CHAPTER VI

STEREOCHEMISTRY OF SOME ELEMENTS OTHER THAN CARBON

§1. Shapes of molecules. Many elements other than carbon form compounds which exhibit optical isomerism. Since the criterion for optical activity must be satisfied, *viz.* the molecule must not be superimposable on its mirror image, it therefore follows that the configurations of the various molecules can never be planar.

In Vol. I, Ch. II, the theory of shapes of molecules has been explained on the basis that all electrons (shared and unshared) in the valency shell of the central atom arrange themselves in pairs of opposite spin which keep as far apart as possible. Furthermore, it was assumed that deviations from regular shapes arise from electrostatic repulsions between electron pairs in the valency shell as follows:

lone-pair—lone-pair > lone-pair—bond-pair > bond-pair—bond-pair.

It was also assumed that a double (and triple) bond repels other bond-pairs more than does a single bond. The following two tables illustrate these ideas.

Shapes of molecules containing single bonds

Number of electrons in valency shell	Number of bonding pairs	Number of lone-pairs	Hybrid orbitals used	Shape of molecule	Examples
2	2	0	sp	Linear	$HgCl_2$
3	3	0	sp^2	Triangular plane	BCl_3
4	4	0	sp^3	Tetrahedron	CH_4
	3	1	sp^3	Trigonal pyramid	NH_3
	2	2	sp^3	V-shape	H_2O
5	5	0	sp^3d	Trigonal bipyramid	PCl_5
6	6	0	sp^3d^2	Octahedron	SF_6

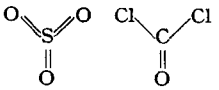
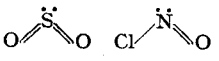
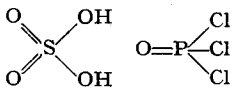
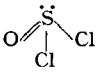
When dealing with molecules containing multiple bonds (treated in terms of σ- and π-bonds), the shapes may also be predicted in a similar fashion if it is assumed that the electron-pairs (2 in a double and 3 in a triple bond) occupy only *one* of the positions in the various arrangements described in the above table, *i.e.*, a multiple bond is treated as a " single " bond. This means that the shape of the molecule is determined by the number of σ-bonds and lone-pairs only; the π-bonds are " fitted in " afterwards (p. 144).

§2. STEREOCHEMISTRY OF NITROGEN COMPOUNDS

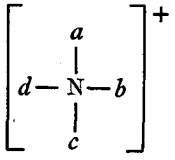
According to the electronic theory of valency, nitrogen can be tercovalent or quadricovalent unielectrovalent; in both of these states nitrogen, as the " central " atom, can exhibit optical activity.

§2a. Quaternary ammonium salts. Originally, the valency of nitrogen in quaternary ammonium salts was believed to be quinquevalent; later,

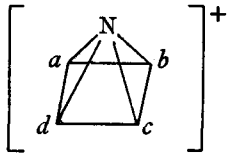
Shapes of molecules containing multiple bonds

Total number of σ-bonds and lone-pairs	Number of σ-bonds	Number of lone-pairs	Shape of molecule	Examples
2	2	0	Linear	O=C=O; H—C≡N
3	3	0	Triangular plane	
	2	1	Triangular plane	
4	4	0	Tetrahedron	
	3	1	Trigonal pyramid	

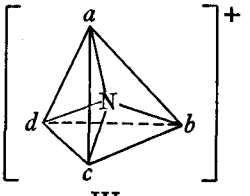
however, it was shown that one valency was different from the other four. Thus, using the formula, $[Nabcd]^+X^-$, for quaternary ammonium salts, and assuming that the charge on the nitrogen atom has no effect on the configuration of the cation, the cation may be considered as a five-point system similar to that of carbon in compounds of the type C*abde*. This similarity is based on the assumption that the four valencies in the ammonium ion are equivalent, and this assumption is well substantiated experimentally and also theoretically. Hence there are three possible configurations for the cation $[Nabcd]^+$, I, II and III (*cf.* §3a. II). If the cation is planar (I),



I



II

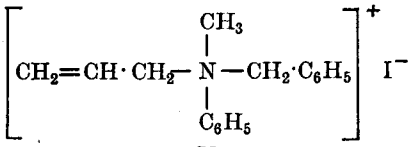


III

then it would not be resolvable; it would be resolvable, however, if the configuration is pyramidal (II) or tetrahedral (III). Le Bel (1891) claimed to have partially resolved *iso*butylethylmethylpropylammonium chloride, IV, by means of *Penicillium glaucum* (*cf.* §10 iii. II), but later work apparently

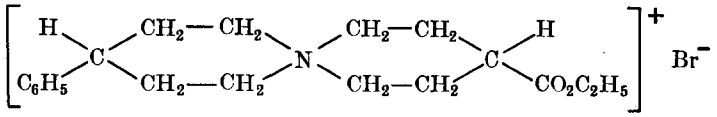
IV

showed this was wrong. The first definite resolution of a quaternary ammonium salt was that of Pope and Peachey (1899), who resolved allylbenzylmethylphenylammonium iodide, V, by means of (+)-bromocamphorsulphonic acid. This was the first case of optical activity due to a " central "

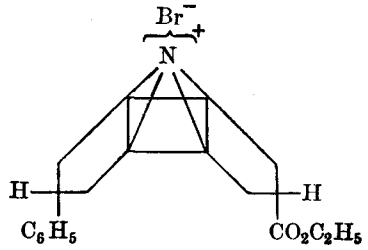


V

atom other than carbon. This resolution was then followed by the work of Jones (1905), who resolved benzylethylmethylphenylammonium iodide. Thus the ammonium ion cannot be planar, but must be either pyramidal or tetrahedral. Bischoff (1890) had proposed a pyramidal structure, and this configuration was supported by Jones (1905) and Jones and Dunlop (1912). On the other hand, Werner (1911) had suggested the tetrahedral configuration, and this was supported by Neagi (1919) and Mills and Warren (1925). It was, however, Mills and Warren who gave the most conclusive evidence that the configuration is tetrahedral. Their evidence is based on the following argument. Compounds of the type *ab*C=C=C*ab* are resolvable since carbon is " tetrahedral " (see allenes, §6. V), and if nitrogen is also " tetrahedral ", then the compound *ab*C=N=C*ab* should be resolvable, but will not be resolvable if the nitrogen is pyramidal. Mills and Warren prepared 4-carbethoxy-4'-phenylbispiperidinium-1 : 1'-spiran bromide, and resolved it. If the configuration of this molecule is VI, *i.e.*, a spiran, then it possesses no elements of symmetry, and hence will be resolvable; if the configuration is VII (*i.e.*, pyramidal), then it will possess a vertical plane of symmetry,

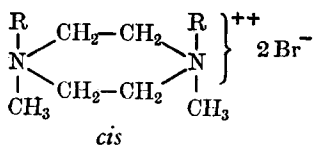


VI

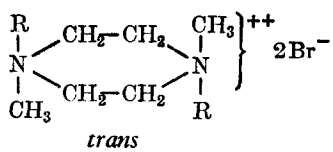


VII

and hence will be optically inactive. Since the compound was resolved, the configuration must be tetrahedral, *i.e.*, VI. This tetrahedral configuration has been confirmed by physico-chemical studies (see §2b). More recently, Hanby and Rydon (1945) have shown that the diquaternary salts of dimethylpiperazine exhibit geometrical isomerism, and this is readily explained on the tetrahedral configuration of the four nitrogen valencies (*cf. cyclo*hexane, §11. IV).

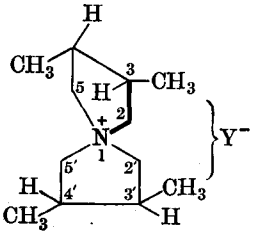


cis

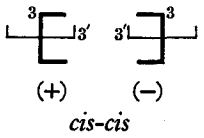


trans

It has already been mentioned (§6. II) that McCasland and Proskow (1956) prepared a spiro-nitrogen compound which contained no plane or centre of symmetry, but was nevertheless optically inactive because it contained an alternating axis of symmetry. We shall now examine this compound (VIII; Y^- is the *p*-toluenesulphonate ion) in more detail. This molecule can exist in four diastereoisomeric forms, three active and one

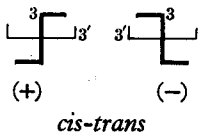


VIII



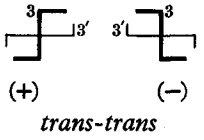
cis-cis

IX



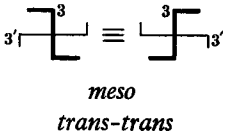
cis-trans

X



trans-trans

XI



meso
trans-trans

XII

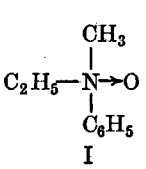
meso. All four have been prepared, and are depicted as shown in IX, X, XI and XII. The co-axis of each spiran is assumed to be perpendicular to the plane of the paper, and the intersecting lines represent the two rings. The short appendages show whether the two substituents (methyl) are *cis* or *trans*. The ring nearer the observer's eye is indicated by the heavy line, and a uniform orientation has been adopted: the front ring is always vertical, and the back horizontal ring with at least one substituent directed upwards and the *cis* ring placed at the back in the case of the *cis/trans* ring combination.

Racemisation of optically active quaternary ammonium salts is far more readily effected than that of carbon compounds containing an asymmetric carbon atom, *i.e.*, compounds of the type C*abde*. The mechanism of the racemisation of the ammonium salts is believed to take place by dissociation into the amine, which then rapidly racemises (§2c):

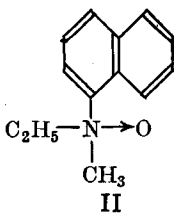
$$N abcd\}\ ^+X^- \rightleftharpoons N abc + dX$$

Recombination of the racemised amine with *d*X results in the racemisation of the quaternary compound (see §4a).

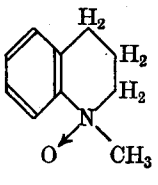
§2b. Tertiary amine oxides. In tertiary amine oxides, *abc*NO, the nitrogen atom is joined to four different groups, and on the basis that the configuration is tetrahedral, such compounds should be resolvable. In



I



II



III

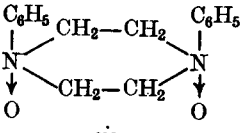
1908, Meisenheimer resolved ethylmethylphenylamine oxide, I, and this was then followed by the resolution of other amine oxides, *e.g.*, ethylmethyl-1-naphthylamine oxide, II, and kairoline oxide, III.

The evidence in favour of the structure IV as opposed to that of V is based on dipole moment measurements and on the fact that such compounds can be resolved. It should be noted that the pyramidal structure would

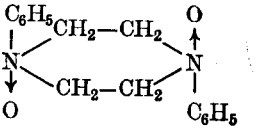
$$R_3N\rightarrow O \quad or \quad R_3\overset{+}{N}-\bar{O} \qquad\qquad R_3N{=}O$$
$$\text{IV} \qquad\qquad \text{V}$$

also account for the optical activity of these compounds as well as the tetrahedral. Consequently these compounds cannot be used as a criterion for the pyramidal or tetrahedral configuration of the nitrogen atom. However, by analogy with the quaternary ammonium salts, the configuration of amine oxides may be accepted as tetrahedral. Further evidence for this is as follows. The electronic configuration of nitrogen is $(1s^2)(2s^2)(2p^3)$. For nitrogen to be quinquevalent, the " valence state " will be derived from the arrangement $(1s^2)(2s)(2p^3)(3s)$. Now the amount of energy required to promote an electron from a $2s$ to a $3s$ orbital appears to be too large for it to occur, and consequently nitrogen is (apparently) never quinquevalent. The valence state of nitrogen is thus achieved by the *loss* of one $2s$ electron and then hybridisation of the $2s$ and $2p^3$ orbitals, *i.e.*, nitrogen becomes quadricovalent unielectrovalent, and the four bonds (sp^3 bonds) are arranged tetrahedrally. The charged nitrogen atom is isoelectronic with carbon, and so one can expect the formation of similar bonds. Furthermore, evidence obtained by an examination of the vibration frequencies of the ammonium ion indicates that the configuration of this ion is tetrahedral.

Recently, Bennett and Glynn (1950) have obtained two geometrical isomers of 1 : 4-diphenylpiperazine dioxide; this is readily explained on the tetrahedral configuration of nitrogen (*cf.* §2a).

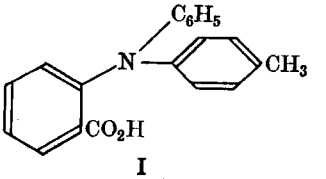


cis



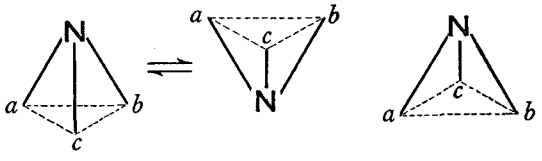
trans

§2c. Amines. If the tertiary amine molecule, N*abc*, is planar, it will be superimposable on its mirror image, and therefore cannot be optically active. All attempts to obtain tertiary amines in optically active forms have failed up to the present time, *e.g.*, Kipping and Salway (1904) treated secondary amines, R·NH·R′, with (±)-benzylmethylacetyl chloride; if the three valencies of the nitrogen atom are not planar, then the base will be a racemic modification, and on reaction with the acid chloride, the following four substituted amides should be formed: B_+A_+, B_-A_-, B_+A_-, B_-A_+, *i.e.*, a mixture of two pairs of enantiomorphs. Experiments carried out with, *e.g.*, methylaniline and benzylaniline gave *homogeneous* products. Meisenheimer *et al.* (1924) attempted to resolve *N*-phenyl-*N*-*p*-tolylanthranilic acid, I, and also failed. In view of these failures, it would thus appear that



I

the tertiary amine molecule is planar. Physico-chemical methods, *e.g.*, dipole moment measurements, infra-red absorption spectra studies, etc., have, however, shown conclusively that the configuration of ammonia and of tertiary amines is tetrahedral. Thus ammonia has been shown to have a dipole moment of 1·5 D; had the molecule been planar, the dipole moment would have been zero. Furthermore, the nitrogen valency angles in, *e.g.*, trimethylamine have been found to be 108°, thus again showing that the amine molecule is not planar. Why, then, cannot tertiary amines be resolved? Is it a question of experimental technique, or is there something inherent in the tertiary amine molecule that makes it impossible to be resolved? Meisenheimer (1924) explained the failure to resolve as follows. In the tertiary amine molecule, the nitrogen atom oscillates rapidly at right angles above and below the plane containing the groups *a*, *b* and *c* (see Fig. 1); II and III are the two extreme forms, and they are mirror



II III IV

FIG. 6.1.

images and not superimposable (IV is III " turned over ", and it can be seen that IV is the mirror image of II). Thus this oscillation brings about very rapid optical inversion. This oscillation theory is supported by evidence obtained from the absorption spectrum of ammonia (Barker, 1929; Badger, 1930), and the frequency of the oscillation (and therefore the inversion) has been calculated to be $2{\cdot}3 \times 10^{10}$ per second (Cleeton *et al.*, 1934).

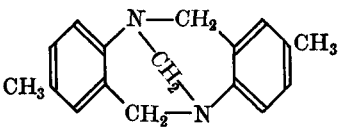
In the foregoing explanation for the racemisation of amines, it has been assumed that the nitrogen valency angles and the bond lengths change. This inversion of amines, however, is better represented as an " umbrella " switch of bonds, *i.e.*, the bond lengths remain unaltered and only the nitrogen valency angles change. This interpretation is more in keeping with the facts, *e.g.*, as the groups *a*, *b* and *c* increase in weight, the frequency of the inversion of the molecule decreases.

Theoretical calculations have shown that an optically active compound will not racemise spontaneously provided that the energy of activation for the change of one enantiomorph into the other is greater than 12–15 kg.cal./mole. The two forms, II and III, have been shown to be separated by an energy barrier of about 6 kg.cal./mole, and consequently the two forms are readily interconvertible.

It has already been mentioned (§2b) that the electronic configuration of the nitrogen atom is $(1s^2)(2s^2)(2p^3)$. According to Hund's rule, electrons tend to avoid being in the same orbital as far as possible (see Vol. I, Ch. II). Thus, in ammonia and its derivatives, bonds are formed by pairing with the three single orbitals $2p_x$, $2p_y$ and $2p_z$. Since these are mutually at right angles, the configuration of the ammonia molecule will be a *trigonal pyramid*, *i.e.*, a pyramid with a triangular base, with the nitrogen atom situated at one corner. Oscillation of the nitrogen atom brings about inversion in the tertiary amines, N*abc*. This picture of the configuration of the ammonia molecule, however, requires modification. The valency angles in ammonia have been shown to be approximately 107°. The deviation from the value of 90° (on the assumption that the bonds are pure $2p$ orbitals) is too great to be accounted for by repulsion between the hydrogen atoms. As we have seen (§1), according to modern theory the orbitals in ammonia

are sp^3, one orbital being occupied by the lone-pair. The deviation of the valency angle of 107° from the tetrahedral value of 109° 28′ has been explained by the greater repulsion between a lone-pair and a bond-pair than between a bond-pair and a bond-pair.

In view of what has been said above, it appears that tertiary amines of the type N*abc* will never be resolved. Now, Kincaid and Henriques (1940), on the basis of calculations of the energy of activation required for the inversion of the amine molecule, arrived at the conclusion that tertiary amines are incapable of resolution because of the ease of racemisation, but if the nitrogen atom formed a part of a *ring* system, then the compound would be sufficiently optically stable to be isolated. This prediction was confirmed by Prelog and Wieland (1944), who resolved Tröger's base, V,



V

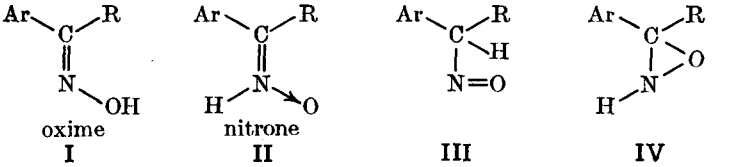
by chromatographic adsorption on D-lactose (*cf.* §10 vi. II). In this compound, the nitrogen is tervalent, but the frequency of oscillation has been brought to zero by having the three valencies of nitrogen as part of the ring system.

Roberts *et al.* (1958) have examined *N*-substituted ethyleneimines (see Vol. I) by NMR spectroscopy. Their results support the "umbrella" switch of bonds, and these authors believe that optical resolution of this type of compound may be possible below −50°.

§2d. Oximes. In 1883, Goldschmidt found that benzil dioxime,

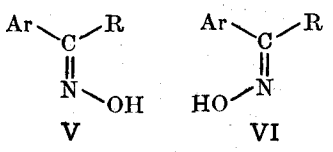
$$C_6H_5{\cdot}C(=NOH){\cdot}C(=NOH){\cdot}C_6H_5,$$

could be converted into an isomeric form by boiling it in ethanolic solution; and then, in 1889, Meyer *et al.* isolated a third isomer of this compound. Beckmann, also in 1889, found that benzaldoxime existed in two isomeric forms, and from that time many aromatic oximes were shown to exist in two isomeric forms. The existence of isomerism in aromatic oximes was first explained by structural isomerism, two of the following four structures corresponding to the two isomers (where R is an alkyl or an aryl group); II is the modern way of writing the nitrone structure (originally, it was



oxime — I; nitrone — II; III; IV

written with quinquevalent nitrogen, the nitrogen being linked to the oxygen by a double bond). Hantzsch and Werner (1890), however, suggested that the isomerism of the oximes was geometrical and not structural. According to these authors, nitrogen is tervalent (in oximes), and is situated at one corner of a tetrahedron with its three valencies directed towards the other three corners; consequently the three valencies are not coplanar (*cf.* tertiary amines). These authors also assumed that there is no free rotation about the C=N double bond (*cf.* §2. IV), and therefore proposed configurations V and VI for the two isomers:



V VI

Many facts are in favour of geometrical isomerism, *e.g.*,

(i) If Ar = R, then isomerism disappears.

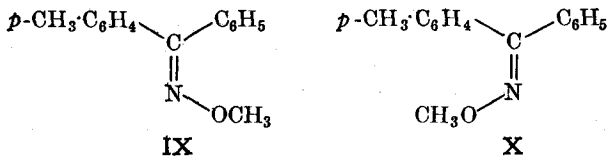
(ii) III and IV would be optically active; this is not found to be so in practice.

(iii) Absorption spectra measurements show that the two isomers have identical structures.

