shown is also possible (this is formed by the addition of a molecule of water to the colour

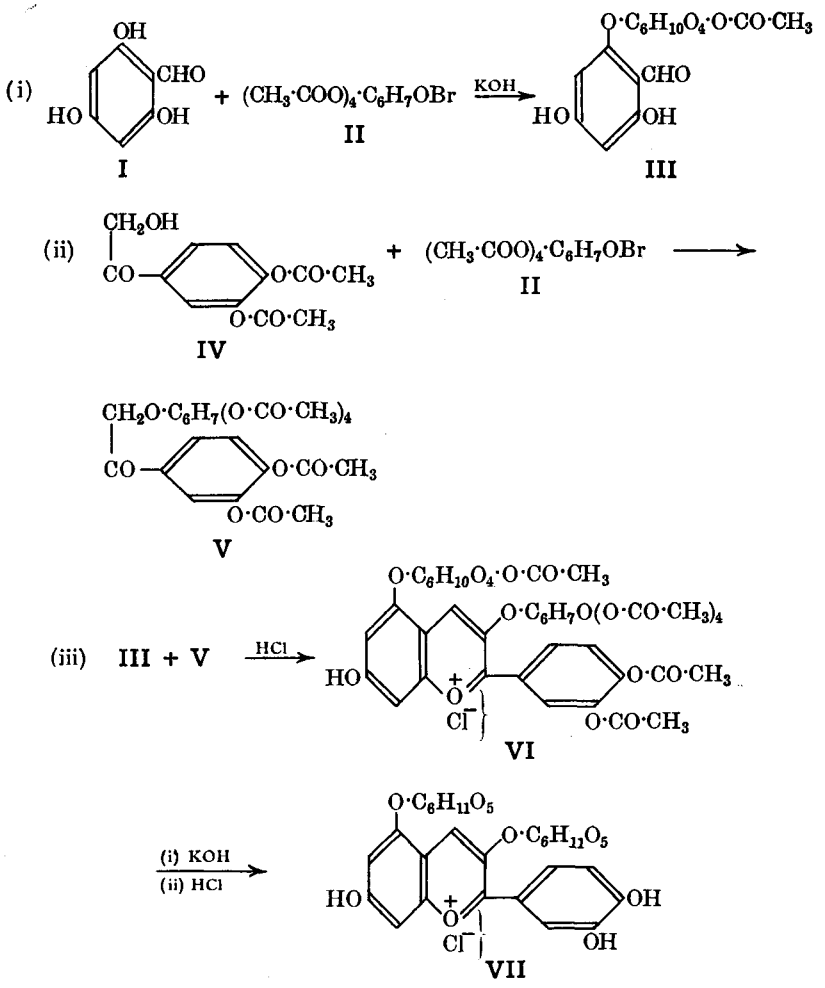


pseudo-base

base), and according to Markakis, it is this form which probably predominates at the isoelectric point. This structure has an interrupted conjugated bond system, and hence will be less coloured than the colour base itself.

Cyanin was the first anthocyanin to be isolated and its structure determined. It has been synthesised by Robinson *et al.* (1932). Phloroglucin-

aldehyde, I, is condensed with tetra-acetyl-α-bromoglucose, II (*cf.* §24. VII), in acetone solution to which has been added aqueous potassium hydroxide; the product is 2-*O*-monoacetyl-β-glucosidylphloroglucinaldehyde, III. ω-Hydroxy-3 : 4-diacetoxyacetophenone, IV, is also condensed with tetra-acetyl-α-bromoglucose (II) in benzene solution to give ω-*O*-tetra-acetyl-β-glucosidoxy-3 : 4-diacetoxyacetophenone, V. Compounds III and V are then dissolved in ethyl acetate and the solution saturated with hydrogen chloride; the product, VI, is treated first with cold aqueous potassium hydroxide and then with hydrochloric acid, whereby cyanin chloride, VII, is produced.

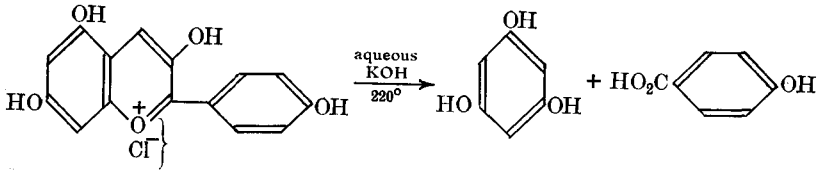


§6. Pelargonidin chloride, $C_{15}H_{11}O_5Cl$. This is formed, together with two molecules of glucose, when pelargonin chloride is hydrolysed with hydrochloric acid.

$$C_{27}H_{31}O_{15}Cl + 2H_2O \xrightarrow{HCl} C_{15}H_{11}O_5Cl + 2C_6H_{12}O_6$$

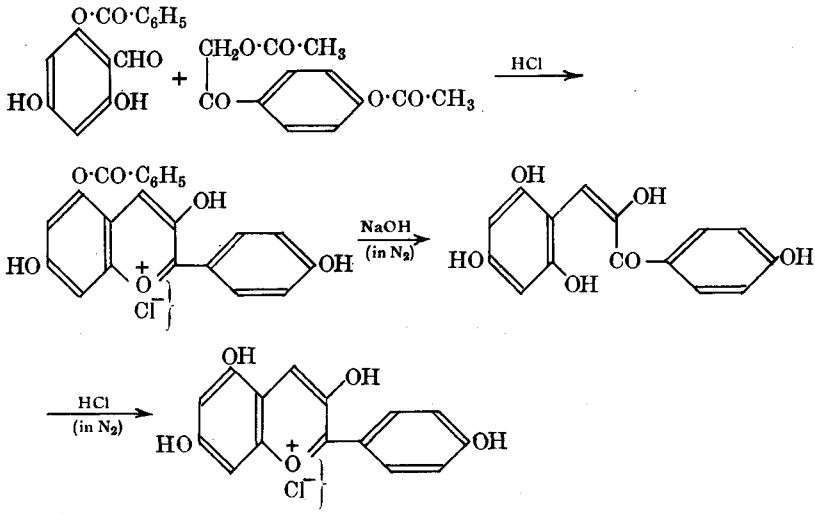
Since pelargonidin chloride forms a tetra-acetate, the molecule contains four hydroxyl groups. Furthermore, since there are no methoxyl groups

present, the potassium hydroxide fusion or boiling with concentrated potassium hydroxide solution may be used to degrade the compound; the products are phloroglucinol and *p*-hydroxybenzoic acid, and so the structure is probably as shown:

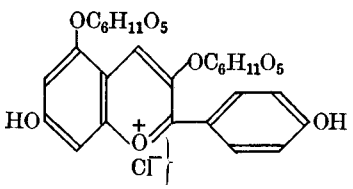


pelargonidin chloride

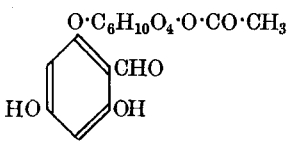
This structure has been confirmed by synthesis, *e.g.*, Robinson *et al.* (1928).



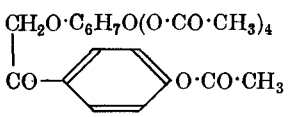
Pelargonin chloride, I, has been synthesised by Robinson *et al.* (1932) from 2-*O*-monoacetyl-*β*-glucosidylphloroglucinaldehyde, II, and *ω*-*O*-tetraacetyl-*β*-glucosidoxy-4-acetoxyacetophenone, III (*cf.* cyanin chloride, §5).



I

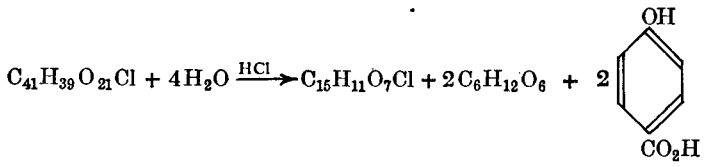


II

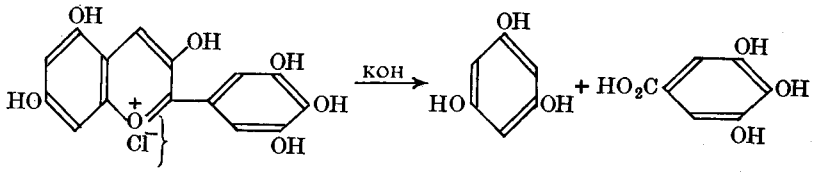


III

§7. **Delphinidin chloride,** $C_{15}H_{11}O_7Cl$, is obtained, together with two molecules of glucose and two molecules of *p*-hydroxybenzoic acid, when delphinin chloride is hydrolysed with hydrochloric acid.



Delphinidin chloride contains six hydroxyl groups, and no methoxyl groups; on fusion with potassium hydroxide, the products are phloroglucinol and gallic acid.



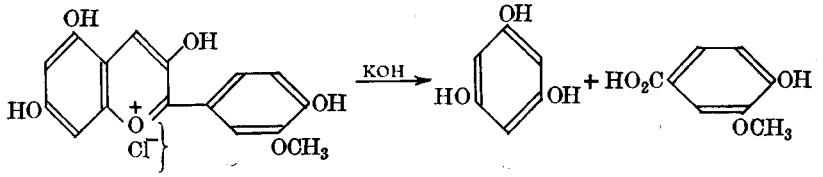
delphinidin chloride

This structure has been confirmed by synthesis, starting from 2-benzoylphloroglucinaldehyde and $\omega : 3 : 4 : 5$-tetra-acetoxyacetophenone (Robinson *et al.*, 1930).

§8. **Peonidin chloride,** $C_{16}H_{13}O_6Cl$, is produced, together with two molecules of glucose, when peonin chloride is hydrolysed with hydrochloric acid.

$$C_{28}H_{33}O_{16}Cl + 2H_2O \xrightarrow{HCl} C_{16}H_{13}O_6Cl + 2C_6H_{12}O_6$$

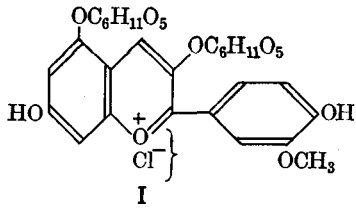
When heated with hydrogen iodide in the presence of phenol, peonidin chloride is demethylated to give cyanidin chloride. Thus peonidin is the monomethyl ether of cyanidin. Heating peonidin chloride with potassium hydroxide solution produces 4-hydroxy-3-methoxybenzoic acid and phloroglucinol. Thus:



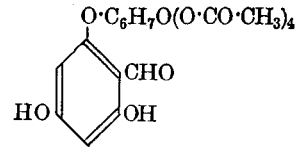
peonidin chloride

This structure has been confirmed by synthesis from 2-benzoylphloroglucinaldehyde and ω : 4-diacetoxy-3-methoxyacetophenone (Robinson *et al.*, 1926).

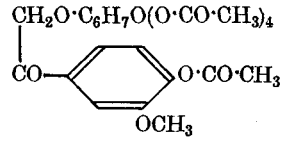
Peonin chloride, I, has been synthesised by Robinson *et al.* (1931), using 2-*O*-tetra-acetyl-β-glucosidylphloroglucinaldehyde, II, and ω-tetra-acetyl-β-glucosidoxy-4-acetoxy-3-methoxyacetophenone, III.



I



II

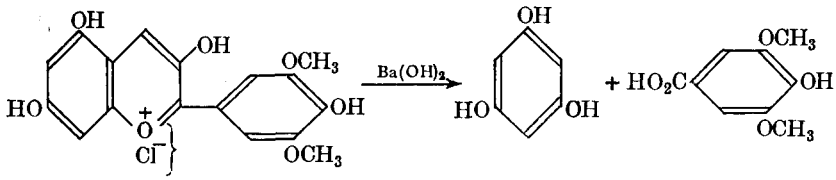


III

§9. Malvidin chloride, $C_{17}H_{15}O_7Cl$, is produced, together with two molecules of glucose, when malvin chloride is hydrolysed with hydrochloric acid.

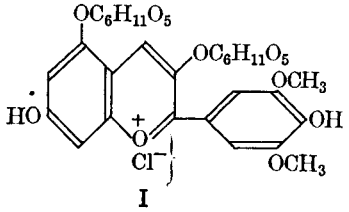
$$C_{29}H_{35}O_{17}Cl + 2H_2O \xrightarrow{HCl} C_{17}H_{15}O_7Cl + 2C_6H_{12}O_6$$

Malvidin chloride contains four hydroxyl groups and two methoxyl groups. When degraded by boiling barium hydroxide solution in an atmosphere of hydrogen, the products are phloroglucinol and *syringic acid* (4-hydroxy-3 : 5-dimethoxybenzoic acid). Thus:

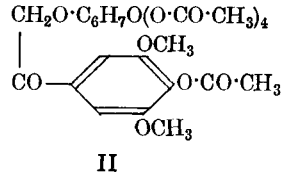


malvidin chloride

Robinson *et al.* (1928) confirmed this structure by synthesis, starting from 2-benzoylphloroglucinaldehyde and *ω*-acetoxy-4-benzyloxy-3 : 5-dimethoxyacetophenone (*cf.* §10). Robinson *et al.* (1932) have also synthesised **malvin chloride,** I, by condensing 2-*O*-tetra-acetyl-*β*-glucosidylphloroglucinaldehyde with *ω*-*O*-tetra-acetyl-*β*-glucosidoxy-4-acetoxy-3 : 5-dimethoxyacetophenone, II.