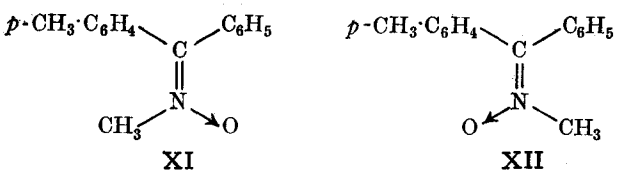
As pointed out above, Hantzsch and Werner chose structure I as the formula for the oximes, but examination of II shows that this would also satisfy the requirements for geometrical isomerism; structure I was chosen because oximes were known to contain the group >C=NOH. Later work, however, has shown that the problem is not so simple as this; methylation of an oxime (with methyl sulphate) usually produces a mixture of two compounds, one of which is the *O*-methyl ether, VII, and the other the *N*-methyl ether, VIII. These two are readily distinguished by the fact

VII VIII

that on heating with hydriodic acid, VII gives methyl iodide, whereas VIII gives methylamine. Thus, Semper and Lichtenstadt (1918) obtained *four* methyl derivatives of phenyl *p*-tolyl ketoxime, IX–XII. On treatment

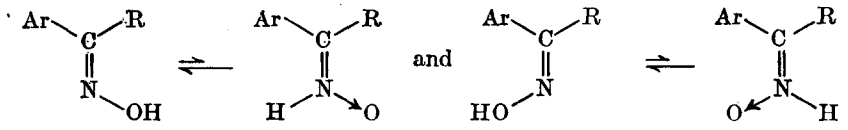


IX X

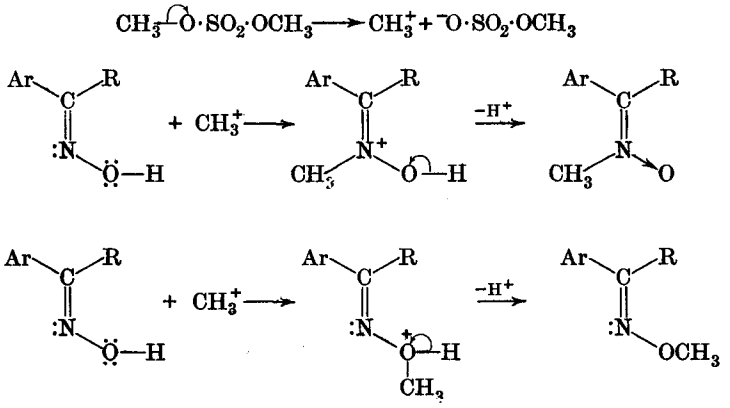


XI XII

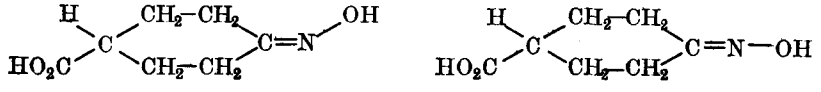
with concentrated hydriodic acid, two of these compounds gave methyl iodide, and therefore correspond to the *O*-methyl derivatives, IX and X; the other two compounds gave methylamine, and therefore correspond to the *N*-methyl derivatives, XI and XII. Thus it appears that oximes can exist in forms I and II. Brady (1916) considered that oximes in solution are a tautomeric mixture of I and II (*oximino-nitrone diad system*). Ultra-violet absorption spectra studies show that the spectra of the oximes are the same as those of the *O*-methyl ethers, whereas those of the *N*-methyl ethers are entirely different. Hence, if oximes are tautomeric mixtures of I and II, the equilibrium must lie almost completely on the oxime side, *i.e.*,



It is possible, however, that none of the nitrone form is present, but its methyl derivative is formed during the process of methylation. If we assume that methyl sulphate provides methyl carbonium ions, then it is possible that these ions attack the nitrogen atom (with its lone-pair) or the oxygen atom (with its two lone-pairs). This would result in the formation of the *N*- and *O*-methyl ethers, without having to postulate the existence of the oximino-nitrone tautomeric system.



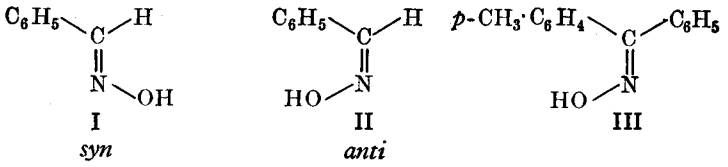
In the foregoing account, the geometrical isomerism of the oximes is based on the assumption that the nitrogen atom, in the oximino-form, exhibits the trigonal pyramidal configuration. Further proof for this configuration is obtained from the examination of the oxime of *cyclo*hexanone-4-carboxylic acid (XIII*a* or *b*). If the three nitrogen valencies are non-planar (*i.e.*, the



XIII*a* XIII*b*

N—O bond is not collinear with the C=N double bond), the configuration is XIII*a*, and it will therefore be optically active. If, however, the three nitrogen valencies are coplanar and symmetrically placed, then the configuration will be XIII*b*, and this will not be optically active, since it possesses a plane of symmetry. Mills and Bain (1910) prepared this oxime and resolved it; hence its configuration must be XIII*a*. This is readily explained on the modern theory of valency (§2c).

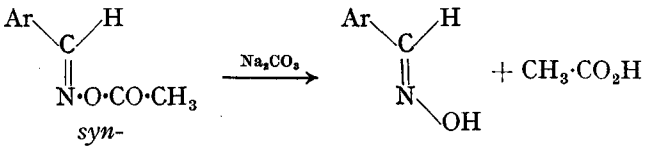
§2e. Nomenclature of the oximes. In oxime chemistry the terms *syn* and *anti* are used instead of the terms *cis* and *trans*. When dealing with aldoximes, the *syn*-form is the one in which both the hydrogen atom and the hydroxyl group are on the same side; when these groups are on opposite sides, the configuration is *anti*. Thus I is *syn*- and II is *anti*-benzaldoxime. With ketoximes, the prefix indicates the spatial relationship between the



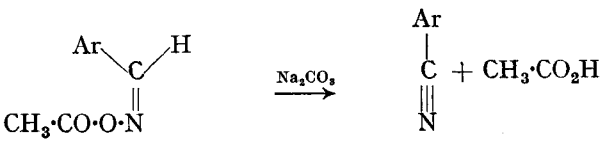
I *syn* II *anti* III

first group named and the hydroxyl group (*cf.* §4. IV). Thus III may be named as *syn-p*-tolyl phenyl ketoxime or *anti*-phenyl *p*-tolyl ketoxime.

§2f. Determination of the configuration of aldoximes. As we have seen, aromatic aldoximes can be obtained in two geometrical isomeric forms, the *syn* and the *anti*. Aliphatic aldoximes, however, appear to occur in one form only, and this is, apparently, the *anti*-form. The problem, then, with aromatic aldoximes, is to assign configurations to the stereoisomeric forms. The two forms (of a given aldoxime) resemble each other in many ways, but differ very much in the behaviour of their acetyl derivatives towards aqueous sodium carbonate. The acetyl derivative of one isomer regenerates the aldoxime; this form is known as the α-isomer. The other isomer, however, eliminates a molecule of acetic acid to form an aryl cyanide; this form is known as the β-isomer. Hantzsch and Werner (1890) suggested that the β-form readily eliminates acetic acid because the hydrogen atom and the acetoxy-group are close together, *i.e.*, the β-isomer is the *syn*-form. Such a view, however, is contrary to many experimental results (*cf.* §5 xi. IV), *i.e.*, the experimental results are:

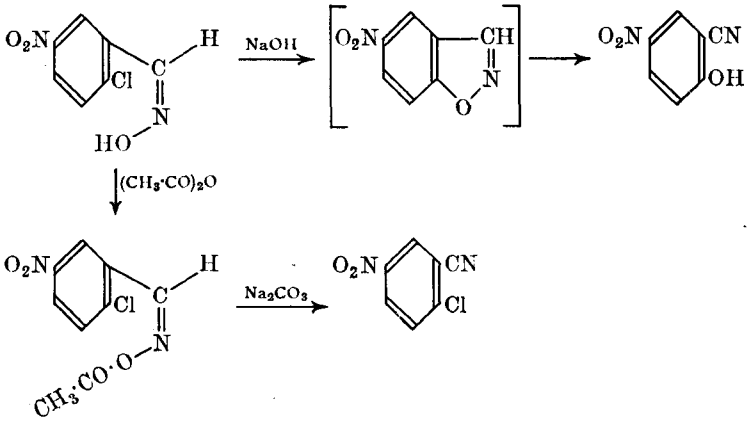


syn-



anti-

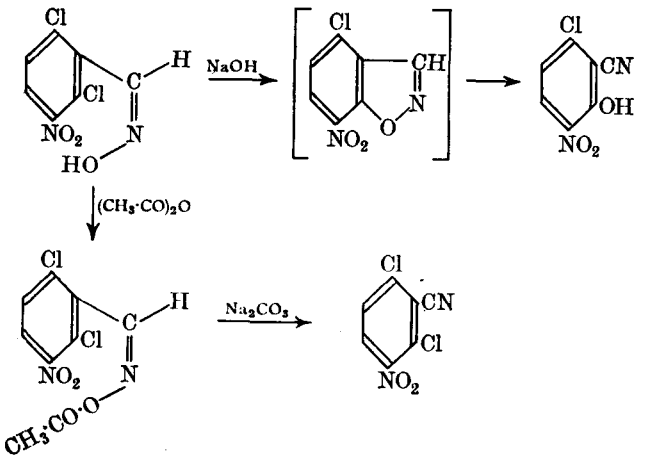
Brady and Bishop (1925) found that only one of the two isomers of 2-chloro-5-nitrobenzaldoxime readily gave ring closure on treatment with



sodium hydroxide. It therefore follows that this form is the *anti*-isomer (*cf.* method of cyclisation, §5 i. IV). It was also found that it was this isomer that gave the cyanide on treatment with acetic anhydride followed by aqueous sodium carbonate. Thus *anti*-elimination must have occurred, *i.e.*, the *β*-isomer is the *anti*-form. These reactions may be formulated as shown at foot of previous page.

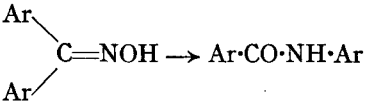
Actually, the ring compound produced, the 5-nitrobenz*iso*-oxazole, is unstable, and rearranges to nitrosalicylonitrile.

In a similar manner, Meisenheimer (1932) found that of the two isomeric 2 : 6-dichloro-3-nitrobenzaldoximes, it was the *anti*-isomer that gave ring closure, and was also the one that gave the cyanide. Hence, if *anti*-elimination is used as the criterion for these reactions, the configurations



of the *syn*- and *anti*-forms can be determined. It might be noted here, in passing, that since the *syn*-form was originally believed to form the cyanide, the configurations of the isomers in the literature up to 1925 (*i.e.*, before Brady's work) are the reverse of those accepted now.

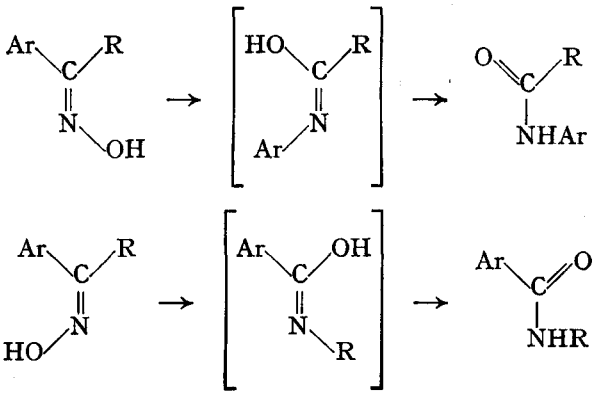
§2g. Determination of the configuration of ketoximes. The configurations of ketoximes have been mainly determined by means of the **Beckmann rearrangement** (1886). Aromatic ketoximes, *i.e.*, ketoximes containing at least one aromatic group, occur in two forms; aliphatic ketoximes appear to occur in one form only. When treated with certain acidic reagents such as sulphuric acid, acid chlorides, acid anhydrides, phosphorus pentachloride, etc., ketoximes undergo a molecular rearrangement, resulting in the formation of an acid amide:



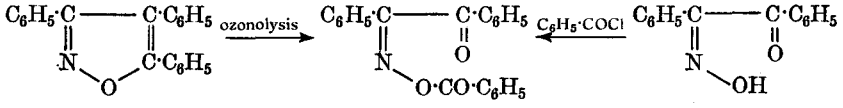
This rearrangement is known as the *Beckmann rearrangement* or *Beckmann transformation*. The best method is to treat an ethereal solution of the oxime with phosphorus pentachloride at a temperature below −20°. On the other hand, Horning *et al.* (1952) have found that a very good method for effecting the Beckmann rearrangement is to heat the oxime in polyphosphoric acid at 95° to 130°.

Hantzsch (1891) suggested that the course of the rearrangement indicated

the configuration of the oxime, and assumed that the *syn*-exchange of groups occurred since they were closer together in this isomer. This, again, was shown experimentally to be the reverse, *i.e.*, it is the *anti*-rearrangement that occurs, and not the *syn*; thus:

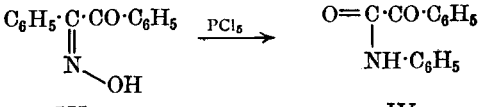


Meisenheimer (1921) subjected triphenyl*iso*-oxazole, I, to ozonolysis, and thereby obtained the benzoyl-derivative of *anti*-phenyl benzil monoxime, II. This configuration is based on the reasonable assumption that the ozonolysis proceeds without any change in configuration. Furthermore, the monoxime designated the β-isomer gave II on benzoylation, and so the configuration



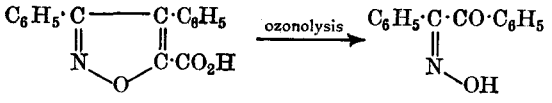
I II III

of the β-isomer, III, is determined. Meisenheimer then subjected this β-oxime (*i.e.*, the *anti*-phenyl oxime) to the Beckmann rearrangement, and obtained the anilide of benzoylformic acid, IV; thus the exchange of groups



III IV

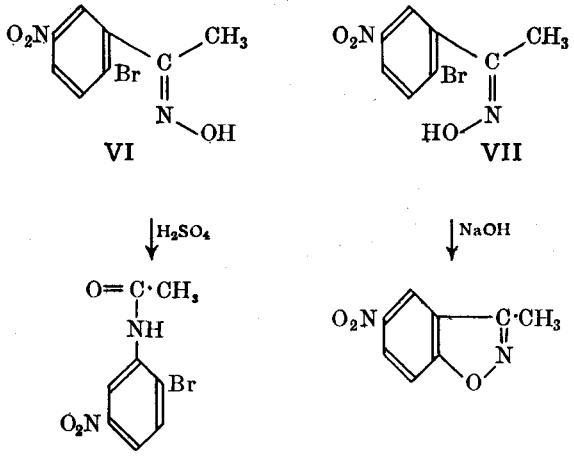
must occur in the *anti*-position. The configuration of the β-monoxime, III, is confirmed by the fact that it may be obtained directly by the ozonolysis of 3 : 4-diphenyl*iso*-oxazole-5-carboxylic acid, V (Kohler, 1924). Meisenheimer *et al.* (1925) also demonstrated the *anti*-rearrangement as follows.



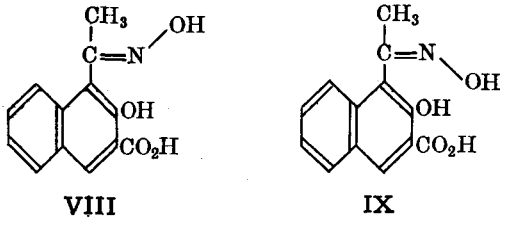
V III

The α-oxime of 2-bromo-5-nitroacetophenone is unaffected by sodium hydroxide, whereas the β-isomer undergoes ring closure to form 3-methyl-5-nitrobenz*iso*-oxazole; thus the α-oxime is the *syn*-methyl isomer VI, and the β-oxime the *anti*-methyl isomer VII. When treated with sulphuric acid or phosphorus pentachloride, the α-oxime underwent the Beckmann

rearrangement to give the *N*-substituted acetamide; thus the exchange occurs in the *anti*-positions.



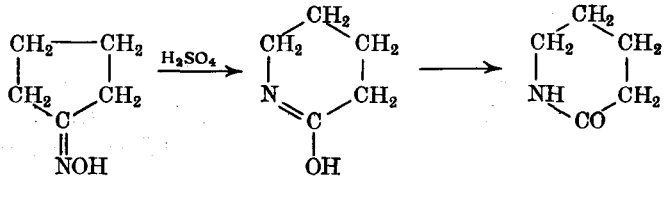
Further evidence for the *anti*-exchange of groups in the Beckmann rearrangement has been obtained by studying the behaviour of compounds exhibiting restricted rotation about a single bond, *e.g.*, Meisenheimer *et al.* (1932) prepared the two isomeric oximes of 1-acetyl-2-hydroxynaphthalene-3-carboxylic acid, VIII and IX, and of these two forms only one was resolvable. This resolvable isomer must therefore be IX, since asymmetry



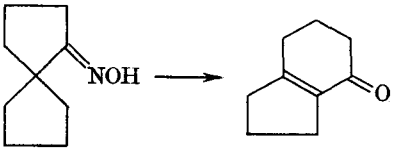
due to restricted rotation is possible only with this form (*cf.* §3. V). Meisenheimer found that the ethyl ester of IX, on undergoing the Beckmann rearrangement, gave the amide Ar·CO·NH·CH_3 (where Ar is the naphthalene part of the molecule), whereas the ethyl ester of VIII gave the amide CH_3·CO·NH·Ar. These results are in agreement with the *anti*-exchange of groups in each case.

Thus the evidence is all in favour of the *anti*-exchange of groups in the Beckmann rearrangement, and hence by using this principle, the Beckmann rearrangement may be used to determine the configuration of ketoximes.

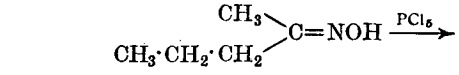
An interesting application of the Beckmann rearrangement is in the formation of heterocyclic rings, *e.g.*, when *cyclo*pentanonoxime is subjected to the Beckmann rearrangement, the nitrogen atom enters the ring (thus producing ring expansion) to form 2-piperidone (see also §2h).



On the other hand, Hill *et al.* (1956) have shown that the oximes of some spiro-ketones undergo abnormal Beckmann rearrangements in the presence of polyphosphoric acid, *e.g.*, spiro-[4 : 4]-nonanone-1-oxime gives hydrind-8 : 9-en-4-one:



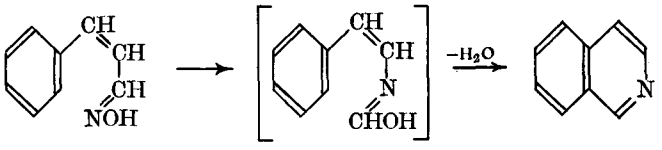
Although aliphatic ketoximes are not known in two isomeric forms, some may produce two products when subjected to the Beckmann rearrangement, *e.g.*, the oxime of pentan-2-one gives *N*-propylacetamide and *N*-methylbutyramide. The reason for this is uncertain; possibly oximes of this type are actually a mixture of the two forms; or alternatively, they exist in one



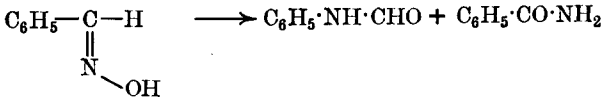
$$CH_3{\cdot}CO{\cdot}NH{\cdot}CH_2{\cdot}CH_2{\cdot}CH_3 + CH_3{\cdot}CH_2{\cdot}CH_2{\cdot}CO{\cdot}NH{\cdot}CH_3$$

stable form which, during the Beckmann rearrangement, is partially converted into the labile form which then undergoes the rearrangement (*cf.* benzaldoxime, below).

Whereas the majority of ketoximes undergo the Beckmann rearrangement, it appears that few aldoximes do so. In an attempt to prepare quinoline by the dehydration of cinnamaldoxime with phosphorus pentoxide, Bamberger and Goldschmidt (1894) actually obtained *iso*quinoline; the formation of the latter compound and not the former can only be reasonably explained on the assumption that the oxime first undergoes the Beckmann rearrangement, and the rearranged product then undergoes ring closure to form *iso*-quinoline. Recently, Horning *et al.* (1952) have shown that aldoximes can



be made to undergo the Beckmann rearrangement under the influence of polyphosphoric acid, *e.g.*, *syn*-benzaldoxime gives a mixture of formanilide and benzamide, the latter being produced by the conversion of the *syn*-

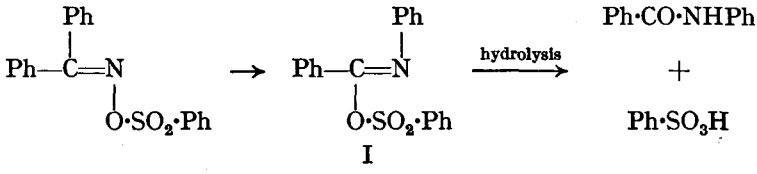


syn-isomer

anti-isomer

form into the *anti*; *anti*-benzaldoxime gives benzamide only. These results are in agreement with the configurations obtained by other methods (see §2f).

§2h. Mechanism of the Beckmann rearrangement. This rearrangement is an example of the 1,2-shift in which the migration origin is carbon and the migration terminus is nitrogen (see also 1,2-shifts, Vol. I, Ch. V). As we have seen above (§2g), an integral part of the rearrangement is the *anti* migration of the group. Since the oxime itself does not rearrange, it is reasonable to suppose that some intermediate is formed between the oxime and the reagent used to effect the rearrangement, and it is this intermediate which then rearranges. Kuhara *et al.* (1914, 1916) prepared the benzenesulphonate of benzophenone oxime and showed that this readily underwent rearrangement in neutral solvents in the absence of any acid catalyst to give an isomeric compound which, on hydrolysis, gave benzanilide and benzenesulphonic acid; thus:



I

Kuhara assigned structure I to this intermediate on the fact that its absorption spectrum was almost identical with that of the compound prepared by reaction between *N*-phenylbenzimidoyl chloride and silver benzenesulphonate:

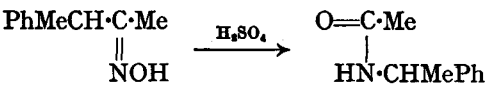
$$Ph{\cdot}CCl{=}NPh + AgO{\cdot}SO_2{\cdot}Ph \longrightarrow I + AgCl$$

Kuhara (1926) also showed that the rate of rearrangement of the benzophenone oxime ester is faster the stronger the acid used to form the ester; the order obtained was:

$$Ph{\cdot}SO_3H > CH_2Cl{\cdot}CO_2H > Ph{\cdot}CO_2H > Me{\cdot}CO_2H$$

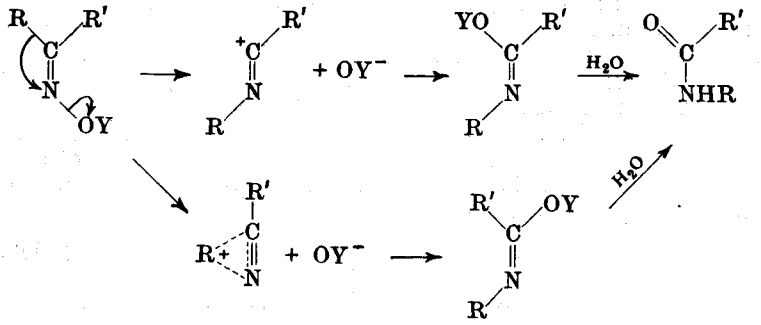
Chapman (1934) showed that the rate of rearrangement of benzophenone oxime picryl ester is faster in polar than in non-polar solvents. Thus the work of Kuhara and Chapman is strong evidence that the rate-determining step in the rearrangement is the ionisation of the intermediate.