I



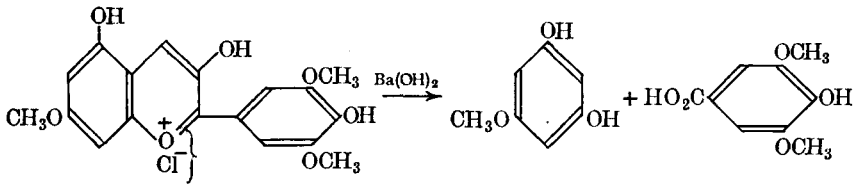
II

§10. Hirsutidin chloride, $C_{18}H_{17}O_7Cl$, is produced by the hydrolysis of hirsutin chloride with hydrochloric acid; two molecules of glucose are also produced.

$$C_{30}H_{37}O_{17}Cl + 2H_2O \xrightarrow{HCl} C_{18}H_{17}O_7Cl + 2C_6H_{12}O_6$$

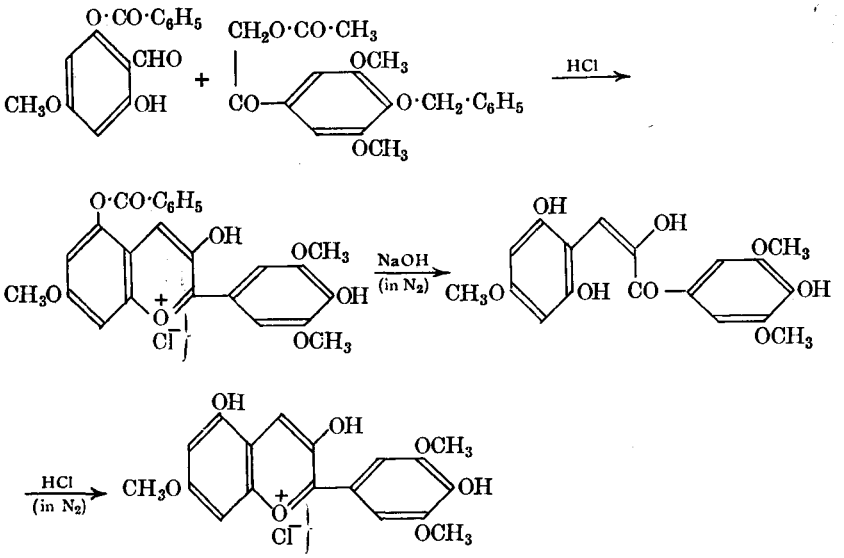
Hirsutidin chloride contains three hydroxyl groups and three methoxyl groups. Its structure is shown from the fact that on hydrolysis with barium

hydroxide solution in an atmosphere of hydrogen, the products are monomethylphloroglucinol and syringic acid. The formation of these products

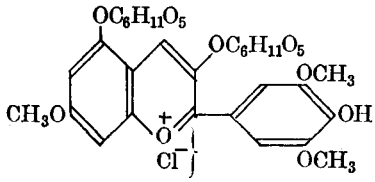


hirsutidin chloride

does not prove conclusively that the methoxyl group at position 7 is actually there; had this position been interchanged with the hydroxyl group at position 5, monomethylphloroglucinol would still have been obtained (*cf.* §3). The formula given for hirsutidin chloride, however, has been confirmed by synthesis, starting from 2-benzoyl-4-*O*-methylphloroglucinaldehyde and ω-acetoxy-4-benzyloxy-3 : 5-dimethoxyacetophenone (Robinson *et al.*, 1930).



Hirsutin chloride has also been synthesised by Robinson *et al.* (1932) from 2-*O*-tetra-acetyl-β-glucosidyl-4-*O*-methylphloroglucinaldehyde and ω-*O*-tetra-acetyl-β-glucosidoxy-4-acetoxy-3 : 5-dimethoxyacetophenone.

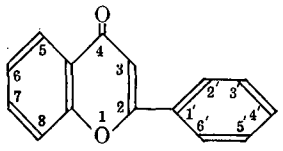


hirsutin chloride

FLAVONES

§11. Introduction. The flavones, which are also known as the **anthoxanthins,** are yellow pigments which occur in the plant kingdom. Flavones

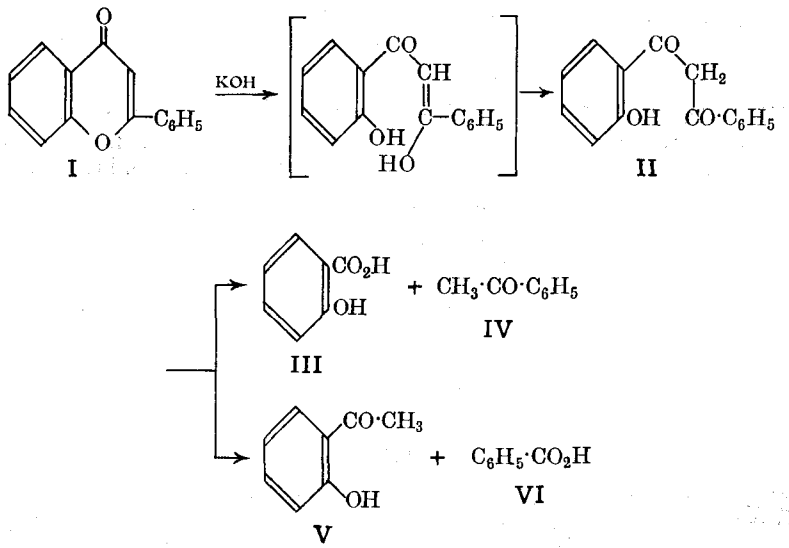
occur naturally in the free state, or as glycosides (the aglycon is the *anthoxanthidin* and the sugar is glucose or rhamnose), or associated with tannins. Chemically, the flavones are very closely related to the anthocyanins; the flavones are hydroxylated derivatives of *flavone* (2-phenyl-4-chromone) which may be partially alkylated. In almost all cases positions 5 and 7 are