Now let us consider the migration of the R or Ar group. This could be either intermolecular or intramolecular, but Kenyon *et al.* have shown it to be the latter; *e.g.*, in 1946, Kenyon *et al.* showed that when (+)-α-phenylethyl methyl ketoxime is treated with sulphuric acid the product, *N*-α-phenylethylacetamide, is almost 100 per cent. optically pure. Thus the migrating group never separates during the rearrangement, since if it did a racemised product would have been obtained. Furthermore, this retention of optical activity might be cited as evidence for the formation of a bridged-ion during the migration, since in such an ion the migrating group is not free and the "new partial" bond is formed on the *same* side as the bond which is breaking (see below).

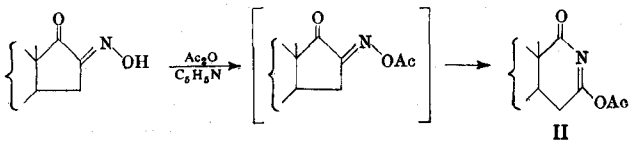


Another problem that arises here is: Does the anion separate completely during the ionisation or does it also migrate intramolecularly? The work of Kuhara and Chapman strongly suggests complete separation, and this is supported by the work of Brodskii *et al.* (1941), who found that when benzophenone oxime was treated with phosphorus pentachloride and then with water enriched with the isotope ^{18}O, the benzanilide obtained contained some of this isotope. Thus the oxygen atom of the oxime group

must have been completely removed in the ionisation stage (see below). The following mechanism is in agreement with all of the above facts (Y is PCl_4, MeCO, etc.); the lower set of equations is the alternative route *via*



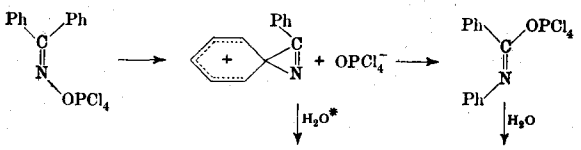
a bridged-ion. It might also be noted that when acid is used as the rearranging reagent, OY is probably OH_2^+. Support for this mechanism is the evidence obtained for the intermediate formation of the imidoyl ester (RN = CR·OY); compound II was obtained by Heard *et al.* (1959), who examined the rearrangement of a 17-keto-16-oxime (a steroid; Ch. XI):



II

It has been shown that when the migrating group is aryl, the rate of the rearrangement is accelerated when there is an electron-releasing group, *e.g.*, Me, in the *p*-position. This may be cited as evidence to support the formation of a bridged-ion (at least for migrating aryl groups).

On the basis of the above mechanism, we can now explain Brodskii's results as follows:

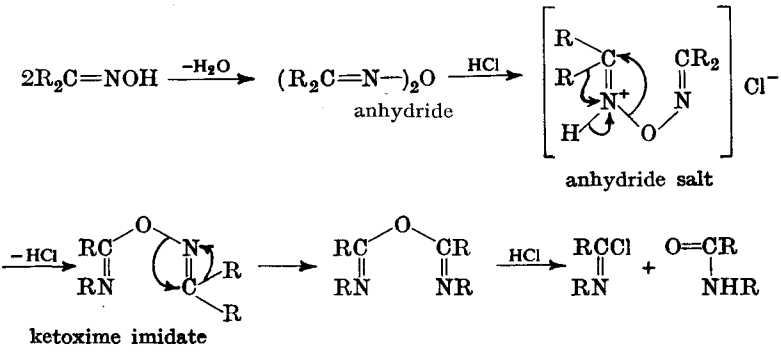


Stephen *et al.* (1956) have shown that *one* molecule of phosphorus pentachloride, phosphoryl chloride, thionyl chloride, or benzenesulphonyl chloride rearranges *two* molecules of the ketoxime to yield the corresponding amide and imidoyl chloride in approximately equimolecular amounts, *e.g.*,

$$2R_2C{=}NOH + PCl_5 \longrightarrow R{\cdot}CO{\cdot}NH{\cdot}R + R{\cdot}CCl{=}N{\cdot}R + POCl_3 + HCl$$

It has also been shown that hydrogen chloride is essential during the rearrangement, but that it does not itself cause the rearrangement of the

oxime. On the basis of these results, Stephen *et al.* have proposed the following mechanism for the Beckmann rearrangement of ketoximes. The reagent first produces some acid amide and imidoyl chloride, and the latter then dehydrates unchanged ketoxime to the anhydride which then reacts as shown:



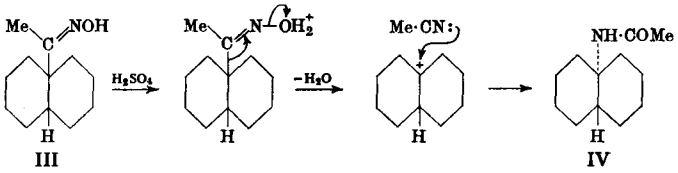
It is also suggested that other reagents which effect the Beckmann rearrangement may function as dehydrating agents for the formation of the ketoxime anhydride.

When a *trace* of the reagent is used, a large yield of amide is obtained. The mechanism is believed to be the same as that given above, provided that in the initial stage there is sufficient to form a trace of the ketoxime anhydride in the presence of hydrogen chloride. Rearrangement of the anhydride will now take place as above with the formation of the imidoyl chloride which can then dehydrate ketoxime to anhydride, itself being converted into the amide:

$$2R_2C{=}NOH + R{\cdot}CCl{=}N{\cdot}R \longrightarrow (R_2C{=}N{-})_2O + R{\cdot}CO{\cdot}NH{\cdot}R + HCl$$

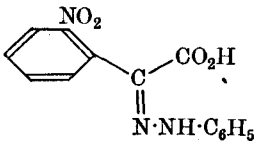
Thus the yield of amide increases at the expense of the imidoyl chloride.

It can be seen from the foregoing account that two mechanisms appear possible for the Beckmann rearrangement. Both are intramolecular, but now an intermolecular mechanism has also been proposed by Hill *et al.* (1962) who have reported an example in which the migrating group had the *inverted* configuration in the amide. These authors examined the rearrangement of 9-acetyl-*cis*-decalin oxime and have suggested the following mechanism:

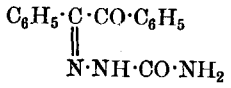


The authors identified methyl cyanide as a product of the reaction of III with phosphorus pentachloride, and also showed that methyl cyanide and *cis*-β-decalol in sulphuric acid gave IV.

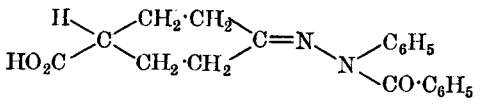
§2i. Stereoisomerism of some other tervalent nitrogen compounds containing a double bond. There are several other types of compounds besides the oximes in which the nitrogen atom is linked by a double bond. The other atom joined by this double bond may be a carbon atom (as in the oximes), or another nitrogen atom, and in both cases stereoisomerism is possible; *e.g.*, Krause (1890) obtained two isomeric forms of the phenylhydrazone of *o*-nitrophenylglyoxylic acid, I, and Hopper (1925) isolated two



I



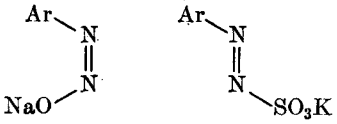
II



III

isomers of the monosemicarbazone of benzil, II. Mills and Bain (1914) resolved III; this is resolvable because of the non-planar configuration of the three nitrogen valencies (*cf.* the oximes, §2d). Karabatsos *et al.* (1962) have examined the NMR spectra of a number of ketone dinitrophenylhydrazones and semicarbazones, and have distinguished between the *syn*- and *anti*-forms, and have also calculated the amounts of each in solution. Phillips (1958) had already examined aldoximes by means of their NMR spectra.

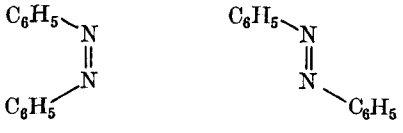
Many cases of geometrical isomerism are known in which the two forms are due to the presence of a nitrogen–nitrogen double bond. Examples of this type which have been most extensively studied are the diazoates, IV, the diazosulphonates, V, and the diazocyanides, VI (see Vol. I, Ch. XXIV, for an account of these compounds).



IV V VI

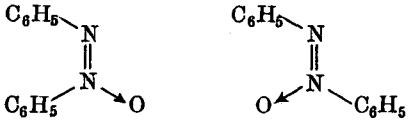
syn-form *anti*-form *anti*-form

Azobenzene is also an example of this type, and according to Hartley (1938), " ordinary " azobenzene is the *anti*-form.



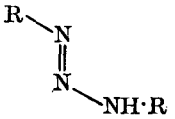
syn-azobenzene m.p. 71·4° *anti*-azobenzene m.p. 68°

Azoxybenzene (in which one nitrogen atom is tercovalent and the other quadricovalent) also exists in two geometrical isomeric forms, the *anti*-isomer being " ordinary " azoxybenzene.



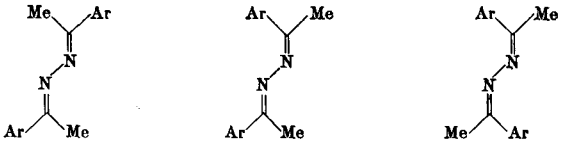
syn-azoxybenzene m.p. 86° *anti*-azoxybenzene m.p. 36°

Recently, Le Fèvre *et al.* (1951) have measured the dipole moments and the ultraviolet absorption spectra of a number of triazens, and have concluded that these compounds exist in the *anti*-configuration about the nitrogen–nitrogen double bond, *i.e.*, the configuration is:



These authors also believe that this *anti*-form is converted into an equilibrium mixture of the *anti*- and *syn*-forms when exposed to sunlight.

Harley-Mason *et al.* (1961) have offered evidence to show that they have isolated the three theoretically possible geometrical isomers of *o*-nitroacetophenone azine (Ar = *o*-$NO_2C_6H_4$-):



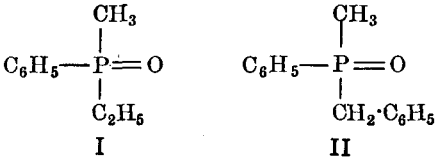
Their evidence was based on infra-red, ultraviolet and NMR spectra. This compound appears to be the first example of the isolation and characterisation of all three possible geometrical isomers of an azine.

§3. STEREOCHEMISTRY OF PHOSPHORUS COMPOUNDS

Nitrogen, as we have seen, can exhibit covalencies of 3 and 4; phosphorus (and arsenic), however, can exhibit covalencies of 3, 4, 5 and 6, and consequently gives rise to more possible configurations than nitrogen. In tercovalent compounds the valency disposition is tetrahedral (sp^3), one orbital being occupied by a lone-pair; and in quinquevalent compounds the valency disposition is trigonal bipyramidal (sp^3d). In quadricovalent unielectrovalent compounds one electron is transferred from the phosphorus or arsenic atom to the anion and the valency disposition is tetrahedral (sp^3) (see also §4b). When there are double bonds present, one is a σ- and the other is a π-bond; thus, in $POCl_3$, the shape is tetrahedral (see also §1).

§3a. Tercovalent phosphorus compounds. Since the electronic configuration of phosphorus is $(1s^2)(2s^2)(2p^6)(3s^2)(3p^3)$, it might be expected that suitable tercovalent compounds, R_3P, could be resolved, since the configuration would be a trigonal pyramid (*cf.* §2c). No tertiary phosphines, however, have yet been resolved, and the reason for this appears to be the same as for tertiary amines, *viz.*, that the phosphorus atom is in a state of oscillation. Calculation has shown that the frequency of this oscillation in phosphine is 5×10^6; this is slower than that of nitrogen ($2{\cdot}3 \times 10^{10}$), and if it could be brought to zero, then tertiary phosphines would be resolvable. Increasing the weight of the groups slows down the oscillation in phosphorus compounds, *e.g.*, replacement of the three hydrogen atoms by deuterium atoms changes the frequency to 6×10^3. It seems possible, therefore, that very large groups might produce phosphines which would be resolvable; and if not zero in these compounds, the oscillation certainly can be expected to be zero in ring compounds (*cf.* nitrogen, §2c). Thus, if chemical difficulties can be overcome, tercovalent phosphorus compounds would be resolvable (see also §4c.)

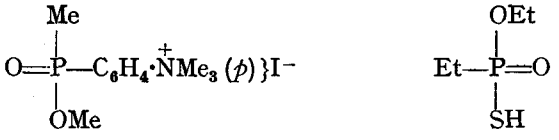
§3b. Quadricovalent and quinquevalent phosphorus compounds. The earliest phosphorus compounds to be resolved were the phosphine oxides, *e.g.*, Meisenheimer *et al.* (1911) resolved ethylmethylphenylphosphine



I II

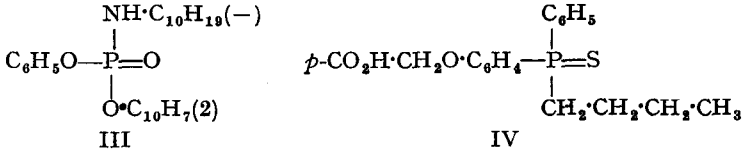
oxide, I, and benzylmethylphenylphosphine oxide, II. Recent measurements of the P—O (and As—O) bond length indicate that this bond is a double bond.

Some phosphine oxides that have been resolved recently are:



(McEwen *et al.*, 1956) (Aaron *et al.*, 1958)

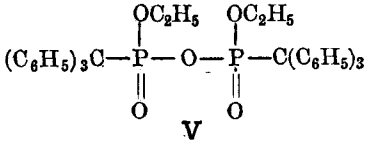
Kipping (1911) obtained two optically active forms of the *N*-(−)-menthyl derivative of 2-naphthylphenylphosphoramidic acid, III, and Davies and Mann (1944) resolved *n*-butylphenyl-*p*-carboxymethoxyphenylphosphine sulphide, IV.



III IV

Michalski *et al.* (1959) have prepared the phosphorus sulphenyl chloride, (EtO)EtP(=O)·SCl, in its (+)- and (−)-forms, and Green *et al.* (1961) have partially resolved phenylethylphosphinothiolic acid, PhEtP(=O)·SH.

Another interesting phosphorus compound from the point of view of optical isomerism is ethyl triphenylmethylpyrophosphonate, V. If the two phosphorus atoms are asymmetric, then V contains two similar asymmetric carbon atoms, and so its structure corresponds to the molecule *Cabd·Cabd*.



V

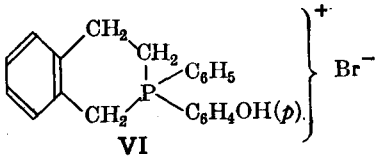
Thus there will be one racemic modification (composed of the pair of enantiomorphs) and one *meso*-form (*cf.* §7d. II). Hatt (1933) obtained two forms of compound V; both were inactive and so correspond to the racemic modification and the *meso*-form, but it was not possible to tell which was which.

Many attempts have been made to resolve quaternary phosphonium compounds, but until recently, all these attempts failed. This failure is attributed to the occurrence in solution of a "dissociation-equilibrium", which causes very rapid racemisation (see §4a).

$$[abcd\mathrm{P}]^{+}\mathrm{X}^{-} \rightleftharpoons abc\mathrm{P} + d\mathrm{X}$$

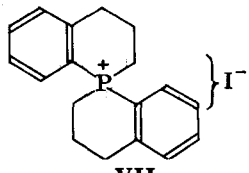
The earlier attempts to resolve phosphonium compounds were always carried

out on compounds containing at least one alkyl group; consequently dissociation in solution could occur, thereby resulting in racemisation. Holliman and Mann (1947) overcame this difficulty by preparing a much more stable type of phosphonium compound; these workers prepared a salt in which the phosphorus atom was in a ring, *viz.*, 2-phenyl-2-*p*-hydroxyphenyl-1 : 2 : 3 : 4-tetrahydro-*iso*phosphinolinium bromide, VI, and resolved it.



VI

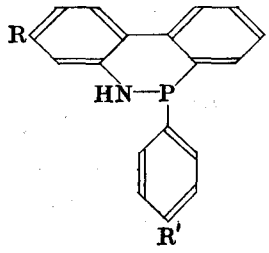
The resolution of 3-covalent compounds of phosphorus does not *prove* that the phosphorus atom has a tetrahedral configuration; it only proves that the phosphorus atom cannot be in the same plane as the other four groups



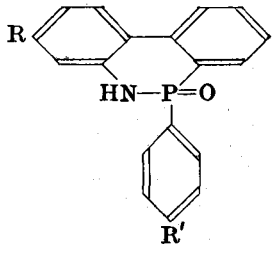
VII

attached to it. Mann *et al.* (1955), however, have now synthesised P-spiro-bis-1 : 2 : 3 : 4-tetrahydrophosphinolinium iodide (VII) and resolved it into (+)- and (−)-forms which have high optical stability. The phosphorus atom is not asymmetric in this compound; it is the *tetrahedral* disposition of the four valencies which produces the dissymmetric cation (*cf.* nitrogen, §2a; see also §4b).

Campbell *et al.* (1960) have prepared a series of azaphosphaphenanthrene (IX; *e.g.*, R = H, R′ = NMe_2), but could not resolve them. When the



IX



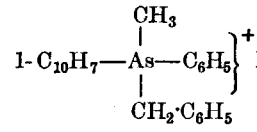
X

phosphine IX was oxidised with hydrogen peroxide, the phosphine oxide obtained, X, was resolved. Reduction of the (+)-oxide with lithium aluminium hydride gave the (−)-phosphine IX, and in the same way the reduction of the (−)-oxide gave the (+)-phosphine IX. It is not certain whether the optical activity in IX is due to an asymmetric tervalent phosphorus atom or to a rigid puckering of the molecular framework.

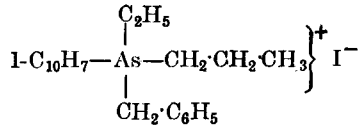
§4. STEREOCHEMISTRY OF ARSENIC COMPOUNDS

Arsenic, like phosphorus, can exhibit covalencies of 3, 4, 5 and 6; consequently these two elements show a great similarity to each other, and differ from nitrogen which has a maximum covalency of 4.

§4a. Quadricovalent and quinquevalent arsenic compounds. The first resolution of an arsonium compound was carried out by Burrows and Turner (1921). These workers obtained a solution of benzylmethyl-1-



I

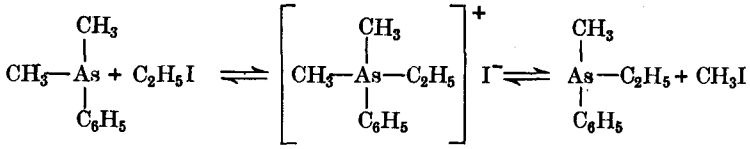


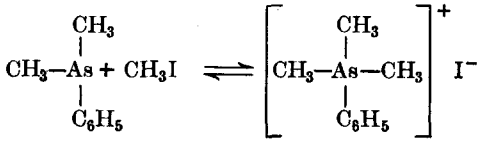
II

naphthylphenylarsonium iodide, I, that had a rotation of +12°, but racemised rapidly (in solution). Similarly, Kamai (1933) isolated the (+)-form of benzylethyl-1-naphthyl-*n*-propylarsonium iodide, II, which also racemised rapidly in solution. This rapid racemisation is believed to be due to a "dissociation-equilibrium" in solution. This explanation was suggested by Pope and Harvey (1901) to account for the racemisation of certain ammonium salts, but definite evidence for this theory was provided by Burrows and Turner (1921) in their work on arsonium salts. If this dissociation-equilibrium occurs, then in solution there will be:

$$[abcd\text{As}]^+\text{I}^- \rightleftharpoons abc\text{As} + d\text{I}$$

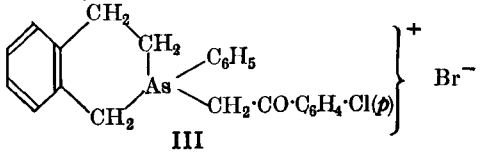
Burrows and Turner showed that when dimethylphenylarsine is treated with ethyl iodide, the expected ethyldimethylphenylarsonium iodide is





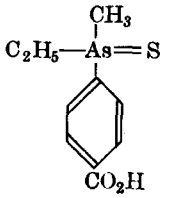
obtained, but at the same time a considerable amount of trimethylphenylarsonium iodide is also formed. These results are readily explained by the dissociation-equilibrium theory.

Since all the arsonium compounds investigated contained at least one alkyl group, Holliman and Mann (1943) prepared an arsonium compound with the arsenic atom in a ring, in the hope of stabilising the compound (*cf.* phosphorus, §3b). These authors prepared 2-*p*-chlorophenacyl-2-phenyl-1 : 2 : 3 : 4-tetrahydro-*iso*arsinolinium bromide, III, resolved it, and found that it did not racemise in solution at room temperature.

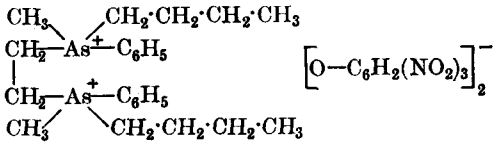


III

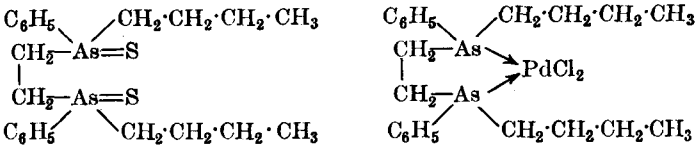
Although phosphine oxides of the type *abc*PO have been resolved (§3b), similar arsine oxides have not; the reason for this is obscure. On the other hand, arsine sulphides have been resolved, *e.g.*, Mills and Raper (1925) resolved *p*-carboxyphenylmethylethylarsine sulphide, IV.



IV



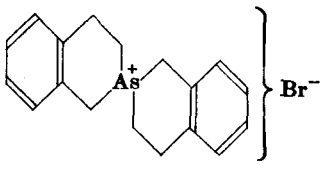
V



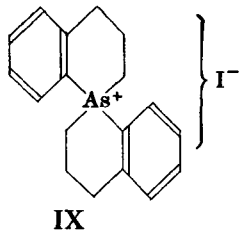
VI VII

Chatt and Mann (1939) prepared ethylene-1 : 2-bis(*n*-butylmethylphenylarsonium) picrate, V, ethylene-1 : 2-bis(*n*-butylphenylarsine sulphide), VI, and ethylene-1 : 2-bis(*n*-butylphenylarsine)-dichloropalladium, VII, and obtained each compound in two forms. Each of these compounds is of the type C*abd*·C*abd*, and hence each should exist in one racemic modification and one *meso*-form (*cf.* §7d. II). As has already been stated, two forms of each were isolated; both were inactive, but the authors had no evidence for deciding which was which.

It has already been pointed out above that Holliman and Mann prepared



VIII

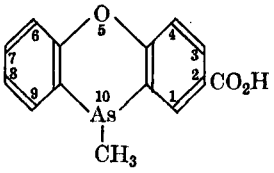


IX

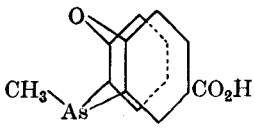
the optically stable arsonium compound III. These authors, in 1945, also resolved an arsonium compound of the spiran type, *viz.*, As-spiro-bis-1 : 2 : 3 : 4-tetrahydro-*iso*arsinolinium bromide, VIII. This does not contain an asymmetric arsenic atom; the optical activity is due to the asymmetry of the molecule (the two rings are perpendicular to each other), and this is evidence that the four valencies of arsenic are arranged tetrahedrally (see also §4b). Mann *et al.* (1960) have also resolved compound IX.

§4b. Tercovalent arsenic compounds. The electronic configuration of arsenic is $(1s^2)(2s^2)(2p^6)(3s^2)(3p^6)(3d^{10})(4s^2)(4p^3)$. Thus the configuration of tercovalent arsenic compounds will be a trigonal pyramid (*cf.* phosphorus,

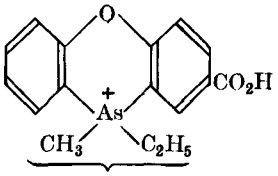
§3a). Physico-chemical evidence (X-ray analysis, spectroscopy and electron diffraction) has shown that in tercovalent compounds the arsenic atom is at the apex of a tetrahedron, and that the intervalency angle is $100 \pm 4°$. It has also been shown that the arsenic is in a state of oscillation, the frequency of this oscillation through the plane of the three hydrogen atoms in arsine being 16×10^4. This is slower than that of phosphorus (5×10^6), and very much slower than that of nitrogen ($2{\cdot}3 \times 10^{10}$). Thus, preventing the oscillation of the arsenic atom, possibly by attachment to very large groups, should lead to the isolation of optically active tercovalent compounds. So far, however, all attempts to resolve compounds of the type As*abc* have failed (*cf.* nitrogen and phosphorus). On the other hand, tercovalent arsenic compounds in which arsenic has two of its valencies occupied in a ring compound have been resolved; the ring structure prevents oscillation of the arsenic atom (*cf.* Tröger's base, §2c). Thus Lesslie and Turner (1934) resolved 10-methylphenoxarsine-2-carboxylic acid, I. These authors suggested that the assymetry of the molecule is due to the presence of a folded structure about the O—As axis, as well as the asymmetry due to the presence of an asymmetric arsenic atom (see structure II). This molecule



I

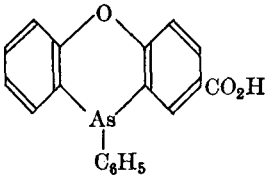


II

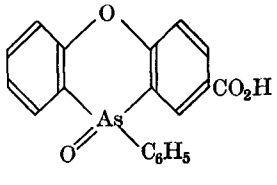


III

and its mirror image are not superimposable. It might be noted, however, that the position of the methyl group with respect to the O—As axis is uncertain (*cf.* the arsanthrens, below). This folded structure is reasonable in view of the fact that the valency angle of oxygen is also approximately 104°; if the molecule were planar, then the valency angles of both arsenic and oxygen would be in the region of 120°, which is a very large increase from the normal valency angle. When each enantiomorph of II is treated with ethyl iodide, the *same* racemised product is obtained. This is due to the fact that when the arsonium compound, III, is formed, the asymmetric quaternary arsenic atom is racemised owing to the dissociation-equilibrium.



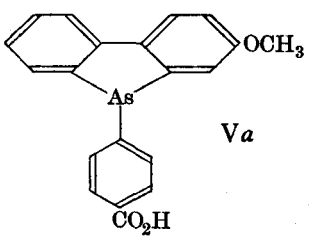
IV



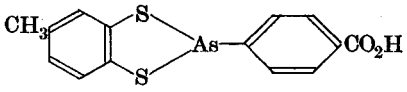
V

Lesslie and Turner (1936) also resolved 10-phenylphenoxarsine-2-carboxylic acid, IV. This compound was very stable, and oxidation to the arsine oxide, V, gave a completely racemised product.

Campbell *et al.* (1956) have resolved some substituted 9-arsafluorenes, *e.g.*, 9-*p*-carboxyphenyl-2-methoxy-9-arsafluorene (V*a*). Campbell (1956) has also resolved 2-*p*-carboxyphenyl-5-methyl-1 : 3-dithia-2-arsaindane (V*b*). This compound is optically stable in chloroform solution, but is racemised in aqueous sodium hydroxide. Campbell believes that this racemisation is due to the fission of the As—S bonds by aqueous alkali, and subsequent reversal of the reaction by acid, a type of behaviour observed in triaryl

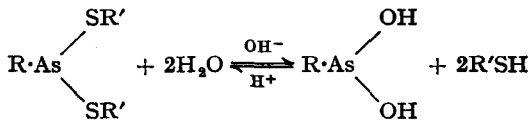


Va

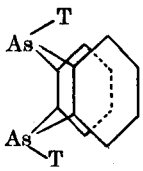


V*b*

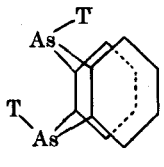
thioarsenites (Klement *et al.*, 1938). Furthermore, Cohen *et al.* (1931) have shown that in sodium hydroxide solution, alkyl thioarsenites exist in equilibrium with thiol and arsenoxide:



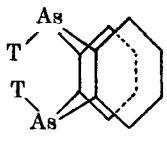
Chatt and Mann (1940) prepared 5 : 10-di-*p*-tolyl-5 : 10-dihydroarsanthren, and pointed out that if the valency angle of arsenic remains constant at its normal angle (of approximately 100°), then the structure will be folded, and consequently the three geometrical isomers, VI, VII and VIII, are apparently possible (T represents the *p*-tolyl group). Chatt and Mann also



VI

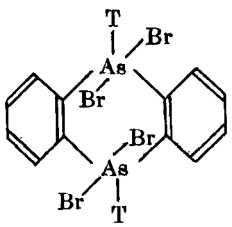


VII



VIII

pointed out that evidence obtained from models constructed to scale showed that the two *p*-tolyl radicals (T) in VIII would almost be coincident, and hence this isomer cannot exist. These authors isolated two optically inactive forms, but were unable to say which was which. When each compound was treated with bromine, both gave the *same* tetrabromide which, on hydrolysis, gave only one tetrahydroxide. The loss of isomerism in the tetrabromide (and in the tetrahydroxide) may be explained as follows. Bromination of VI and VII converts tercovalent arsenic into quinquecovalent arsenic, and in the latter state the ring valency angles of the arsenic become 120°, and so the arsanthren nucleus is *now* planar. Thus both the



IX

forms VI and VII would give the same tetrabromide, IX (the same is true for the tetrahydroxide); the tetrabromide should thus be planar, the configuration of each arsenic atom being trigonal bipyramidal in the 5-covalent state (Fig. 2).

Quinquevalent phosphorus and arsenic can make use of the $3d$ or $4d$ orbitals, respectively (*cf.* nitrogen, §2b). Thus nitrogen has a maximum covalency of 4, whereas that of phosphorus and arsenic is 5 or 6, *e.g.*, the covalency of 6 is exhibited by phosphorus in *solid* phosphorus pentachloride; X-ray diffraction shows this " molecule " (in the solid state) is $PCl_4^+ PCl_6^-$.

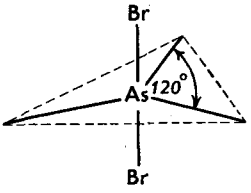


FIG. 6.2.

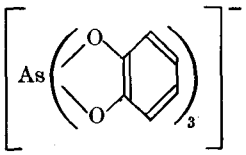
Phosphorus, which is $(1s^2)(2s^2)(2p^6)(3s^2)(3p^3)$ in the ground state, may become $(1s^2)(2s^2)(2p^6)(3s)(3p^3)(3d)$ in its " valence state ", since the $3s$ and $3d$ orbitals have energy levels which are close together. Kimball (1940) showed, by calculation, that this arrangement, *i.e.*, sp^3d, could give rise to the stable trigonal bipyramidal configuration. This consists of three equivalent coplanar orbitals pointing towards the corners of an equilateral triangle, and two orbitals perpendicular to this plane (see Fig. 2). Electron diffraction studies of the vapours of phosphorus pentachloride and pentafluoride indicate the trigonal bipyramidal configuration in these molecules. The phosphonium ion might possibly be formed from this trigonal bipyramid by the transference of one of the electrons, or by the transference of a $3s$ electron and hybridisation of the $(3s)(3p^3)$ orbitals; in either case the tetrahedral configuration of the phosphonium ion can be asymmetric, but only in the case of the hybridisation of the $(3s)(3p^3)$ orbitals will the four bonds be equivalent. Since the properties of phosphonium compounds are in agreement with the equivalence of the four bonds, it therefore appears, on theoretical grounds, that the tetrahedral configuration with the phosphorus atom at the centre is the probable one.

From the experimental side, the preparation of optically active spiro-compounds of phosphorus (§3b) and of arsenic (§4a) proves the tetrahedral configuration of these atoms. Earlier work by Mann *et al.* (1936, 1937) has also definitely established this configuration. These authors prepared compounds of the type $[R_3As—CuI]_4$ by combination of tertiary arsines or phosphines with cuprous iodide (or silver iodide); in these compounds the phosphorus or arsenic is 4-covalent, and X-ray analysis studies of the arsenic compound showed that the arsenic atom is at the centre of a tetrahedron. Since the corresponding phosphorus compounds are isomorphous, the configuration of the phosphorus is also tetrahedral.

In the *solid* state, phosphorus and arsenic compounds may contain a negatively charged phosphorus or arsenic atom, *e.g.*, $PCl_4^+ PCl_6^-$ (see above). In this condition, the phosphorus acquires an electron to become

$$---(3s)(3p^3)(3d^2),$$

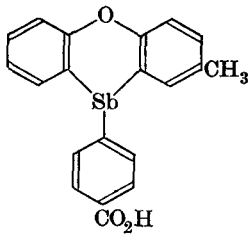
and the arsenic also acquires an electron to become $---(4s)(4p^3)(4d^2)$. In both cases the configuration is octahedral (six sp^3d^2 bonds), *e.g.*, the following compound has been resolved (Rosenheim *et al.*, 1925).



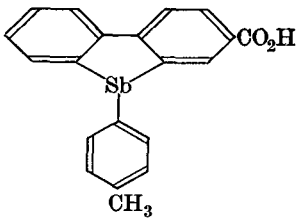
Harris *et al.* (1956) have shown that a negatively charged phosphorus atom can also exist in *solution*; these authors showed that triphenyl phosphite dibromide ionises in methyl cyanide solution as follows:

$$2P(OPh)_3Br_2 \rightleftharpoons [P(OPh)_3Br]^+ + [P(OPh)_3Br_3]^-$$

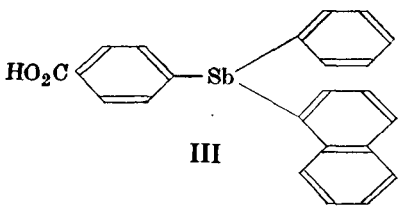
§4c. Stereochemistry of antimony compounds. Some optically active tervalent antimony compounds have been prepared, the phenoxstibine (I) and the stibiafluorene (II; Campbell, 1947, 1950). The asymmetry in I is probably due to the folding about the O—Sb axis (*cf.* phenoxarsines, §4b). Campbell *et al.* (1958) have also resolved the stibine (III).



I



II



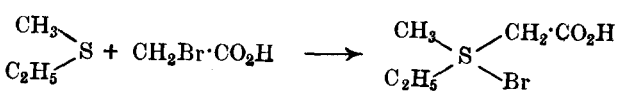
III

It is of interest to note, in this connection, that calculations by Weston (1954) have led him to the conclusion that tervalent antimony, arsenic and sulphur compounds should be stable to inversion at room temperature. On the other hand, similar compounds of phosphorus would be optically stable only at low temperatures, and those of nitrogen not at all.

§5. STEREOCHEMISTRY OF SULPHUR COMPOUNDS

Various types of sulphur compounds have been obtained in optically active forms, and although the general picture of the configurations of these molecules is quite clear, the details of the nature of the bonds of the central sulphur atom are in a state of flux (see §5e).

§5a. Sulphonium salts. Pope and Peachey (1900) prepared carboxymethylethylmethylsulphonium bromide by the reaction between ethyl methyl sulphide and bromoacetic acid, and formulated the reaction as follows:



At this time (before the electronic theory of valency, 1916), sulphur was believed to be quadricovalent, and so Pope and Peachey accounted for the optical activity of this compound (see below) by assuming that the sulphur atom was at the centre of a tetrahedron, *i.e.*, the configuration was similar to carbon. According to the electronic theory of valency, however, sulphur

is tercovalent unielectrovalent in sulphonium salts, and the valency disposition is (sp^3), one orbital being occupied by a lone-pair of electrons (Fig. 3). This molecule is not superimposable on its mirror image, and hence can, at least theoretically, exist in two optically active forms. This bromide was treated with silver (+)-camphorsulphonate and the salt

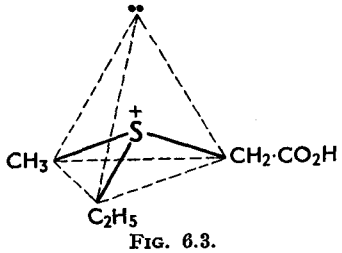
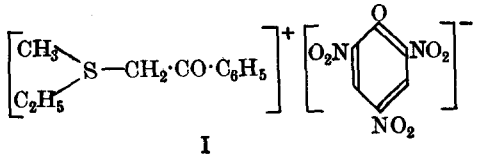


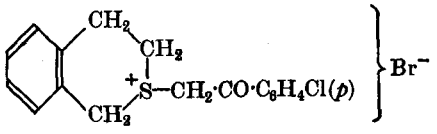
FIG. 6.3.

obtained was fractionally crystallised from a mixture of ethanol and ether. Pope and Peachey found that the (+)-sulphonium camphorsulphonate was the less soluble fraction, and had an M_D of +68°. Since the rotation of the (+)-camphorsulphonate ion is about +52°, this leaves +16° as the contribution of the sulphonium ion to the total rotation (see §12. I). Although this does not prove conclusively that the sulphur compound is optically active, it is certainly strong evidence in its favour. Final proof was obtained by replacement of the camphorsulphonate ion by the platinichloride ion to give $[CH_3(C_2H_5)\cdot S\cdot CH_2\cdot CO_2H]_2{}^+PtCl_6^=$; this compound had an $[\alpha]_D$ of +4·5° in water. In a similar way, Smiles (1900) prepared ethylmethylphenacylsulphonium picrate, I, in two optically active forms, one



I

with an $[\alpha]_D$ of +8·1° and the other −9·2°. A more recent example of an optically active sulphonium salt is one with the sulphur atom in a ring; this compound, II, was obtained as the optically active ion with the picrate Mann and Holliman, 1946).



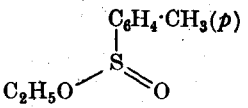
II

§5b. Sulphinic esters. Phillips (1925) partially resolved sulphinic esters, $R\cdot SO_2R'$, by means of the kinetic method of resolution (§10 vii. II). Two molecules of ethyl *p*-toluenesulphinate were heated with one molecule of (−)-menthyl alcohol or (−)-*sec.*-octyl alcohol, *i.e.*, the sulphinate was subjected to alcoholysis. Now, if the sulphinate is a racemic modification, then the (+)- and (−)- forms will react at *different* rates with the optically active alcohol (see §§2, 7b. II). Phillips actually found that the (+)-ester reacted faster than the (−)-ester. If we represent the ester by E, the alcohol by A, and unchanged ester by E_r, then the following equation symbolises the alcoholysis:

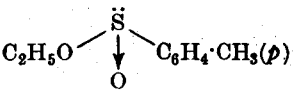
$$(+)E + (-)E + (-)A \rightarrow [(+)E(-)A] + [(-)E(-)A] + (+)E_r + (-)E_r$$

Since [(+)E(−)A] is greater than [(−)E(−)A], it therefore follows that $(+)E_r$ is less than $(-)E_r$; thus a partial resolution has occurred. The unchanged ester, having a lower boiling point than the new ester, distilled off first; this contained more of the (−)-form. The residual ester (the higher boiling fraction) was then heated with a large excess of ethanol; alcoholysis again occurred, this time the (−)-alcohol (menthol or octyl) being displaced to regenerate the original ethyl *p*-toluenesulphinate. This resulted in a fraction containing more of the (+)-form.

To account for the optical activity of these sulphinates, the older formula I, with quadricovalent sulphur linked to the oxygen atom by a double bond, was replaced by formula II, in which the sulphur atom is at the centre of the tetrahedron, but one corner is occupied by a lone-pair of electrons



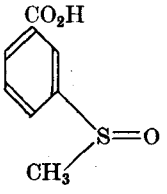
I



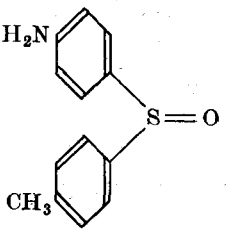
II

(*cf.* Fig. 3). In I, the sulphur atom was considered to be at the centre of a tetrahedron, and the molecule is *flat*, and consequently is superimposable on its mirror image. Molecule II, however, is asymmetric, and so is optically active. Recent evidence, however, is now in favour of structure I, and the molecule is *not* flat (see §5e). The formulæ of sulphoxides, etc., will therefore be written with double bonds.

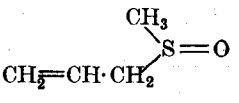
§5c. Sulphoxides. Sulphoxides of the type R·SO·R′ have also been resolved; sulphoxides I and II were resolved by Phillips *et al.* (1926), and Karrer *et al.* (1951) obtained III in the (−)-form and the racemic modification.



I



II

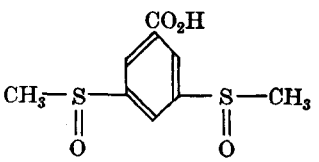


III

Bell and Bennett (1927) investigated disulphoxides of the type

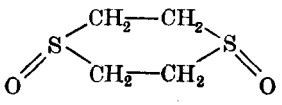
$$CH_3{\cdot}SO{\cdot}CH_2{\cdot}CH_2{\cdot}SO{\cdot}CH_3.$$

This molecule contains two similar asymmetric carbon atoms and so is of the type *Cabd·Cabd*. Thus it should exist in one racemic modification and one *meso*-form. Bell and Bennett failed to resolve this compound, but succeeded in resolving the following disulphoxide.

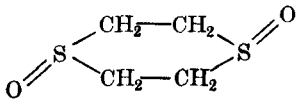


If the former disulphoxide (the dioxide of a 1 : 4-dithian) is converted

into the corresponding ring compound (*i.e.*, into a cyclic 1 : 4-dithian), then two geometrical isomers are possible, neither of which is resolvable; these two forms have been isolated by Bell and Bennett (1927, 1929). Shearer (1959) has examined the *trans*-form by X-ray analysis and showed that the ring is in the chair form with the S=O in *trans* and axial positions.