flavone

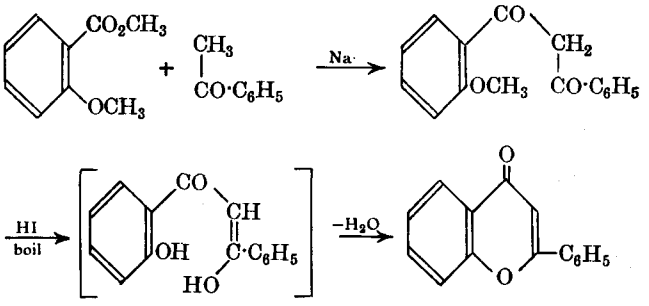
hydroxylated, and frequently one or more of positions 3′, 4′ and 5′. The general method of ascertaining the structure of the flavones is similar to that used for the anthocyanins: the number of free phenolic groups and the number of methoxyl groups are first determined, and then the products obtained by alkaline fusion or hydrolysis are examined. Finally, the structure is confirmed by synthesis. Recently, Simpson *et al.* (1954) have shown that methoxyflavones may be demethylated selectively by hydrobromic acid, the relative rates being $3' > 4' > 7$. These authors have also shown that the relative rates of methylation of flavone-hydroxyl groups with methyl sulphate and sodium hydrogen carbonate in acetone solution are $7 > 4' > 3' > 3$. With methyl sulphate and aqueous alcoholic sodium carbonate, the exact reverse of this order is obtained. These results thus offer a method of ascertaining the positions of methoxyl groups in various methoxyflavones.

§12. Flavone, $C_{15}H_{10}O_2$, occurs naturally as " dust " on flowers, leaves, etc. When boiled with concentrated potassium hydroxide solution, flavone, I, gives a mixture of four products, salicylic acid (III), acetophenone (IV), *o*-hydroxyacetophenone (V) and benzoic acid (VI). The products, which are produced in the pairs III and IV, and V and VI, arise from the fact that the opening of the pyrone ring produces *o*-hydroxydibenzoylmethane, II, which then undergoes scission in two different ways (II is a β-diketone).

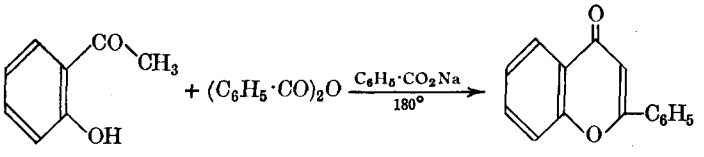


In general, all the flavones give a mixture of four products when degraded with potassium hydroxide. The intermediate *o*-hydroxy-*β*-diketone can be isolated if *cold* alkali or an ethanolic solution of sodium ethoxide is used. On the other hand, if a normal solution of barium hydroxide is used as the degrading agent, then the products are usually salicylic acid and acetophenone (Simonis, 1917).

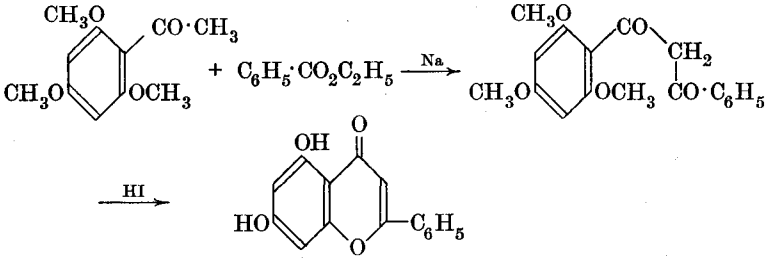
The structure given for flavone has been confirmed by synthesis. Many syntheses are known, *e.g.*, *the Kostanecki synthesis* (1900). This is a general method for synthesising flavones, and consists in condensing the ester of an alkylated salicylic acid with an acetophenone in the presence of sodium (this is an example of the Claisen condensation; this synthesis is a reversal of the formation of III and IV). Thus, for flavone itself, the reaction is carried out with methyl *o*-methoxybenzoate and acetophenone.



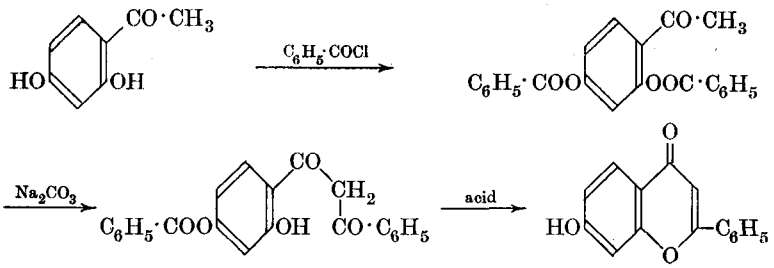
The most useful general synthetic method for preparing flavones is that of Robinson (1924). This is a reversal of the formation of V and VI; an *o*-hydroxyacetophenone is heated at about 180° with the anhydride and sodium salt of a substituted benzoic acid, *e.g.*, flavone:



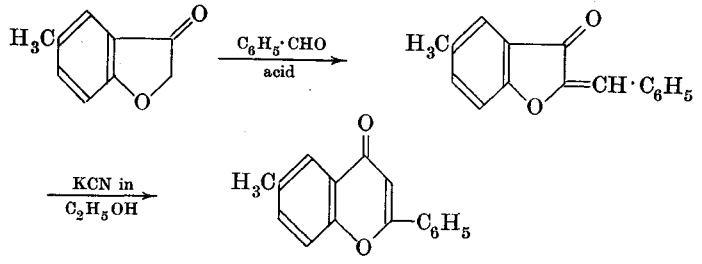
Another general method which is also a reversal of the formation of V and VI is illustrated by the preparation of *chrysin* (5 : 7-dihydroxyflavone) from 2 : 4 : 6-trimethoxyacetophenone and ethyl benzoate.



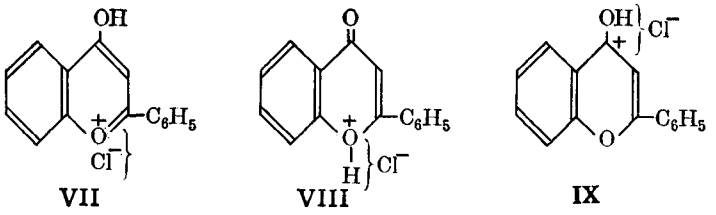
This preparation involves a Claisen condensation, and the following is also another general method which involves an " internal " Claisen condensation.