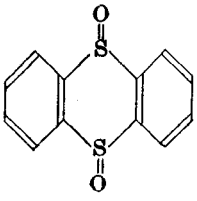


cis

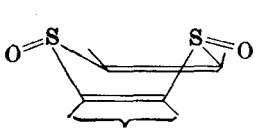


trans

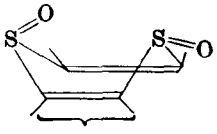
Thianthren dioxide, IV, also exists in two geometrical isomeric forms, α, m.p. 284°, $\mu = 1{\cdot}7$ D; and β, m.p. 249°, $\mu = 4{\cdot}2$ D (Bergmann *et al.*, 1932). On the basis of these dipole moments, Bergmann assigned the *trans*-configuration to the α-form and the *cis*-configuration to the β-form. Hosoya *et al.* (1957) have examined the α-form by X-ray analysis and showed it was boat-shaped (only this part of the molecule is shown in the diagrams), with the molecule folded along the S—S axis (*cf.* the dithian dioxides above). These authors also showed that this α-form has the *anti-cis*-configuration



IV



α-form



β-form

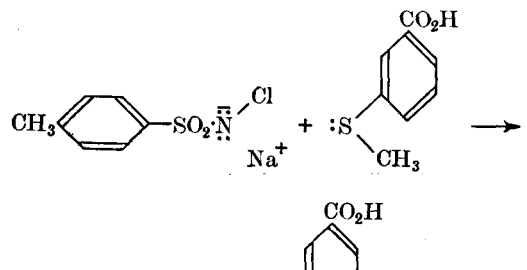
of the two S=O bonds. The β-form is therefore assumed to be a *trans*-form. Thus the configurations are the *reverse* of those given by Bergmann. When either of these disulphoxides is oxidised to the disulphone, both give the same compound (Hosoya, 1958).

It is of interest to note, in connection with optically active sulphoxides, that Schmid and Karrer (1948) have isolated *sulphoraphen* from its glycoside which occurs in radish seed. These authors showed that sulphoraphen is a lævorotatory oil which owes its optical activity to the presence of a sulphoxide group.

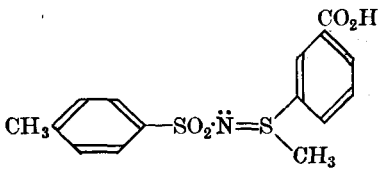
$$CH_3{\cdot}SO{\cdot}CH{=}CH{\cdot}CH_2{\cdot}CH_2{\cdot}NCS$$

sulphoraphen

§5d. Sulphilimines. Chloramine T reacts with alkyl sulphides to form sulphilimines, *e.g.*,



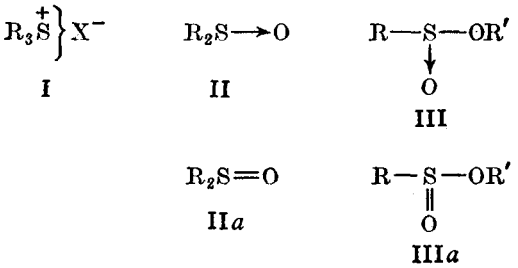
The electronic structure of this molecule appears to be uncertain; one possibility has been given above, and in this one the sulphur atom is asymmetric (it is of the type that occurs in the sulphonium salts). An alternative electronic structure is:



In this structure, the sulphur atom can still be asymmetric (see §5e). This sulphilimine has been resolved by Kenyon *et al.* (1927).

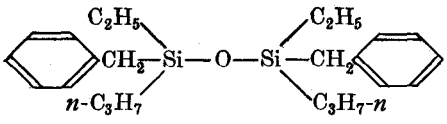
It seems likely that sulphilimines are resonance hybrids of the above two contributing structures.

§5e. The valency disposition of the sulphur atom. The electronic configuration of sulphur is $(1s^2)(2s^2)(2p^6)(3s^2)(3p^4)$. As we have seen, the older formulæ (II*a*) and (III*a*) for sulphoxides and sulphinic esters were replaced by Phillips by (II) and (III) respectively. However, in the light of more recent work, these compounds are now believed to contain double bonds, *e.g.*, the length of the S—O bond in sulphoxides and sulphones is shorter than the single S—O bond.

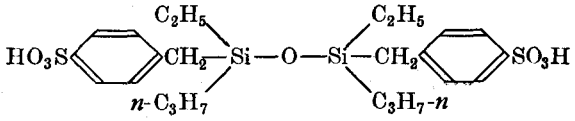


It has already been pointed out that these multiple bond formulæ were rejected on the grounds that such molecules, on the assumption that the sulphur atom was quadrivalent and at the centre of a tetrahedron, would be flat and hence not optically active. If we consider the shapes of optically active sulphur compounds from the point of view of the ideas discussed in §1, then in the formulæ (I), (II*a*) and (III*a*), the sulphur atom has one lone-pair of electrons (these are not shown in the formulæ), three σ- and one π-bond. Thus the bond spatial arrangement will be tetrahedral, the lone-pair occupying one of these orbitals. Consequently each molecule will be a trigonal pyramid and is not superimposable on its mirror image when all three groups are different. It might be noted here that the double bonds are composed of one σ- and one d_π-p_π bond. In these compounds the *d* orbitals are produced by promotion of one 3*s* and one 3*p* electron to 3*d*; this is possible because of the small energy differences between the orbitals concerned. In sulphonium salts, since only three single bonds and one lone-pair are present, the hybridisation is sp^3 (tetrahedral); one electron has been transferred to the halogen atom, thereby producing the positively charged sulphonium ion.

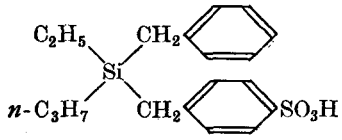
§6. Stereochemistry of silicon compounds. Kipping (1907) prepared benzylethylpropylsilicyl oxide, I, and isolated one form of it. If the silicon atom has a tetrahedral configuration, this molecule is of the type C*abd*·C*abd*,



I

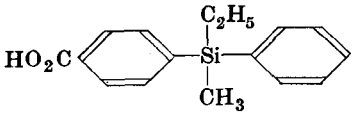


II



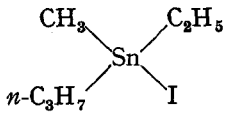
III

i.e., it should exist in (+)-, (−)- and *meso*-forms. When I was sulphonated to give II, the latter compound was resolved. Challenger and Kipping (1910) also resolved the silane III, and Eaborn *et al.* (1958) have resolved the silane IV.

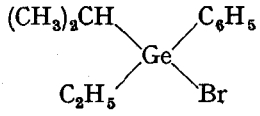


IV

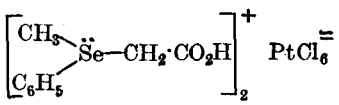
§7. Stereochemistry of tin compounds. Pope and Peachey (1900) obtained ethylmethyl-*n*-propylstannonium iodide in the dextrorotatory form; concentration of the mother liquor also gave this (+)-form. Thus we have an example of asymmetric transformation (§10 iv. II).



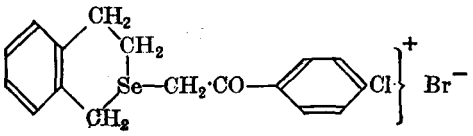
§8. Stereochemistry of germanium compounds. Schwarz and Lewinsohn (1931) obtained the (+)-form of ethylphenyl*iso*-propylgermanium bromide, but failed to get the (−)-form; this latter form appears to racemise in the mother liquor.



§9. Stereochemistry of selenium compounds. Pope *et al.* (1902) resolved carboxymethylmethylphenylselenonium bromide in the same way as the corresponding sulphonium salts (§5a); they obtained the active platinichloride.

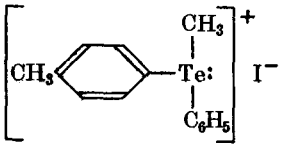


Mann *et al.* (1945) also resolved the following selenonium salt:

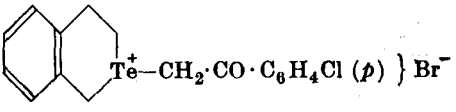


So far, attempts to resolve selenoxides have failed.

§10. Stereochemistry of tellurium compounds. Lowry *et al.* (1929) obtained the optically active forms of methylphenyl-*p*-tolyltelluronium iodide, I, and Mann *et al.* (1945) have resolved II.



I



II

READING REFERENCES

Gilman (Ed.), *Advanced Organic Chemistry*, Wiley (1943, 2nd ed.). Ch. 4, pp. 400–443. Optical Isomerism of Elements other than Carbon.

Dickens and Linnett, Electron Correlation and Chemical Consequences, *Quart. Reviews* (*Chem. Soc.*), **1957, 11,** 291.

Gillespie and Nyholm, Inorganic Stereochemistry, *Quart. Reviews* (*Chem. Soc.*), 1957, **11,** 339.

Organic Reactions, Wiley. Vol. 11 (1960). Ch. 1. The Beckmann Rearrangement.

Mann, *The Heterocyclic Derivatives of P, As, Sb, Bi, and Si*, Interscience Publishers (1950).

Campbell and Way, Synthesis and Stereochemistry of Heterocyclic Phosphorus Compounds, *J.C.S.*, **1960,** 5034.

Abrahams, The Stereochemistry of Sub-group VIB of the Periodic Table, *Quart. Reviews* (*Chem. Soc.*), **1956, 10,** 407.

McCasland and Proskow, Synthesis of an Image-Superposable Molecule which Contains no Plane or Centre of Symmetry, *J. Amer. Chem. Soc.*, 1956, **78,** 5646.

Klyne and de la Mare (Ed.), *Progress in Stereochemistry*, Butterworth. Vol. II (1958). Ch. 6. The Stereochemistry of the Group V Elements.

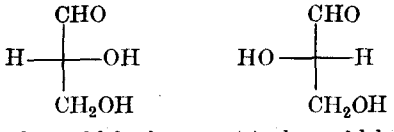
CHAPTER VII

CARBOHYDRATES

THIS chapter is mainly concerned with the stereochemistry of the carbohydrates and the structures of the disaccharides and polysaccharides. It is assumed that the reader is familiar with the open-chain structures and general reactions of the monosaccharides (for an elementary account of these compounds, see Vol. I, Ch. XVIII).

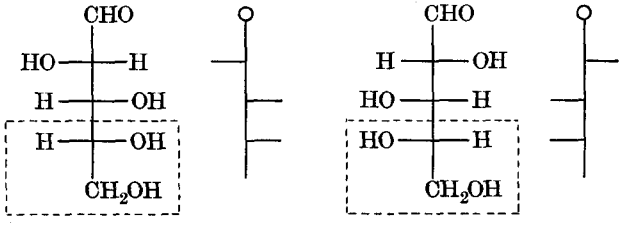
§1. DETERMINATION OF THE CONFIGURATION OF THE MONOSACCHARIDES

Aldotrioses. There is only one aldotriose, and that is glyceraldehyde. As we have seen (§5. II), the enantiomorphs of this compound have been chosen as the *arbitrary* standards for the D- and L- series in sugar chemistry:



D(+)-glyceraldehyde L(−)-glyceraldehyde

The conventional planar diagrams of the sugars are always drawn with the CHO (or $CH_2OH{\cdot}CO$) group at the top and the CH_2OH group at the bottom; the following short-hand notation is also used:

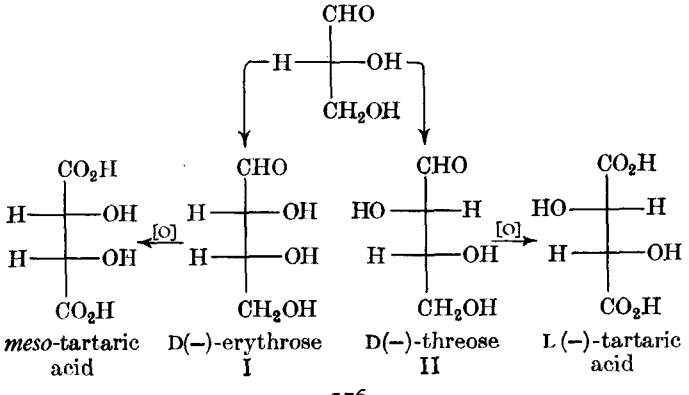


D-series L-series

Aldotetroses. The structural formula of the aldotetroses is

$$CH_2OH{\cdot}CHOH{\cdot}CHOH{\cdot}CHO.$$

Since this contains two unlike asymmetric carbon atoms, there are four



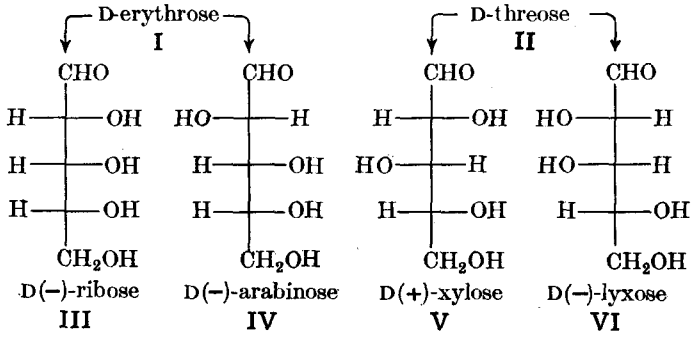
meso-tartaric acid D(−)-erythrose I D(−)-threose II L(−)-tartaric acid

optically active forms (two pairs of enantiomorphs) possible theoretically. All four are known, and correspond to D- and L-threose and D- and L-erythrose. D(+)-Glyceraldehyde may be stepped up by the Kiliani reaction to give D(−)-erythrose and D(−)-threose. The question now is: Which is which? On oxidation, D-erythrose gives *meso*tartaric, and on reduction gives *meso*erythritol. Therefore D-erythrose is I, and consequently II must be D-threose. The configuration of the latter is confirmed by the fact that on oxidation, D-threose gives L(−)-tartaric acid.

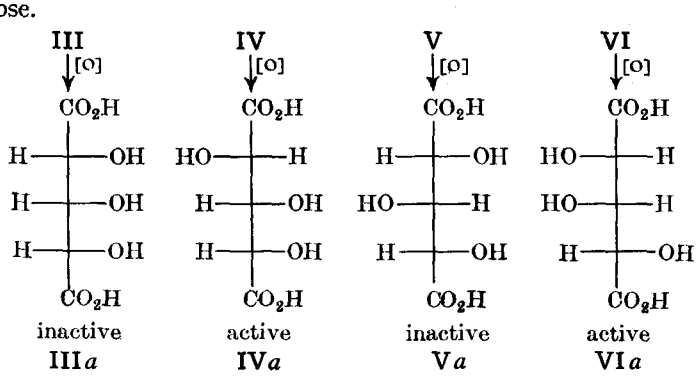
Aldopentoses. These have the structural formula

$$CHO{\cdot}CHOH{\cdot}CHOH{\cdot}CHOH{\cdot}CH_2OH,$$

and since it contains three unlike asymmetric carbon atoms, there are eight optically active forms (four pairs of enantiomorphs). All are known, and correspond to the D- and L-forms of ribose, arabinose, xylose and lyxose. Their configurations may be ascertained by either of the following two methods.



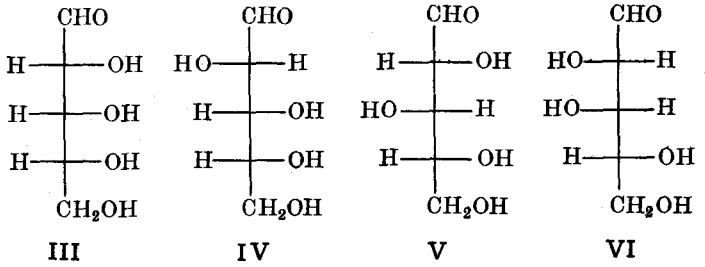
One method starts by stepping up the aldotetroses by the Kiliani reaction. Thus D-erythrose gives D(−)-ribose and D(−)-arabinose; similarly, D-threose gives D(+)-xylose and D(−)-lyxose. III and IV must be ribose and arabinose, but which is which? On oxidation with nitric acid, arabinose gives an optically active dicarboxylic acid (a trihydroxyglutaric acid), whereas ribose gives an optically inactive dicarboxylic acid. When the terminal groups, *i.e.*, CHO and CH_2OH, of III are oxidised to carboxyl groups, the molecule produced (III*a*) possesses a plane of symmetry, and so is inactive. Oxidation of IV gives IV*a*, and since this molecule has no plane (or any other element) of symmetry, it is optically active. Thus III is D-ribose and IV is D-arabinose.



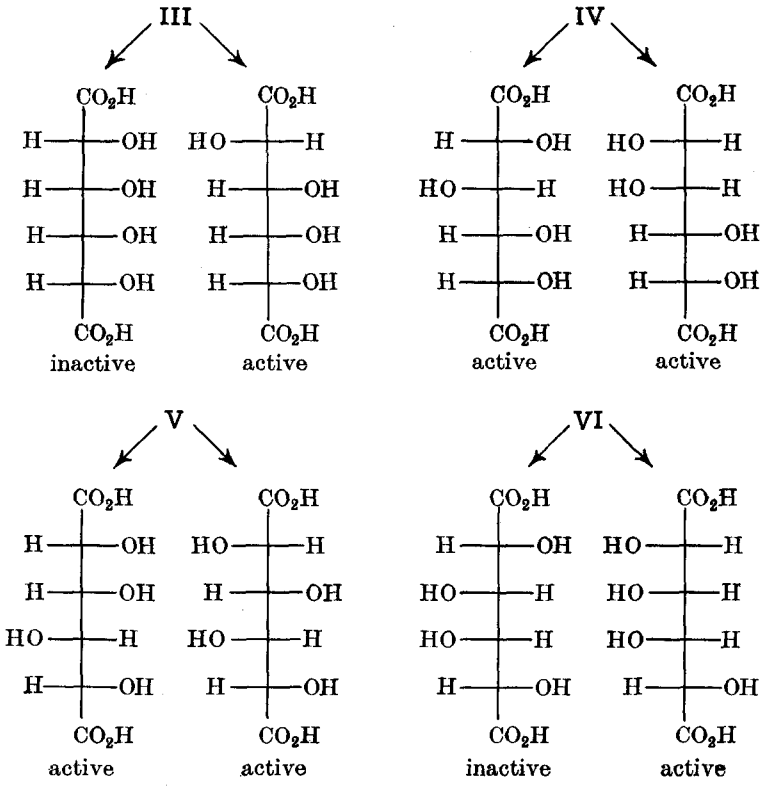
V and VI must be xylose and lyxose, but which is which? The former sugar, on oxidation, gives an optically inactive dicarboxylic acid, whereas

the latter gives an optically active dicarboxylic acid. Therefore V is D-xylose and VI is D-lyxose.

The following is the alternative method of elucidating the configurations of the aldopentoses; it is more in keeping with Fischer's solution to the problem. The structural formula of the aldopentoses can give rise to four pairs of enantiomorphs, the D-forms of which are as follows:



It should be noted that these four configurations have been obtained from first principles (see §7c. II); no recourse has been made to the configurations of the aldotetroses. Arabinose and lyxose, on oxidation with nitric acid, produce optically active dicarboxylic acids (trihydroxyglutaric acids). Therefore these two pentoses must be IV and VI, but we cannot say which is which. Xylose and ribose, on oxidation, produce optically inactive dicarboxylic acids (trihydroxyglutaric acids). Therefore these two pentoses must be III and V, and again we cannot say which is which. When each



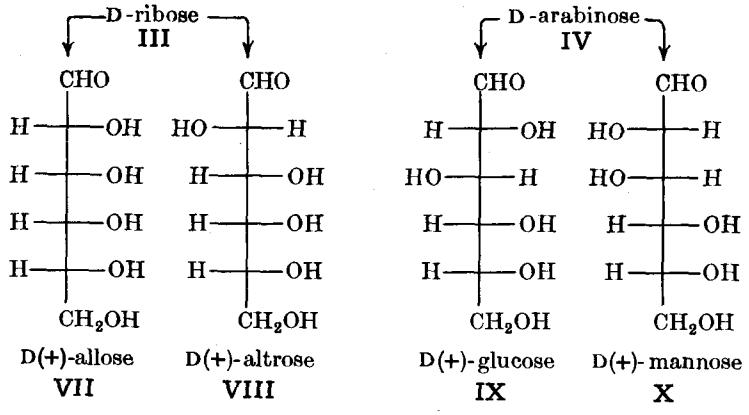
aldopentose is stepped up by one carbon atom (by means of the Kiliani reaction) and then oxidised to the dicarboxylic acid (the terminal groups are oxidised), it is found that arabinose and xylose each give *two* active dicarboxylic acids, whereas ribose and lyxose each give one active and one inactive (*meso*) dicarboxylic acid. The chart at foot of previous page shows the dicarboxylic acids obtained from the configurations III–VI.

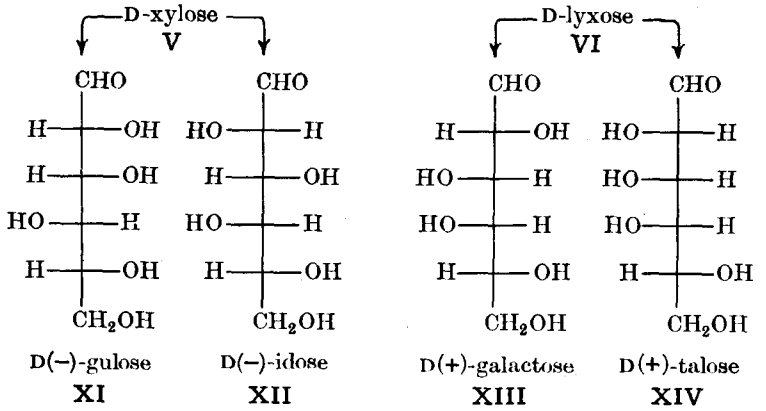
It therefore follows that D-ribose is III, D-arabinose is IV, D-xylose is V and D-lyxose is VI. These configurations are confirmed by the facts that ribose and arabinose give the same osazone, and xylose and lyxose give the same osazone; the only difference between sugars giving the same osazone is the configuration of the second carbon atom, *i.e.*, III and IV are epimers, as are V and VI. It should also be noted that arabinose and lyxose produce the same trihydroxyglutaric acid on oxidation.