A recent method for synthesising flavones is by the ring expansion of 2-benzylidenecoumaran-3-ones (Wheeler *et al.*, 1955), *e.g.*,

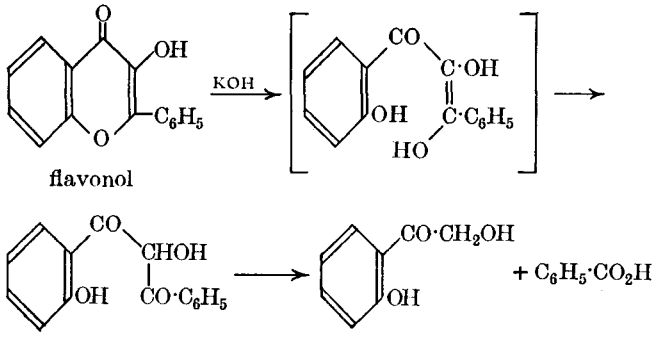


Most flavones are yellow solids which are soluble in water, ethanol and dilute acids and alkalis. The oxonium salts are usually more highly coloured than the free bases; the flavones do not occur naturally as salts (*cf.* anthocyanins). The structure of flavone salts is not certain; VII, VIII and IX are possibilities, and according to calculations of charge distribution (in γ-pyrone salts), IX appears to be most likely (Brown, 1951).

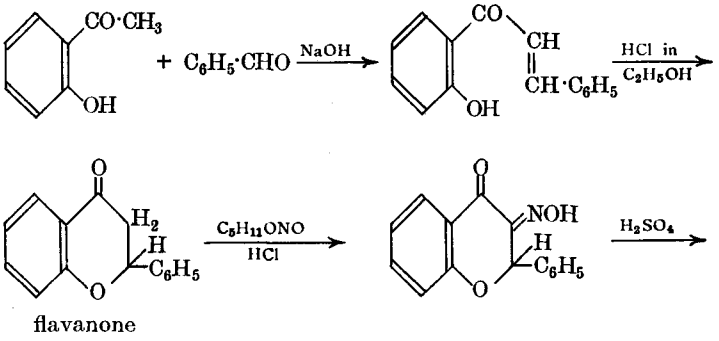


§13. Flavonol (3-hydroxyflavone), $C_{15}H_{10}O_3$. Flavonol is widely distributed in the plant kingdom, usually in the form of glycosides.

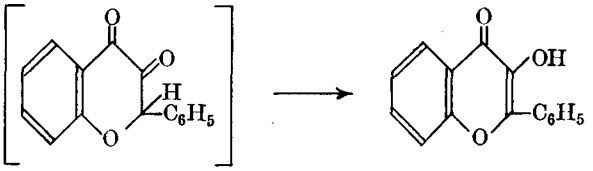
When boiled with an ethanolic solution of potassium hydroxide, flavonol gives *o*-hydroxybenzoylmethanol and benzoic acid. This suggests that flavonol is 3-hydroxyflavone (3-hydroxy-2-phenyl-γ-chromone).



This structure has been confirmed by various syntheses, *e.g.*, Kostanecki *et al.* (1904). This is a general method, and uses the Claisen reaction between *o*-hydroxyacetophenones and substituted benzaldehydes, *e.g.*, flavonol.

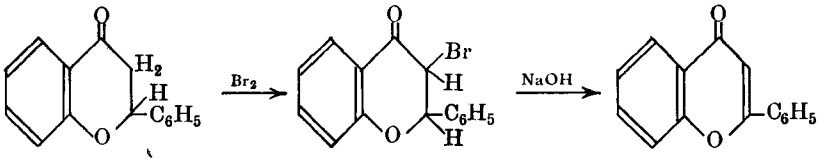


flavanone



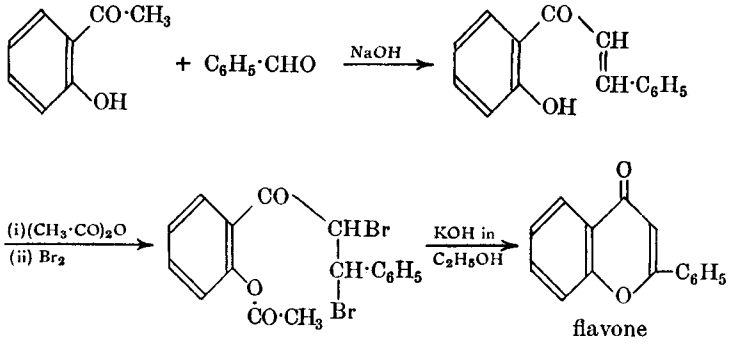
keto form enol form; flavonol

The synthesis, starting from flavanone, has been adapted to the preparation of flavones.



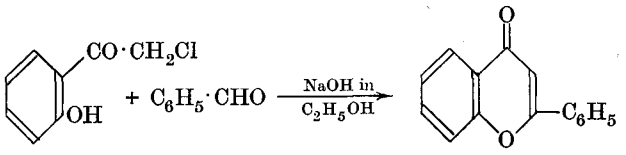
flavanone flavone

An alternative general method for preparing flavones based on the flavonol synthesis is as follows (Kostanecki *et al.*, 1898):



flavone

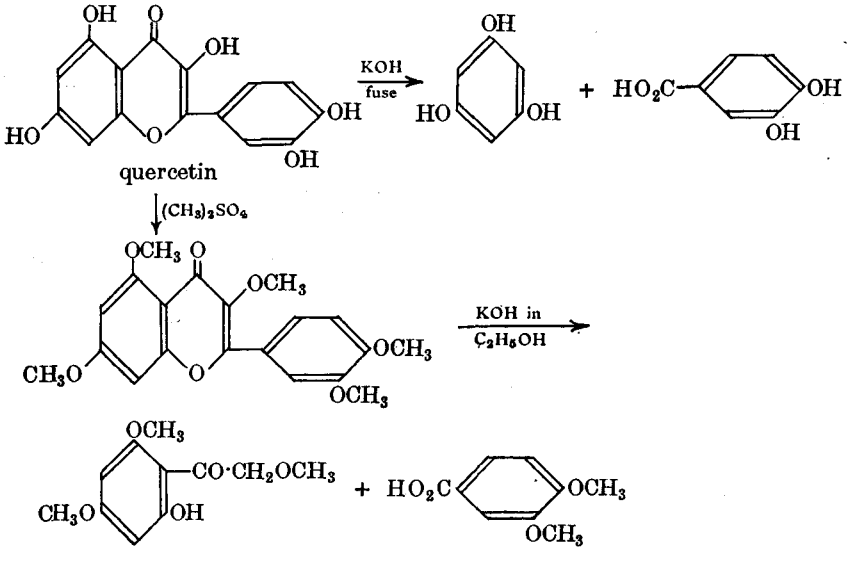
This synthesis has been simplified by Wheeler *et al.* (1955); these authors prepared flavones by condensing ω-chloro-*o*-hydroxyacetophenones with aromatic aldehydes in the presence of ethanolic sodium hydroxide, *e.g.*,



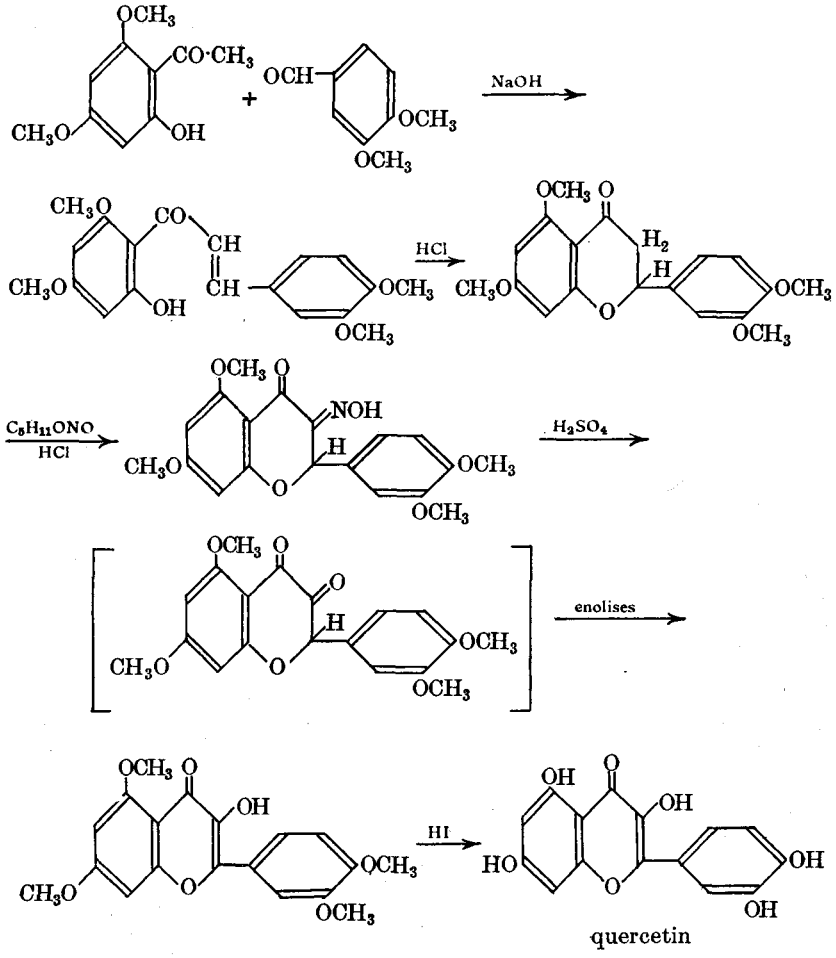
§14. Quercetin, $C_{15}H_{10}O_7$, occurs as the glycoside *quercitrin* in the bark of *Quercus tinctoria*; quercitrin appears to be the most widely distributed natural pigment. On hydrolysis with acids, quercitrin forms quercetin and one molecule of rhamnose.

$$C_{21}H_{20}O_{11} + H_2O \xrightarrow{HCl} C_{15}H_{10}O_7 + CH_3\cdot(CHOH)_4\cdot CHO$$

Quercetin contains five hydroxyl groups; no methoxyl groups are present; on fusion with potassium hydroxide, phloroglucinol and protocatechuic acid are obtained (*cf.* cyanidin, §5). Also, when quercetin is methylated and the product, pentamethylquercetin, boiled with an ethanolic solution of potassium hydroxide, 6-hydroxy-ω : 2 : 4-trimethoxyacetophenone and veratric acid are obtained. These results suggest that quercetin is 3 : 3′ : 4′ : 5 : 7-pentahydroxyflavone.

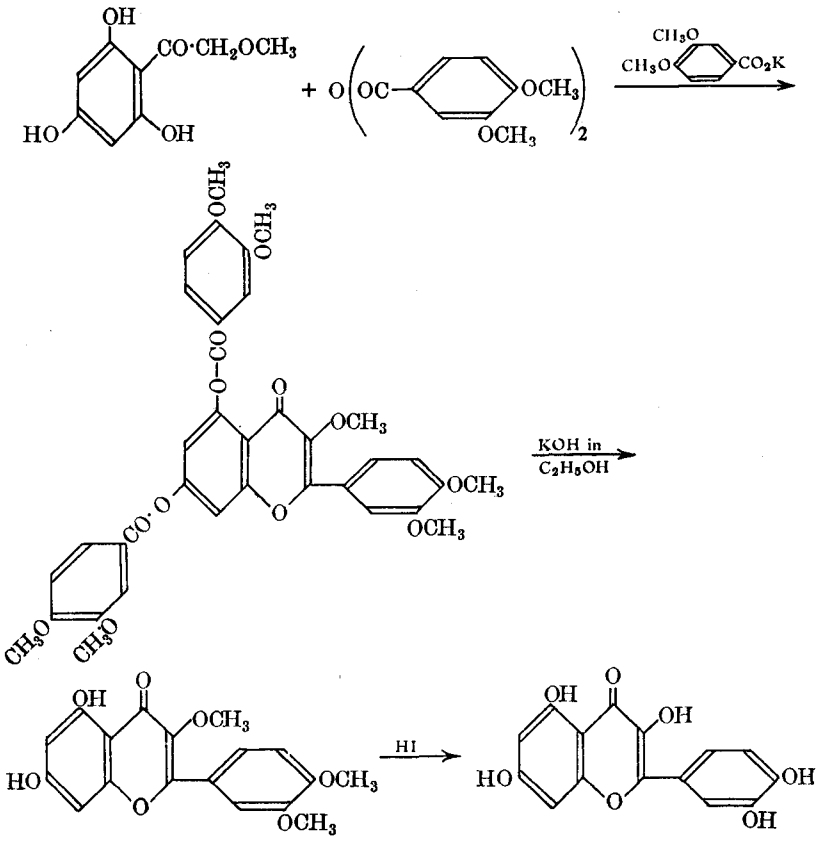


This structure has been confirmed by synthesis, *e.g.*, Kostanecki *et al.* (1904).