Aldohexoses. The structural formula of these compounds is

$$CHO{\cdot}CHOH{\cdot}CHOH{\cdot}CHOH{\cdot}CHOH{\cdot}CH_2OH,$$

and since it contains four unlike asymmetric carbon atoms, there are sixteen optically active forms (eight pairs of enantiomorphs). All are known, and may be prepared by stepping up the aldopentoses: D-ribose gives D(+)-allose and D(+)-altrose; D-arabinose gives D(+)-glucose and D(+)-mannose; D-xylose gives D(−)-gulose and D(−)-idose; and D-lyxose gives D(+)-galactose and D(+)-talose.





VII and VIII must be allose and altrose, but which is which? On oxidation with nitric acid, the former gives an optically inactive (allomucic) and

the latter an optically active (talomucic) dicarboxylic acid. Therefore allose is VII and altrose is VIII.

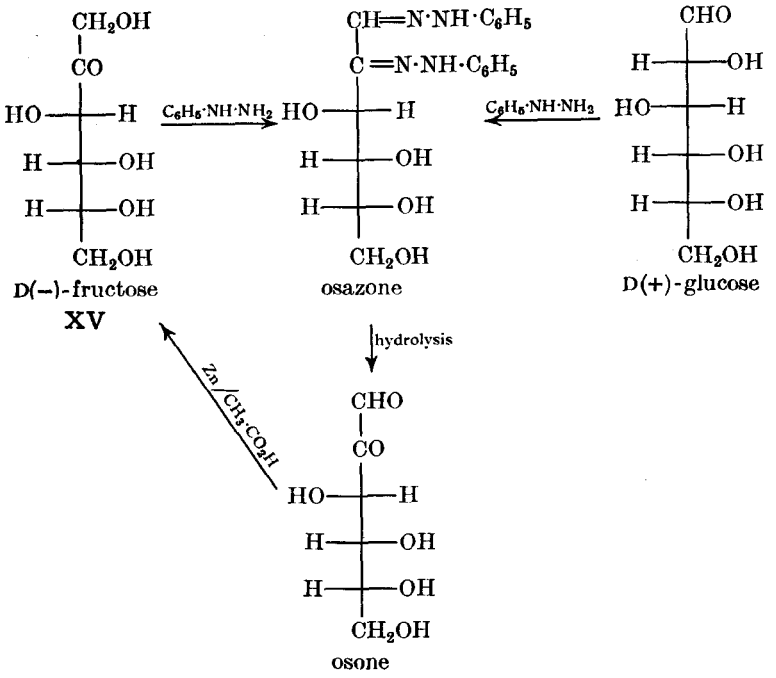
XIII and XIV must be galactose and talose, but which is which? On oxidation with nitric acid, the former gives an optically inactive (mucic) and the latter an optically active (talomucic) dicarboxylic acid. Therefore XIII is galactose and XIV is talose.

The elucidation of the configurations of the remaining four aldohexoses is not quite so simple, since, on oxidation with nitric acid, glucose and mannose *both* give optically active dicarboxylic acids, as also do gulose and idose; in all four configurations (IX, X, XI, XII), replacement of the two terminal groups (CHO and CH_2OH) by carboxyl groups leads to dicarboxylic acids whose structures have no plane (or any other element) of symmetry. It has been found, however, that the dicarboxylic acid from glucose (saccharic acid) is the same as that obtained from gulose (actually the two saccharic acids obtained are enantiomorphous, D-glucose giving D-saccharic acid and D-gulose L-saccharic acid). Since saccharic acid, $CO_2H\cdot(CHOH)_4\cdot CO_2H$, is produced by the oxidation of the terminal groups with the rest of the molecule unaffected, it therefore follows that the "rest of molecule" must be the same for both glucose and gulose. Inspection of formulæ IX, X, XI and XII shows that only IX and XI have the "rest of the molecule" the same; by interchanging the CHO and CH_2OH groups of IX, the enantiomorph of XI, *i.e.*, L-gulose, is obtained. Therefore IX must be glucose (since we know that glucose is obtained from arabinose), and XI must be gulose. Consequently X is mannose and XII is idose.

Ketohexoses. All the ketohexoses that occur naturally have the ketonic group adjacent to a terminal CH_2OH group, *i.e.*, the structural formula of all the natural ketohexoses is

$$CH_2OH\cdot CO\cdot CHOH\cdot CHOH\cdot CHOH\cdot CH_2OH.$$

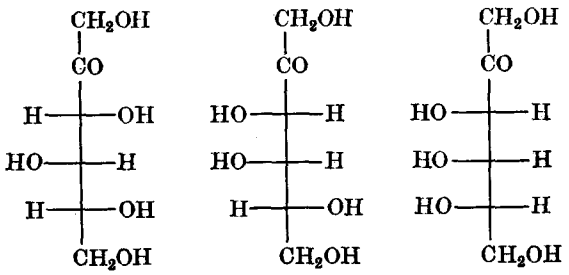
Since this structure contains three dissimilar asymmetric carbon atoms,



there are eight optically active forms (four pairs of enantiomorphs) possible theoretically; of these the following six are known: D(−)- and L(+)-fructose, D(+)- and L(−)-sorbose, D(+)-tagatose and L(−)-psicose. Only D(−)-fructose, L(−)-sorbose and D(+)-tagatose occur naturally.

Fructose. Natural fructose is lævorotatory, and since D-glucose gives the *same* osazone as natural fructose, the latter must be D(−)-fructose. Furthermore, since osazone formation involves only the first two carbon atoms in a sugar, it therefore follows that the configuration of the rest of the molecule in glucose and fructose must be the same. Hence the configuration of D(−)-fructose is XV, and is confirmed by the fact that D(+)-glucose may be converted into D(−)-fructose *via* the osazone (see chart at foot of previous page).

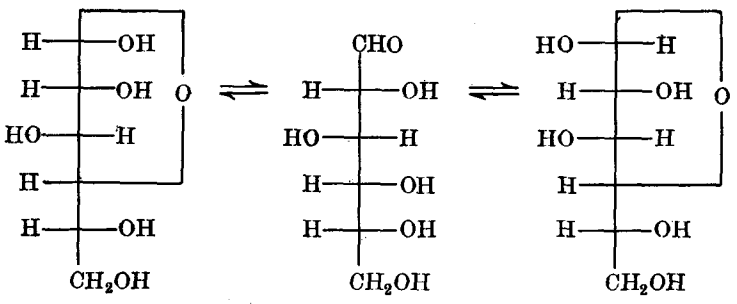
The configurations of the other ketohexoses are:



D(+)-sorbose D(+)-tagatose L(−)-psicose

§2. Ring structure of the monosaccharides. When a monosaccharide is dissolved in water, the optical rotatory power of the solution gradually changes until it reaches a constant value (Dubrunfaut, 1846); *e.g.*, a freshly prepared solution of glucose has a specific rotation of +111°, and when this solution is allowed to stand, the rotation falls to +52·5°, and remains constant at this value. The final stage can be reached more rapidly either by heating the solution or by adding some catalyst which may be an acid or a base. This change in specific rotation is known as **mutarotation**; all *reducing* sugars (except a few ketoses) undergo mutarotation.

To account for mutarotation, Tollens (1883) suggested an oxide ring structure for D(+)-glucose, whereby *two* forms would be produced, since, in the formation of the ring, another asymmetric carbon atom (which can exist in *two* configurations) is produced (*cf.* the Kiliani reaction). Tollens assumed that a five-membered ring (the γ-form) was produced:



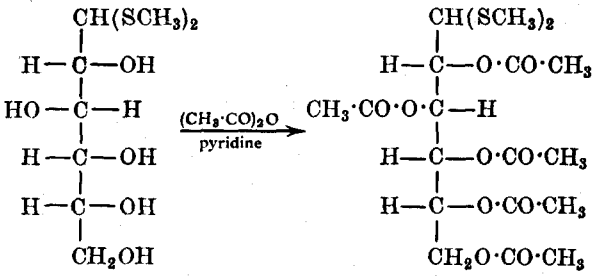
I D(+)-glucose II

The difficulty of this suggestion was that there was no experimental evidence for the existence of these two forms. Tanret (1895), however, isolated two

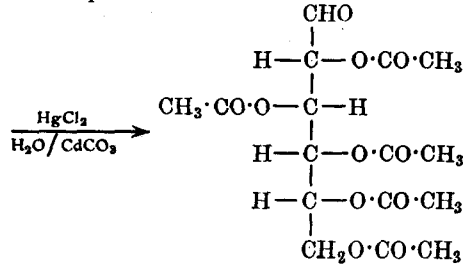
isomeric forms of D(+)-glucose, thus apparently verifying Tollens' supposition (but see §§7a, 7f). The two forms, I and II, are known respectively as α- and β-D(+)-γ-glucose (see also §7b for the nomenclature of these forms).

Ring formation of a sugar is really hemiacetal formation, one alcoholic group of the sugar forming a hemiacetal with the aldehyde group of the *same* molecule, thus producing a ring structure which is known as the *lactol* form of the sugar.

Mechanism of mutarotation. According to Lowry (1925), mutarotation is not possible without the presence of an amphiprotic solvent, *i.e.*, a solvent which can function both as an acid and a base, *e.g.*, water. Thus Lowry and Faulkner (1925) showed that mutarotation is arrested in pyridine solution (basic solvent) and in cresol solution (acidic solvent), but that it takes place in a mixture of pyridine and cresol. It has been assumed that when mutarotation takes place, the ring opens and then recloses in the inverted position or in the original position. There is some evidence for the existence of this open-chain form. The absorption spectra of fructose and sorbose in aqueous solution indicate the presence of open-chain forms; aldoses gave negative results (Bednarczyk *et al.*, 1938). Solutions of glucose and arabinose in 50 per cent. sulphuric acid gave an ultraviolet absorption spectrum containing the band characteristic of the oxo (carbonyl) group (Pascu *et al.*, 1948). Aldoses in solution contain a form which is reducible at the dropping mercury electrode (Cantor *et al.*, 1940). Although the nature of this reducible form has not been established, it is probably the open-chain form, either free or hydrated. Furthermore, a relationship was shown to exist between the amount of this reducible form and the rate of mutarotation. One interpretation of this observation is that the reducible form is an intermediate in mutarotation. Rate constants for the conversion of the ring forms of aldoses to the open-chain form have been calculated from polarographic measurements, and it has also been shown that the energy of activation required to open the pyranose ring is the same for glucose, mannose, galactose, arabinose and xylose (Delahay *et al.*, 1952). The formation of this acyclic intermediate during mutarotation has been confirmed by isotopic evidence (Goto *et al.*, 1941) and by further polarographic evidence

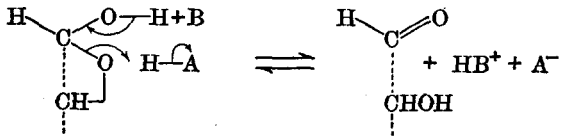


glucose dimethyl mercaptal



(Overend *et al.*, 1957). It is interesting to note in connection with this problem of the existence of the open-chain structure, that ***aldehydo-sugars***, *i.e.*, aldoses in which the aldehyde group is present, can only be *isolated* if all the hydroxyl groups in the *open-chain form* are "protected"; *e.g.*, Wolfrom (1929) prepared 2 : 3 : 4 : 5 : 6-penta-acetylaldehydoglucose as shown at foot of previous page.

The problem now is: What is the mechanism of the formation of the open-chain form from the ring-form? Lowry (1925) suggested that it occurred by the simultaneous addition and elimination of a proton, since both an acid and a base must be present (see above). This concerted mechanism would conform to a third-order reaction:



Swain *et al.* (1952) have shown that the mutarotation of tetramethylglucose, catalysed by phenol and pyridine in benzene solution, is a third-order reaction; this supports the above mechanism. On the other hand, some authors believe that the reaction proceeds in two independent ways, one being an acid-catalysed reaction, and the other a base-catalysed reaction. In this case the mechanism would conform to a second-order reaction. Hill *et al.* (1952) have shown that the mutarotation of glucose in aqueous methanol containing acetate buffers is in better agreement with a second-order reaction than with a third-order.

It can thus be seen that the mechanism of mutarotation cannot be regarded as settled, and it appears likely that the sugar investigated (free or as a derivative) and the experimental conditions may play a part in deciding which mechanism will operate (see §7h).

Preparation of the α- and β-forms of a sugar. Experimentally, it is very difficult to isolate the α- and β-forms of a sugar. The ordinary form of D(+)-glucose is the α-isomer, m.p. 146° and $[\alpha]_D = +111°$; this form may be prepared by crystallising glucose from cold ethanol. The β-isomer, m.p. 148–150°, $[\alpha]_D = +19{\cdot}2°$, can be obtained by crystallising glucose from hot pyridine. Thus the α-form may be converted into the β-, and *vice versa*, during the process of crystallisation; this is an example of asymmetric transformation (§10 iv. II). Both forms show mutarotation, the final value of the specific rotation being +52·5°; this corresponds to a mixture containing about 38 per cent. of the α-isomer, and 62 per cent. of the β-. The two stereoisomeric ring-forms of a sugar are often referred to as *anomers*.

Summary of the evidence for the ring structures of sugars. The cyclic structure of the sugars accounts for the following facts:

(i) The existence of two isomeric forms (anomers) of a given sugar, *e.g.*, α- and β-glucose.

(ii) Mutarotation.

(iii) Glucose and other aldoses do not give certain characteristic reactions of aldehydes, *e.g.*, Schiff's reaction, do not form a bisulphite or an aldehyde-ammonia compound. Recently, however, it has been shown that by preparing Schiff's reagent in a special way, it becomes very sensitive, simple aldoses restoring the pink colour to this solution; the monosaccharide aldoses react strongly, but the disaccharide aldoses react weakly (Tobie, 1942). This reaction with a sensitive Schiff's reagent appears to indicate that some, although a very small amount, of the open-chain form of a sugar is present in solution in equilibrium with the two ring-forms.

(iv) Glucose penta-acetate does not react with hydroxylamine; this

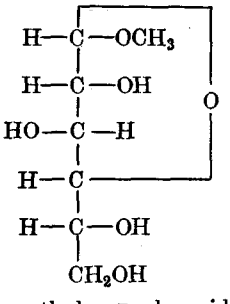
indicates that the aldehyde group is absent in this derivative (glucose itself does form an oxime).

(v) Aldehydes normally form acetals by combination with two molecules of a monohydric alcohol; aldoses (and ketoses) combine with only one molecule of an alcohol. It should be noted, however, that aldoses will combine with *two* molecules of a thiol to form a mercaptal (thioacetal).

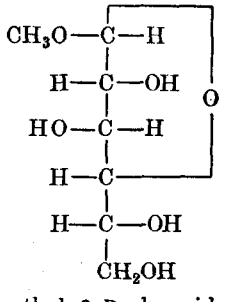
(vi) X-ray analysis definitely proves the existence of the ring structure, and at the same time indicates the size of the ring (see §7f).

§3. **Glycosides.** Just as simple hemi-acetals react with another molecule of an alcohol to form acetals, so can the sugars, in their ring-forms (lactols), react with a molecule of an alcohol to form the acetal derivative, which is known under the generic name of **glycoside;** those of glucose are known as *glucosides*; of fructose, *fructosides*, etc. The hydroxyl group produced at the oxo group by ring formation is known as the *glycosidic hydroxyl group*. This group can be acetylated and methylated, as can all the other hydroxyl groups in the sugar, but the glycoside derivatives are far more readily decomposed by various reagents.

E. Fischer (1893) refluxed glucose in methanol solution in the presence of one-half per cent. hydrochloric acid, and thereby obtained a white crystalline product which contained *one* methyl group (as shown by analysis), and which did not reduce Fehling's solution or mutarotate, and did not form an osazone. Thus the hemiacetal structure is no longer present in this compound; in fact, this compound appears to be an acetal since it is stable in alkaline solution (Fehling's solution). Furthermore, on boiling with dilute inorganic acids, the compound regenerated the original sugar, a reaction again typical of acetals. Ekenstein (1894) isolated a second isomer from the reaction mixture when he repeated Fischer's work, and Fischer explained the existence of these two isomers by suggesting ring structures for the two methyl glucosides, *viz.*,



methyl α-D-glucoside



methyl β-D-glucoside

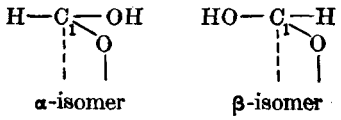
Fischer assumed that these methyl glucosides were five-membered ring systems, basing his assumption on Tollens' suggestion (§2). As we shall see later (§7a), Fischer's assumption is incorrect.

The non-sugar part of a glycoside is known as the *aglycon* (or *aglycone*), and in many glycosides that occur naturally, the aglycon is often a phenolic compound (see §24).

Fischer (1894) found that methyl α-D-glucoside was hydrolysed by the enzyme maltase, and the β-D-glucoside by the enzyme emulsin. Furthermore, Fischer also found that maltase would not hydrolyse the β-glucoside, and that emulsion would not hydrolyse the α-glucoside. Thus the two isomers can be distinguished by the specificity of action of certain enzymes (see also §16, XIII). Armstrong (1903) followed these enyzmic hydrolyses polarimetrically, and showed that methyl α-D-glucoside liberates α-D-glucose,

and that the β-glucoside liberates β-D-glucose; Armstrong found that hydrolysis of the α-glucoside produced a " downward " mutarotation, whereas that of the β-glucoside produced an " upward " mutarotation. It therefore follows that α-D-glucose is stereochemically related to methyl α-D-glucoside, and β-D-glucose to methyl β-D-glucoside.

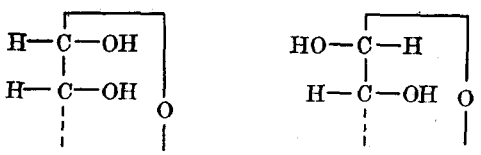
§4. Configuration of C_1 in glucose. The configurations of C_1 in α- and β-D-glucose have been written, in the foregoing account, as:



I II

The question that now confronts us is: What justification is there for this choice, *i.e.*, what is the evidence that enables us to say that the α-isomer (characterised by certain physical constants) actually has the hydrogen atom to the left and the hydroxyl group to the right? Hudson (1909) proposed the *empirical* rule that of an α, β pair of sugars in the D-series, the α-isomer, which has the *higher* dextrorotation (*i.e.*, this physical constant decides which of the two is to be designated α-), has the hydrogen to the left (*i.e.*, I); the β-isomer consequently has the hydrogen atom to the right (II). Thus α-D(+)-glucose is the isomer with the specific rotation +111°, and β-D(+)-glucose is the isomer with the specific rotation +19·2°. If the D-sugar has a negative rotation, then, according to the empirical rule, the β-isomer has the higher negative rotation (*i.e.*, the less positive rotation), *e.g.*, α-D(−)-fructose is the isomer with the specific rotation −20°, and the β-isomer −133°. In the L-sugars, the α-isomer is the one with the *higher* lævorotation, and the other is the β-isomer; thus the α-forms (and the β-forms) of the D- and L-series are enantiomorphous.

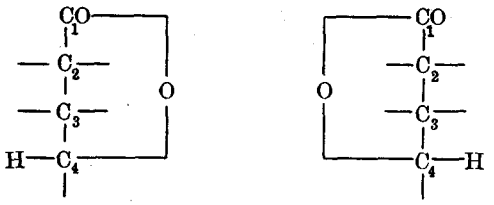
Böeseken (1913) found that when boric acid is added to a solution of a cyclic 1 : 2-glycol, the electrical conductivity of the solution is greater than that of boric acid itself, and that the increase is greater for the *cis*-isomer than for the *trans*- (see Vol. I). This phenomenon has been used to distinguish between the two anomers of D-glucose; the results obtained showed that the conductivity of the isomer called the α (from the above empirical rule), in the presence of boric acid, decreased during mutarotation, whereas the conductivity of the β-isomer increased. This suggests that the α-isomer has configuration III, and the β-isomer IV. Thus we now have physicochemical evidence that the 1 : 2-hydroxyl groups are in the *cis*-position in



III IV

the α-isomer, *i.e.*, there is now some experimental evidence in support of Hudson's empirical rule. These configurations have been confirmed by further work, *e.g.*, Rüber (1931) found that, in general, *trans*-compounds have a higher molecular refraction than the corresponding *cis*-; the molecular refraction of β-D-glucose is greater than that of the α-isomer, and so agrees with the results obtained by the conductivity experiments. The strongest bit of evidence for the configurations of the α- and β-isomers has been obtained from X-ray studies of α-D-glucose (see §7f).

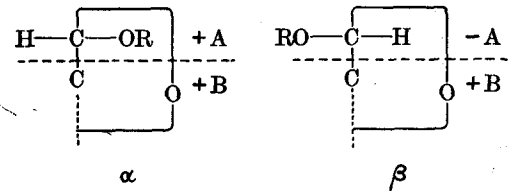
§5. Hudson's lactone rule. Hudson (1910) studied the rotation of the lactones derived from the aldonic acids. Using the usual projection formulæ, the lactone ring will be on the right or left according as the hydroxyl group on C_4 (*i.e.*, the γ-hydroxyl group) is on the right or left, *i.e.*, according as C_4 has a *dextro* or *lævo* configuration:



dextrorotatory lævorotatory

From an examination of 24 lactones derived from aldonic acids, and assuming that they were γ-lactones, Hudson concluded that if the lactone ring was on the right, the compound was dextrorotatory; if the ring was on the left, then lævorotatory.

§6. Hudson's isorotation rules. Hudson (1909, 1930) applied the rule of optical superposition (§12. I) to carbohydrate chemistry, and his first application was to the problem of the configuration of C_1 in the anomers of aldoses. Hudson pointed out that the only structural difference between the α- and β-anomers (of sugars and glycosides) is the configuration of C_1. Thus, representing the rotation of this terminal group as A and that of the rest of the molecule as B, and then taking the α-anomer as the one with the higher positive rotation (in the D-series) we have:



α β

Molecular rotation of the α-anomer $= +A + B$
" " " " β- " $= -A + B$

Thus in every pair of α- and β-anomers the following rules will hold:

Rule 1. The sum of the molecular rotations (2B) will be a constant value characteristic of a particular sugar and independent of the nature of R.

Rule 2. The difference of the molecular rotations (2A) will be a constant value characteristic of R.

As we have seen, the rule of optical superposition does not hold exactly (due to neighbouring action, etc.; see §12. I). In the sugars, however, the rotation of C_1 is affected only to a small extent by changes in the rest of the molecule, and *vice versa*. This is illustrated in the following table, from which it can be seen that the sum of the molecular rotations (2B) for various pairs of glucopyranoside anomers is fairly constant.

C_1 substituent	M_α	M_β	$M_\alpha + M_\beta = 2B$
OH	+202	+34	+236
OCH_3	+309	−66	+243
OC_2H_5 . . .	+314	−69·5	+245·5

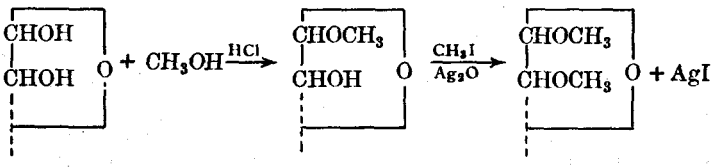
These isorotation rules have been used to ascertain which of an anomeric pair of glycosides is α and which is β, and to determine the type of glycosidic link in disaccharides and polysaccharides.

Lemieux *et al.* (1958), by means of proton magnetic resonance studies, have shown that the configurations assigned to the α- and β-anomers of sugar acetates on the basis of Hudson's rules are correct.

§7. Methods for determining the size of sugar rings. As pointed out previously, Fischer followed Tollens in proposing the γ-oxide ring. There was, however, no experimental evidence for this; the γ-hydroxyl group was chosen as being involved in ring formation by analogy with the ready formation of γ-lactones from γ-hydroxyacids. The problem was further complicated by the fact that Hudson *et al.* (1915) isolated four galactose penta-acetates, none of which had a free aldehyde group. Furthermore, these four compounds were related to each other as pairs, *i.e.*, there were two α- and two β-isomers. The only reasonable explanation for this was that there are *two* ring systems present, but once again there is no evidence to decide the actual sizes of the rings.

The original experimental approach to the problem of determining the size of the ring present in sugars consisted essentially in studying the methylated sugars. A more recent method uses the methyl glycosides (for this method, see §7g). Since methylation is so important in the original method, the following account describes briefly the methods used.

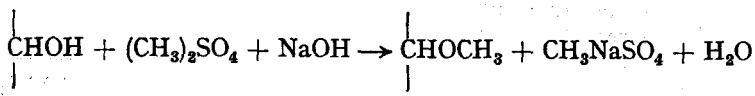
(i) *Purdie's method* (1903). The sugar is first converted into the corresponding methyl glycoside (methanol and hydrochloric acid), and this is then heated with methyl iodide in the presence of *dry* silver oxide; thus:



Purdie's method is only applicable to glycosides and other derivatives in which the *reducing group* is missing or has been protected by substitution. Methylation of a free reducing sugar by this method would result in the oxidation of that sugar by the silver oxide.

In certain cases, thallous hydroxide may be used instead of silver oxide (Fear *et al.*, 1926).

(ii) *Haworth's method* (1915). In this method methyl sulphate and aqueous sodium hydroxide are added to a well-stirred sugar solution at such a rate that the liquid remains practically neutral:



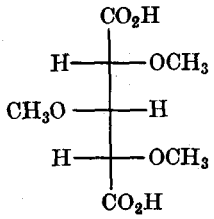
This method is directly applicable to all reducing sugars.

(iii) More recent methods of methylation use sodium and methyl iodide in liquid ammonia, or diazomethane in the presence of moisture.

Having obtained the fully methylated methyl glycoside, the latter is then hydrolysed with dilute hydrochloric acid, whereby the glycosidic methyl group is eliminated. A study of the oxidation products of the methylated sugar then leads to the size of the ring. It should be noted that throughout the whole method, the assumption is made that no methyl groups migrate or that any change in the position of the oxide ring occurs (see, however, later). The number of methyl groups present in the methylated sugar and

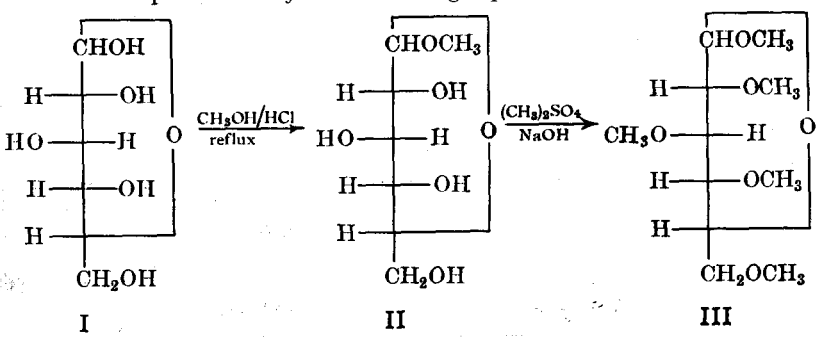
the various oxidation products are determined by the Zeisel method (see Vol. I). Also, these methyl derivatives are often purified by distillation *in vacuo*. Bishop *et al.* (1960) have now separated methylated methyl glycosides by gas chromatography.

§7a. Pyranose structure. This structure is also sometimes referred to as the δ-oxide or amylene oxide ring. As an example of the method used, we shall consider the case of D(+)-glucose (Haworth and Hirst, 1927). D(+)-Glucose, I, was refluxed in methanol solution in the presence of a small amount of hydrochloric acid, and the methyl D-glucoside, II, so produced was methylated with methyl sulphate in the presence of sodium hydroxide to give methyl tetramethyl-D-glucoside, III, and this, on hydrolysis with dilute hydrochloric acid, gave tetramethyl-D-glucose, IV. When this was dissolved in water and then oxidised by heating with excess of bromine at 90°, a lactone, V, was isolated, and this, on further oxidation with nitric acid, gave xylotrimethoxyglutaric acid, VI. The structure of this compound is known, since it can be obtained directly by the oxidation of methylated xylose; thus its structure is VI (see also §7d). The structure of this

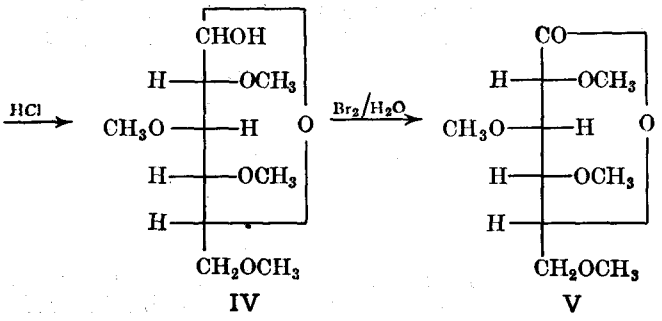


VI

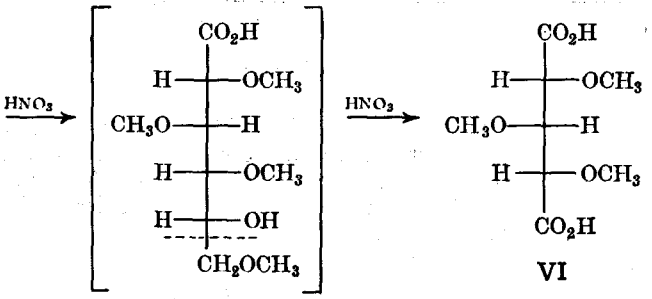
compound is the key to the determination of the size of the ring in the sugar. One of the carboxyl groups in VI must be that which is combined in the formation of the lactone ring in the tetramethylgluconolactone, V. The other carboxyl group is almost certainly the one that has been derived from the non-methylated carbon atom, *i.e.*, from the CHOH group that is involved in the ring formation in the sugar. Therefore there must be *three* methoxyl groups in the lactone ring. Thus the lactone *cannot* be a γ-lactone, and consequently C_5 must be involved in the ring formation. It therefore follows that the lactone, V, must be 2 : 3 : 4 : 6-tetra-*O*-methyl-D-gluconolactone. Working *backwards* from this compound, then IV must be 2 : 3 : 4 : 6-tetra-*O*-methyl-D-glucose, III methyl 2 : 3 : 4 : 6-tetra-*O*-methyl-D-glucoside, II methyl D-glucopyranoside, and I D-glucopyranose (see §7f for the significance of the term pyranose). It should be noted that the question as to whether the sugar is α or β has been ignored; starting with either leads to the same final results. The foregoing experimental results can now be represented by the following equations:



I II III

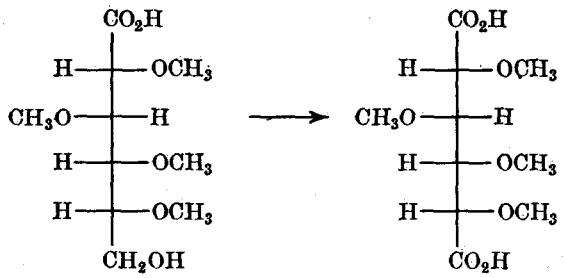


IV V



VI

There is a slight possibility that the ring might have been an ε-ring, *i.e.* the oxide ring involves C_1 and C_6, and that C_5 is converted to the carboxy group with loss of C_6. Haworth, however, made certain that this was not the case by the following method. Had the ring been 1 : 6-, then 2 : 3 : 4 : 5-tetramethylgluconic acid, VII, would have been obtained (instead of V). VII was obtained by Haworth *et al.* (1927) from melibiose and gentiobiose (see §§18, 19) and, on oxidation, gave tetramethylsaccharic acid, VIII, and not the dicarboxylic acid, VI.

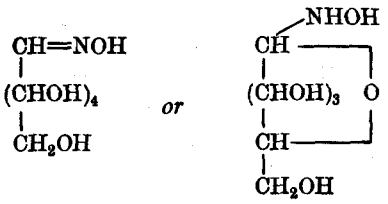


VII VIII

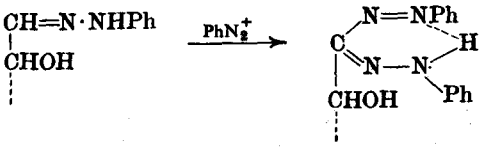
Thus there is a 1 : 5-ring in the tetramethylgluconolactone, tetra-*O*-methylglucose, methyl tetra-*O*-methylglucoside, methyl glucoside, *and therefore in glucose itself.* This conclusion is based on the assumption that no change in the ring position occurs during the methylation of glucose. Thus glucose is a δ- or pyranose sugar.

By similar methods it has been shown that hexoses and pentoses all possess a pyranose structure. There is also a large amount of evidence to

show that the oximes, phenylhydrazones and osazones of hexoses and pentoses may be cyclic or open-chain, *e.g.*, the oxime of glucose:



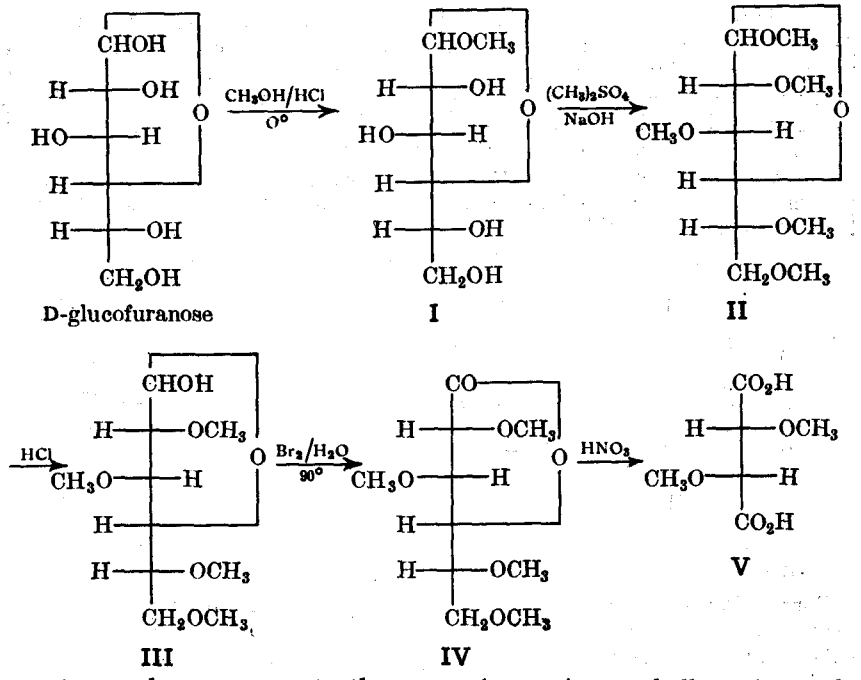
Mester *et al.* (1951–1955) showed that aldose phenylhydrazones react in pyridine solution with solutions of diazonium salts to give brilliant-red sugar diphenylformazans:



Formazan formation proves the acyclic structure of the sugar phenylhydrazones. The cyclic structures do not react, *e.g.*, there are three modifications of D-glucose phenylhydrazone (α, m.p. 159–160°; β, m.p. 140–141°; γ, m.p. 115–116°); two of these do not form formazans, but the third does. Hence the former two are cyclic and the third is acyclic.

§7b. Furanose structure. This structure is also sometimes referred to as the γ-oxide or butylene oxide ring. Fischer (1914) prepared methyl D(+)-glucoside by a slightly modified method, *viz.*, by dissolving D(+)-glucose in methanol, adding one per cent. hydrochloric acid, and then allowing the mixture to stand at 0° (instead of refluxing, as in his first procedure). On working up the product, he obtained a syrup (a crystalline compound was obtained by the first procedure). Fischer called this compound methyl γ-glucoside, and believed it was another isomer of the α- and β-forms; this is the significance of the symbol γ as used by Fischer. This syrup, however, was subsequently shown to be a mixture of methyl α- and β-glucofuranosides, *i.e.*, this glucoside contained a γ- or 1 : 4-ring (Haworth *et al.*, 1927). This syrup, I, when completely methylated (methyl sulphate method), gave a methyl tetra-*O*-methyl-D-glucoside, II, and this, on hydrolysis with dilute hydrochloric acid, gave tetra-*O*-methyl-D-glucose, III. On oxidation with bromine water at 90°, III gave a crystalline lactone, IV, and this, when oxidised with nitric acid, gave dimethyl-D-tartaric (dimethoxysuccinic) acid, V. This compound (V) is the only compound of known structure, and is therefore the key to the determination of the size of the ring in the sugar. Working *backwards* from V, then IV is 2 : 3 : 5 : 6-tetra-*O*-methyl-D-gluconolactone, III is 2 : 3 : 5 : 6-tetra-*O*-methyl-D-glucose, II is methyl 2 : 3 : 5 : 6-tetra-*O*-methyl-D-glucoside, and I is methyl D-glucofuranoside. If we write D-glucose as D-glucofuranose, then the foregoing reactions may be formulated as shown on next page (see §7f for the meaning of furanose).

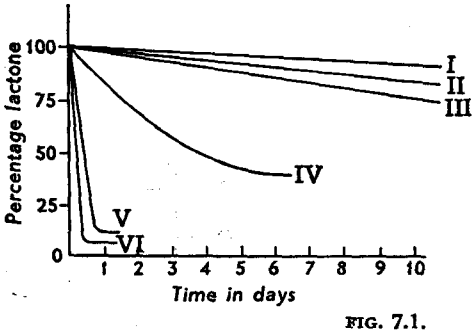
These reactions prove that I, II, III and IV all contain a γ-oxide ring, *i.e.*, the methyl glucoside, I, *prepared at* 0°, has a 1 : 4-ring. This then raises the question: What is the size of the ring in glucose itself? Is it 1 : 4 or 1 : 5? Preparation of the methyl glucoside at reflux temperature gives the 1 : 5-compounds (see §7a); preparation at 0° gives the 1 : 4-compounds. It is therefore not possible to say from these experiments whether glucose itself exists in the pyranose (1 : 5-) or furanose (1 : 4-) forms originally, or whether these two forms are in equilibrium. Further information is neces-



sary to supply an answer to these questions. As we shall see later, the normal form of a sugar is the pyranose structure (see §7f); pyranosides are often referred to as the " normal " glycosides.

By similar methods it has been shown that hexoses and pentoses give methyl glycosides possessing a furanose structure when prepared at 0° (or at room temperature).

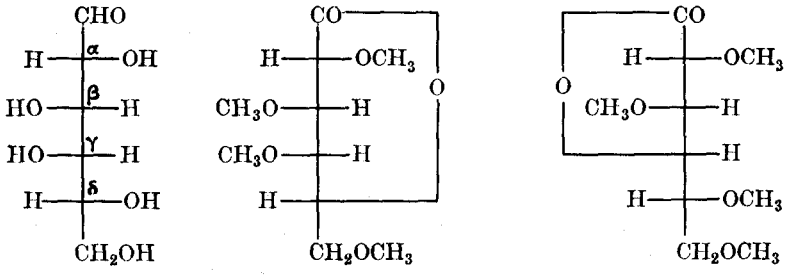
§7c. Determination of ring size by means of lactone formation. As we have seen, glycoside formation at *reflux* temperature leads ultimately to a methylated δ-lactone, whereas at 0° a methylated γ-lactone is obtained. Haworth (1927) examined the rates of hydration of these two types of lactones to the open-chain acids; the rates were measured by changes in the rotation or conductivity. Haworth found that the rate of hydration was much faster in one series than in the other; the δ-lactones were converted almost completely to the acids, whereas the γ-lactones were converted at a much slower rate (see Fig. 1). Thus, by comparing the stabilities (to hydration) of the various methylated lactones, it is possible to say whether the lactone under investigation is γ- or δ-. It is very important to note



I γ-mannolactone
II γ-galactonolactone
III γ-gluconolactone
IV δ-mannonolactone
V δ-gluconolactone
VI δ-galactonolactone

FIG. 7.1.

that this method easily distinguishes a γ- from a δ-lactone, but it does not *prove* one to be γ- and the other δ-. The actual nature of the lactone was proved *chemically*; the fast-changing lactone was shown to be the δ-lactone, and the slow-changing one the γ- (the chemical evidence was obtained by the degradative oxidation already described). However, having once established the relationship between the rate of hydration and the nature of the lactone, *e.g.*, in the case of glucose, mannose, galactose and arabinose, the property can *then* be used to determine the size of the ring in an unknown lactone of a sugar acid.



D-galactose (open-chain) | (+)-lactone; δ-lactone | (−)-lactone; γ-lactone

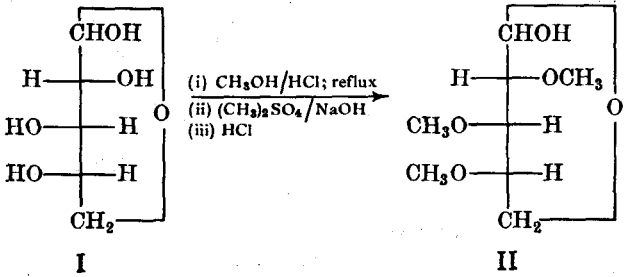
Correlation between the above scheme and Hudson's lactone rule has been demonstrated in certain cases, *e.g.*, galactose. Preparation of the methyl galactoside at reflux temperature, then methylation, hydrolysis, and finally oxidation with bromine water, leads to the formation of a methylated lactone which is dextrorotatory, and since it is a rapidly hydrated lactone, it must be δ-. Preparation of the methyl galactoside at 0°, etc., leads to the formation of a methylated lactone which is lævorotatory and is very stable to hydration. Thus, this lactone will have the ring to the left (Hudson's lactone rule), and hence must be a γ-lactone; at the same time, since it is a slowly hydrated lactone, it must be γ- (see the above formulæ).

§7d. Pyranose and furanose structures of pentoses. The methods used for determining the size of sugar rings have been described with glucose (an aldohexose) as the example. It is also instructive to apply these methods to the aldopentoses. L(+)-Arabinose has been chosen as the example, and the following equations and footnotes should now be readily followed:

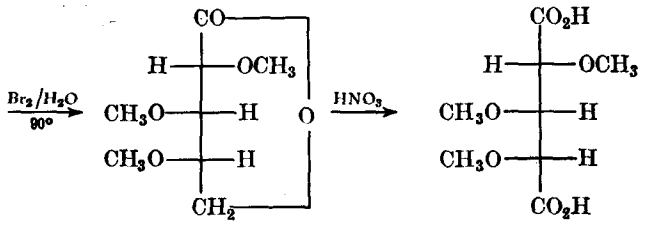
(i) *Glycoside formation at reflux temperature* (Haworth *et al.*, 1927).