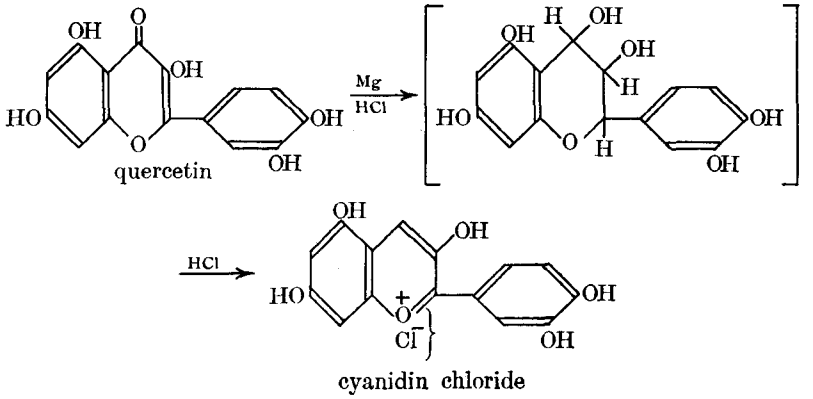
quercetin

Another synthesis is that of Robinson *et al.* (1926); it is a general method for flavonols (*cf.* flavone, §12): ω-methoxyphloroacetophenone is condensed with veratric anhydride in the presence of the potassium salt of veratric acid.



The position of the rhamnose residue in quercitrin has been shown to be **3** (Herzig *et al.*, 1912).

Before leaving this problem of quercetin, let us consider its relationship to cyanidin (§5). As we have seen, the relationship between the two compounds is suggested by the fact that both give the same products when fused with potassium hydroxide. Willstätter *et al.* (1914) reduced quercetin with magnesium in hydrochloric acid containing mercury, and thereby obtained a small amount of cyanidin chloride.



quercetin

cyanidin chloride

Bauer *et al.* (1954) have converted the penta-acetate of quercetin into cyanidin chloride by means of lithium aluminium hydride.

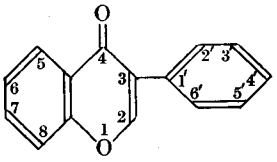
King *et al.* (1957) have shown that the reductive acetylation of a flavonol, followed by the action of hot hydrochloric acid, gives the corresponding anthocyanidin; thus:

$$\text{quercetin} \xrightarrow[\text{(ii) HCl}]{\text{(i) Zn–AcONa; } Ac_2O} \text{cyanidin chloride}$$

This appears to be a useful general method.

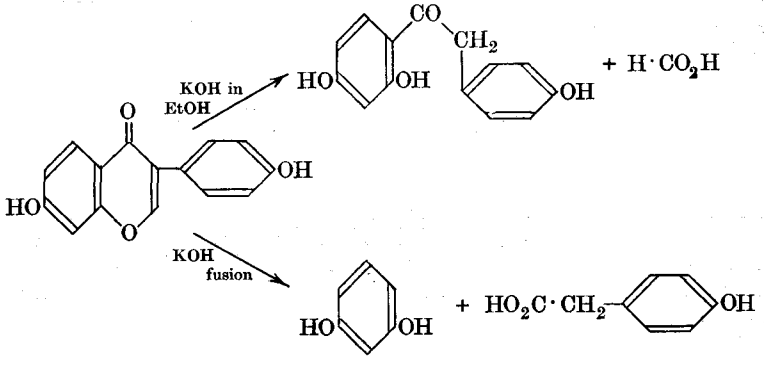
ISO*FLAVONES*

§14a. *iso*Flavones are hydroxylated derivatives of *iso*flavone (3-phenyl-4-chromone) which may be partially alkylated. The *iso*flavones occur



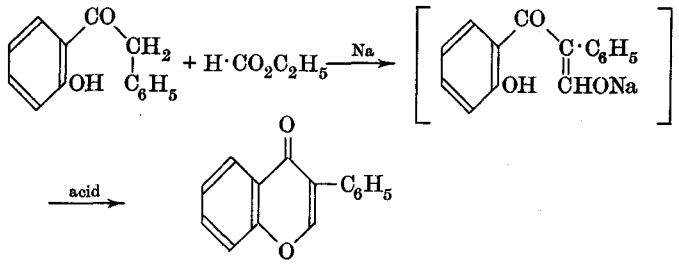
*iso*flavone

naturally, but are not so widespread as the flavones; they occur either in the free state or as glycosides. The general method of ascertaining the structure of *iso*flavones is similar to that used for the flavones (see §§3, 11). Thus fusion with potassium hydroxide breaks down the molecule into two fragments, and hydrolysis with ethanolic potassium hydroxide permits the isolation of intermediates. This may be illustrated with *daidzein* (Walz, 1931):

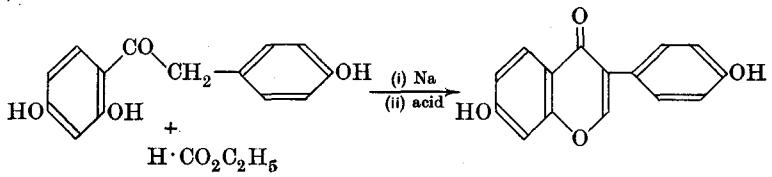


Oxidation with alkaline hydrogen peroxide may also be used in degrading *iso*flavones; recognisable fragments are not usually obtained by this method, but sometimes information may be obtained about the substituents in the 3-phenyl nucleus, *e.g.*, *genistein* (4′ : 5 : 7-trihydroxy*iso*flavone) gives *p*-hydroxybenzoic acid.

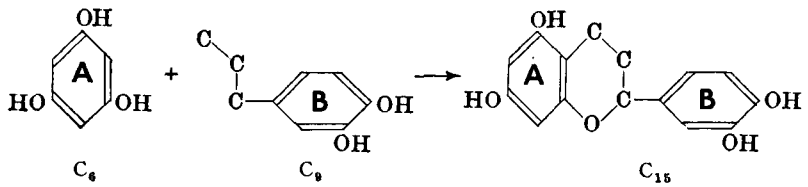
The final proof of the structure of an *iso*flavone lies in its synthesis. A general method of synthesising *iso*flavones is that of Späth *et al.* (1930); *e.g.*, *iso*flavone itself may be synthesised from benzyl *o*-hydroxyphenyl ketone and ethyl formate:



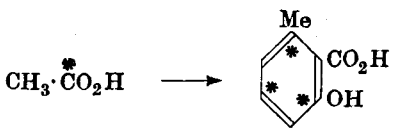
By using substituted ketones, various *iso*flavones may be synthesised, *e.g.*, daidzein from 2 : 4-dihydroxyphenyl *p*-hydroxybenzyl ketone (Wessely *et al.*, 1933):



§14b. Biosynthesis of the flavonoids. Robinson (1936) considered the C_{15} skeleton of flavonoids to be composed of two parts, C_6 and C_3:



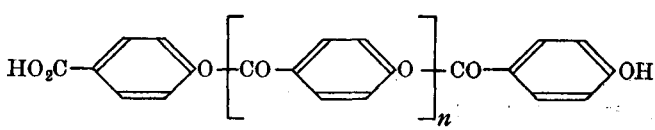
Biosynthetic work has shown that rings A and B are derived from different sources. Birch *et al.* (1955) have carried out the biosynthesis of benzenoid compounds from acetate, *e.g.*, using cultures of *Penicillium griseofulvin*, it was shown that:



Underhill *et al.* (1957), using ^{14}C-labelled compounds, showed that in the biosynthesis of quercetin and cyanidin, rings A and B have different origins; ring A appears to be produced from acetate, but ring B is produced by the shikimic acid pathway (§18. XIII). Biosynthetic studies of cyanidin (Weygand *et al.*, 1957; Grisebach, 1958) also support the origin of phloroglucinol (ring A) from acetate units.

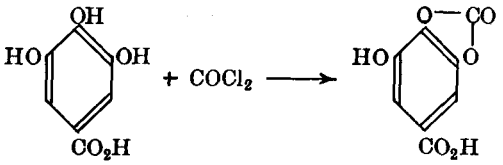
DEPSIDES

§15. Depsides. Phenolic acids, by the interaction of the carboxyl group of one molecule with the hydroxyl group of another, give rise to **depsides:**

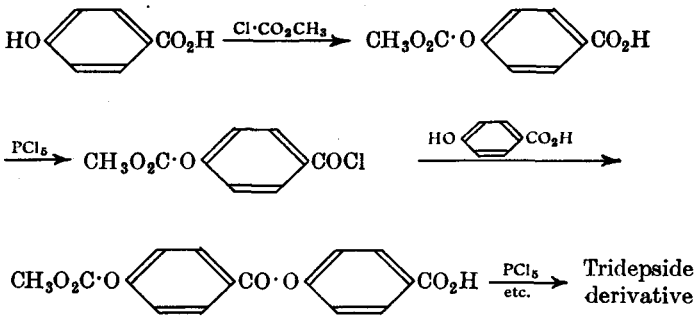


If n is zero, then the molecule is a didepside; if n is 1, then a tridepside; etc. The main sources of the depsides are the lichens.

In order to synthesise depsides in a known fashion, it is necessary to protect hydroxyl groups. Fischer (1919) carried this out by means of acetylation (acetic anhydride) or by introducing a carbomethoxyl group (with methyl chloroformate); two hydroxyl groups in the *ortho*-position may be protected by means of carbonyl chloride, *e.g.*, gallic acid forms the following compound.

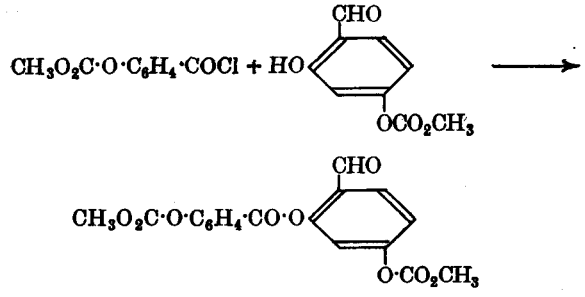


Let us consider the synthesis of a depside from a monohydroxybenzoic acid.



I

I may be hydrolysed to the didepside by means of cold alkali. By using different phenolic acids, it is possible to synthesise a large variety of depsides. When the hydroxyl group is *meta* or *para* to the carboxyl group, the phenolic acid is readily carboxymethylated, but *ortho*-hydroxyl groups are very resistant under the same conditions (steric effect; see Vol. I). Reaction can, however, be brought about by condensing *o*-hydroxyacids with methyl chloroformate in the presence of a base, *e.g.*, dimethylaniline. There is also the further difficulty that *ortho*-hydroxyl groups do not react with acid chlorides (steric effect). This has been overcome by condensing an acid chloride with an *o*-phenolic aldehyde, *e.g.*,



§16. Tannins. These are widely distributed in plants; many are glycosides. One of the best sources of tannin is nutgall. The tannins are colourless non-crystalline substances which form colloidal solutions in water; these

solutions have an astringent taste. Tannins precipitate proteins from solution, and they form a bluish-black colour with ferric salts, a property which is used in the manufacture of ink. Tannins also precipitate many alkaloids from their solutions.

All tannins contain polyhydroxyphenols or their derivatives. Some tannins are hydrolysable by acids, and others are not; those which can be hydrolysed by acid give variable yields of gallic acid.

READING REFERENCES

Gilman (Ed.), *Advanced Organic Chemistry*, Wiley (1943, 2nd ed.). Vol. II. Ch. 18. The Anthocyanins and the Flavones.

Stewart and Graham, *Recent Advances in Organic Chemistry*, Longmans, Green. Vol. II (1948, 7th ed.). (i) Ch. 10. Anthocyanins. (ii) Ch. 11. Depsides and Tannins.

Perkin and Everest, *The Natural Organic Colouring Matters*, Longmans, Green (1918).

Elderfield (Ed.), *Heterocyclic Compounds*, Wiley, Vol. II (1951). (i) Ch. 8. Chromones, Flavones and Isoflavones. (ii) Ch. 9. Chromenols, Chromenes and Benzopyrylium Salts.

Rodd (Ed.), *Chemistry of the Carbon Compounds*, Elsevier. Vol. IVB (1959); pp. 855–; 903–; 935–. Flavonoids, etc.

Bentley, *The Natural Pigments*, Interscience (1960).

Hill, The Synthesis and Structure of Benzopyrylium (Chromylium) Salts, *Chem. Reviews*, 1936, **19**, 27.

Robinson, Natural Colouring Matters and their Analogues, *Chem. and Ind.*, **1933**, 737.

Robinson, Über die Synthese von Anthocyaninen, *Ber.*, 1934, **67A**, 85.

Robinson, Chemistry of the Anthocyanins, *Nature*, 1935, **135**, 732.

Warburton, The *iso*Flavones, *Quart. Reviews* (*Chem. Soc.*), 1954, **8**, 67.

Jain and Seshadri, Nuclear Methylation of Flavones and Related Compounds, *Quart. Reviews* (*Chem. Soc.*), 1956, **10**, 169.

Seshadri, Recent Developments in the Chemistry of Flavonoids, *Tetrahedron*, 1959, **6**, 169.