I is L(+)-arabinopyranose, and since it is *dextrorotatory*, the ring has been drawn to the *right*. This way of drawing the projection formula is based on the observation of Haworth and Drew (1926), who pointed out that if a ring in a sugar is 1 : 5- (*i.e.*, δ-), then Hudson's lactone rule holds good for sugars as for γ-lactones.

II is 2 : 3 : 4-tri-*O*-methyl-L-arabinose.



I II



III IV

III is 2 : 3 : 4-tri-*O*-methyl-L-arabinolactone; it is a δ-lactone as shown by oxidation to IV, and also by the fact that it is of the type that is readily hydrated.

IV is 2 : 3 : 4-L-arabinotrimethoxyglutaric acid (this is the key compound).

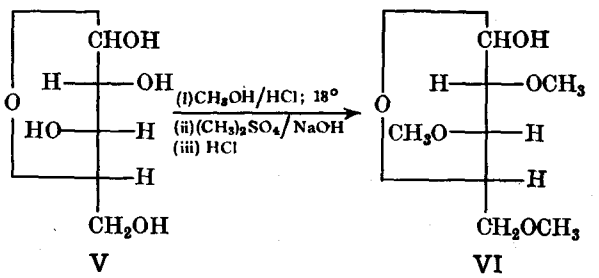
(ii) *Glycoside formation at room temperature* (Haworth *et al.*, 1925, 1927).

V is L-arabinofuranose.

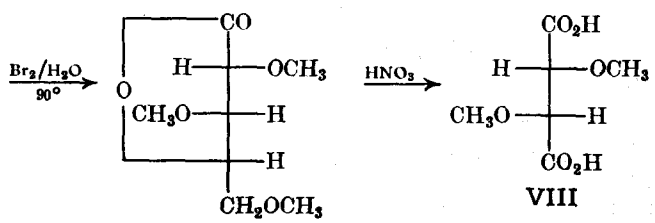
VI is 2 : 3 : 5-tri-*O*-methyl-L-arabinose.

VII is 2 : 3 : 5-tri-*O*-methyl-L-arabinolactone (Hudson's lactone rule, and is slow-changing type).

VIII is dimethyl-D-tartaric acid (this is the key compound).

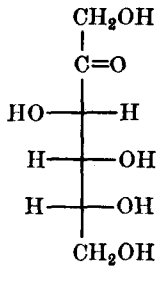


V VI

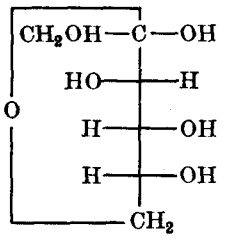


VII VIII

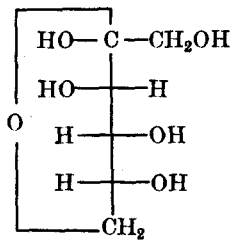
§7e. Ketose ring structures. Only D-fructose will be considered; the method is essentially the same as that for the aldoses, but there is one important variation, and that is in the oxidation of the tetramethylfructose. This cannot be oxidised by bromine water as can the tetramethylaldose; the fructose derivative is first oxidised with dilute nitric acid and then with acid permanganate, and by this means the lactone is obtained. The lactone is then further oxidised by moderately concentrated nitric acid. The following equations and footnotes explain the method, but before giving these, let us first consider the way of writing the projection formula of the ring structure of fructose. The usual open-chain formula is I, and to form the ring the ketone group is involved with C_6 in the pyranose form, and with C_5 in the furanose form; each of these can exist as the α- and β-isomers. When



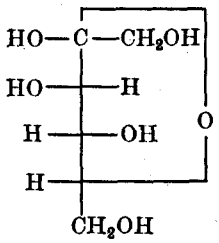
I



II
α-form



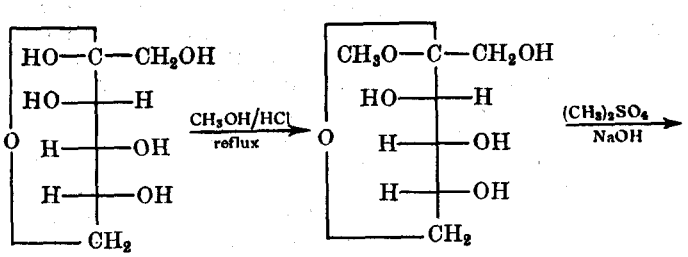
III
β-form



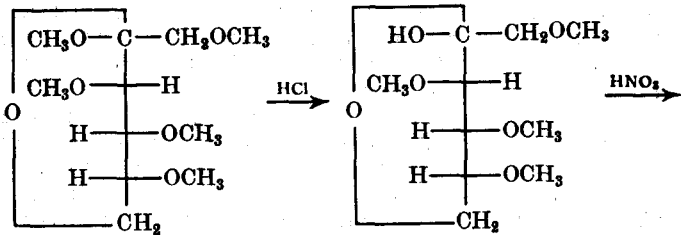
IV
β-form

the ring is closed, then if the hydroxyl group is drawn on the right, this will be the α-isomer (the CH_2OH group now replaces a hydrogen atom in the aldoses). Furthermore, since D-fructopyranose is lævorotatory, the oxide ring is drawn to the left (see the comments on L(+)-arabinopyranose, §7d). Thus α-D(−)-fructopyranose is II, and β-D(−)-fructopyranose is III. The furanose forms are obtained in a similar manner, but in this case the ring must be written to the right since the hydroxyl group on C_5 is on the right; thus β-D-fructofuranose is IV (see also sucrose, §13).

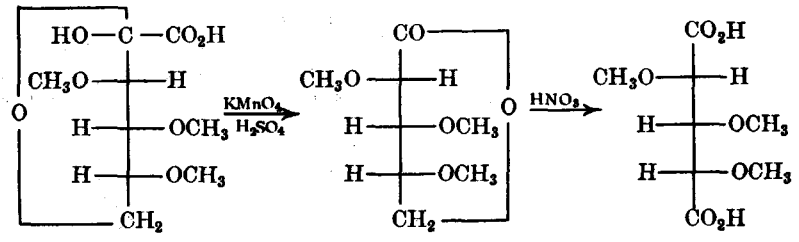
(i) *Glycoside formation at reflux temperature* (Haworth *et al.*, 1926, 1927).
V is β-D(−)-fructopyranose.
VI is methyl β-D-fructopyranoside.
VII is methyl 1 : 3 : 4 : 5-tetra-*O*-methyl-β-D-fructoside.
VIII is 1 : 3 : 4 : 5-tetra-*O*-methyl-β-D-fructose.
IX is 3 : 4 : 5-tri-*O*-methyl-β-D-fructuronic acid (as lactol).
X is 2 : 3 : 4-tri-*O*-methyl-D-arabinolactone; this is a quick-changing lactone, and is therefore a δ-lactone.
XI is D-arabinotrimethoxyglutaric acid.



V VI

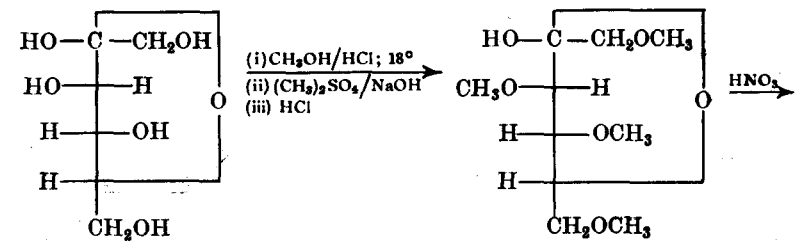


VII VIII

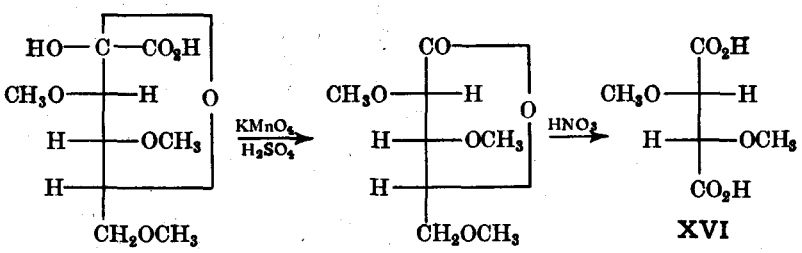


IX X XI

(ii) *Glycoside formation at room temperature* (Haworth *et al.*, 1927).
XII is β-D-fructofuranose.
XIII is 1 : 3 : 4 : 6-tetra-*O*-methyl-β-D-fructose.
XIV is 3 : 4 : 6-tri-*O*-methyl-β-D-fructuronic acid (as lactol).



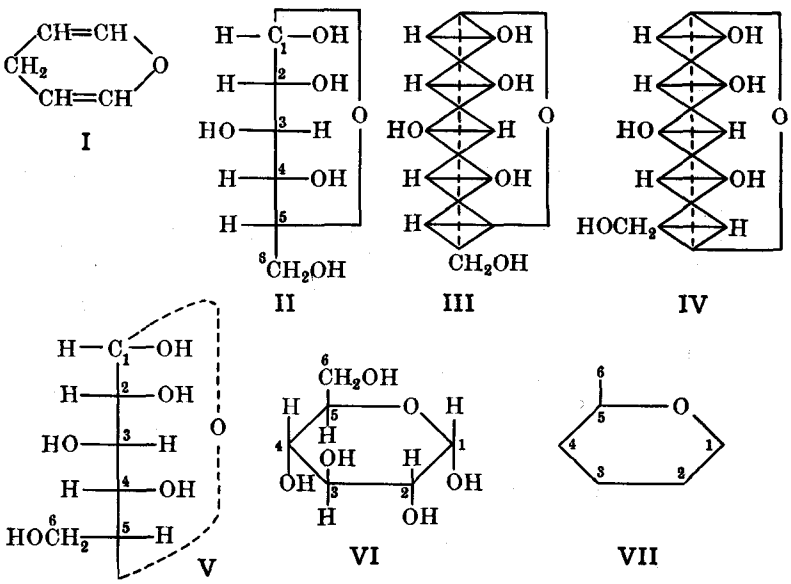
XII XIII



XIV XV XVI

XV is 2 : 3 : 5-tri-*O*-methyl-D-arabinolactone; this is a slow-changing lactone, and so is γ-.
XVI is dimethyl-L-tartaric acid.

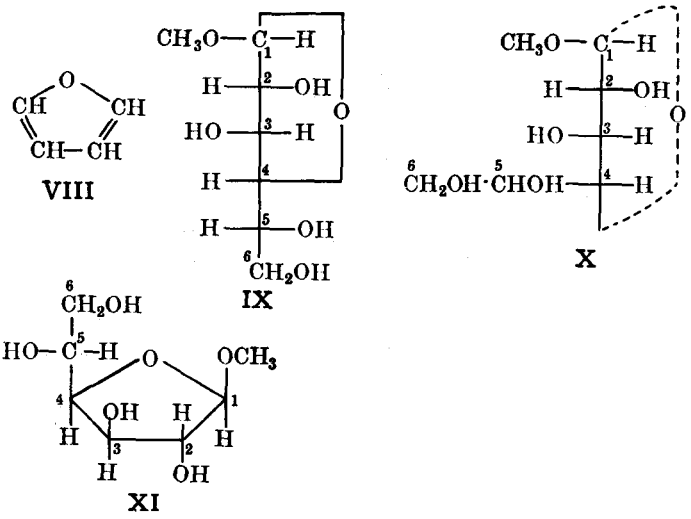
§7f. Conclusion. From the foregoing account it can be seen that the sugars exist as ring structures and not as open chains. Haworth (1926) therefore proposed a hexagonal formula for δ-sugars based on the pyran ring, I. The problem now is to convert the conventional plane-diagrams that we have been using into the *pyranose* formula. Let us take α-D-glucopyranose, II, as our example. The conventional tetrahedral diagram of II is III (see §5. II). Examination of III shows that the point of attachment of the oxide ring at C_1 is *below* the plane of the paper, and that at C_5 it is *above* the plane of the paper. If the tetrahedron with C_5 at its centre is rotated so that the point of attachment of the oxide ring is placed *below* the plane of the paper, III will now become IV, and the oxide ring will now be *perpendicular* to the plane of the paper, *i.e.*, perpendicular to the plane containing all the other groups (these all lie in a plane above the plane of the paper). The conventional plane-diagram of IV is V, but in order to emphasise the fact that the oxide ring is actually perpendicular to the plane of the paper, the part of the ring lying below the plane of the paper is shown by a broken line (the true plane-diagram should have a normal line drawn



as in II). Comparison of V with II shows that where the CH_2OH was originally is now the point of attachment of the oxide ring, the CH_2OH occupying the position where the H atom was, and the latter now where the oxide ring was. Thus, if we consider the conversion of II into V without first drawing III and IV, then in effect *two* Walden inversions have been effected, and consequently the original configuration is retained. V is now transformed into the perspective formula VI by twisting V so that the oxide ring is perpendicular to the plane of the paper and all the other groups are joined to bonds which are parallel to the plane of the paper. By convention, C_1 is placed to the right and the oxygen atom at the right-hand side of the part of the ring furthest from the observer. Sometimes the lower part of the ring, which represents the part nearest to the observer, is drawn in thick lines. Thus, to change V into VI, first draw the hexagon as shown in VI, and then place all the groups on the left-hand side in V above the plane of the ring in VI; all those on the right-hand side in V are placed below the

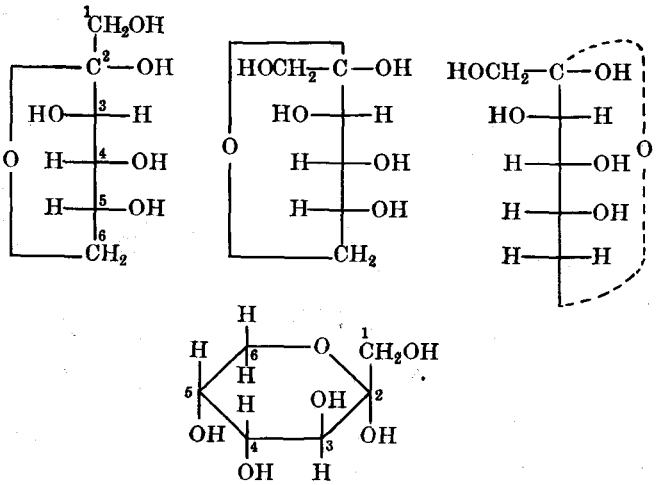
plane of the ring in VI. VII represents a " short-hand representation " of D-glucose.

In a similar manner, Haworth proposed a five-membered ring for γ-sugars based on the furan ring, VIII. Using the above scheme of transformation, the plane-diagram of methyl β-D(+)-glucofuranoside, IX, is first changed into X (*two* changes are carried out), and then X is twisted so as to be represented by XI, in which the oxygen atom is furthest from the observer.

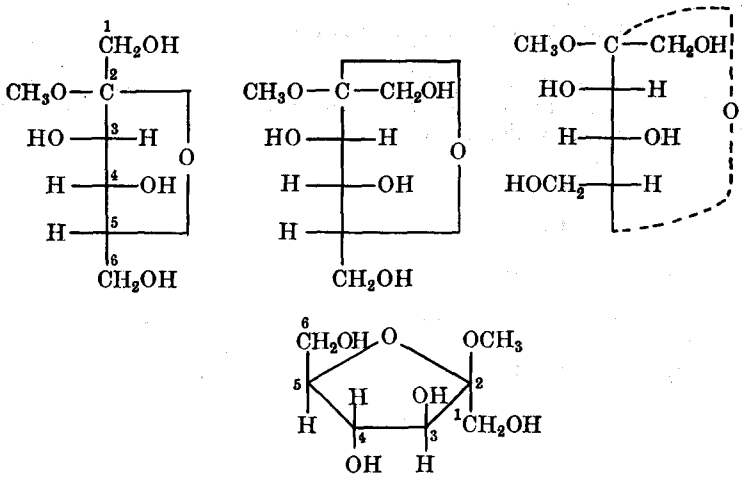


Two other examples which illustrate the conversion into the perspective formula are:

(i) **α-D(−)-fructopyranose.**



(ii) **Methyl β-D(+)-fructofuranoside.**



Actual size of sugar rings. Since glycoside formation under different conditions gives compounds containing different sized rings, the important question then is: What is the size of the ring in the original sugar? Oxidation of an aldose with hypobromite produces an unstable δ-lactone; this is the first product, but slowly changes into the stable γ-lactone (Hudson, 1932). It therefore follows that the size of the ring in *normal* sugars is pyranose. By analogy, ketoses are also believed to exist normally as pyranose compounds. This pyranose structure has been confirmed by X-ray analysis of various crystalline monosaccharides (Cox, 1935). McDonald *et al.* (1950) examined α-D-glucose by X-ray analysis, and confirmed the presence of the six-membered ring, the configuration as found chemically, and also the *cis* arrangement of the 1 : 2-hydroxyl groups in the α-form. Eiland *et al.* (1950) subjected difructose strontium chloride dihydrate to X-ray analysis, and showed the presence of a six-membered ring, and confirmed the configuration found chemically. It might be noted here that furanose sugars have not yet been isolated, but some furanosides have. It is also interesting to note that apparently fructose and ribose *always* occur in compounds as the furanose structure. Barker *et al.* (1959), however, have obtained evidence to show that D-ribose exists as the pyranose form at the moment of dissolution and its mutarotation involves change in size of the ring (*cf.* the fructose residue in sucrose, §13).

§7g. More recent methods for determining the size of the ring in sugars. These methods make use of the fact that periodic acid splits 1 : 2-glycols (Malaprade, 1928); thus periodic acid splits the following types of compounds (see also Vol. I):

$$\text{R·CHOH·CHOH·R}' \xrightarrow{1HIO_4} \text{R·CHO} + \text{R}'\text{·CHO}$$

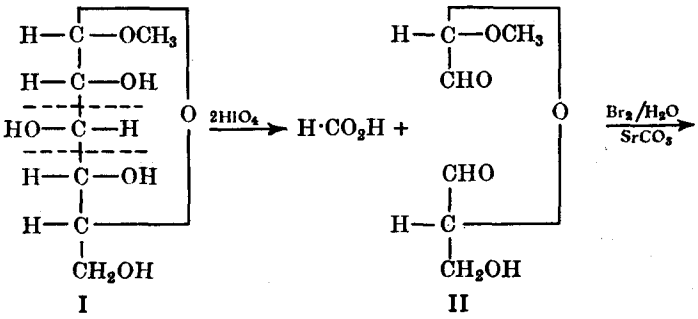
$$\text{R·CHOH·CO·R}' \xrightarrow{1HIO_4} \text{R·CHO} + \text{R}'\text{·CO}_2\text{H}$$

$$\text{R·CO·CO·R}' \xrightarrow{1HIO_4} \text{R·CO}_2\text{H} + \text{R}'\text{·CO}_2\text{H}$$

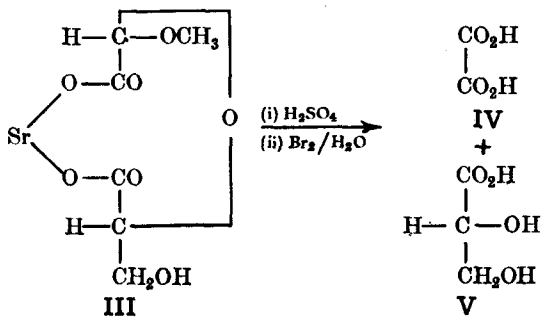
Thus a *free* sugar is broken down completely, *e.g.*,

$$\text{CH}_2\text{OH·CHOH·CHOH·CHOH·CHO} \xrightarrow{4HIO_4} \text{H·CHO} + 4\text{H·CO}_2\text{H}$$

In all of these reactions, one molecule of periodic acid is used for each pair of adjacent alcoholic groups (or oxo groups). Thus, by estimating the periodic acid used, and the formic acid and formaldehyde formed, the number of *free* adjacent hydroxyl groups in a sugar can be ascertained. Hudson (1937, 1939) oxidised " normal " methyl α-D-glucoside, I, with periodic acid, and found that two molecules of periodic acid were consumed, and that one molecule of formic acid was produced. It should be noted that although periodic acid can completely degrade a *free* sugar, the oxide ring in glycosides is sufficiently stable to resist opening by this reagent. The first product

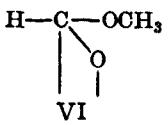


I II



III

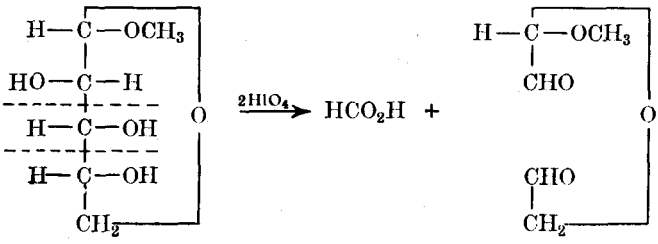
of oxidation of methyl α-D-glucoside was D'-methoxy-D-hydroxymethyldiglycolaldehyde, II, and this, on oxidation with bromine water in the presence of strontium carbonate, gave the crystalline salt, III. III, on acidification with sulphuric acid (for hydrolysis), followed by further oxidation with bromine water, gave oxalic acid, IV, and D(−)-glyceric acid, V. Isolation of II, III, IV and V indicates that the ring in I is δ-; this is also supported by the fact that only one carbon atom was eliminated as formic acid, and that two molecules of periodic acid were consumed. By similar experiments, it has been shown that all methyl α-D-hexosides of the " normal " type consume *two* molecules of periodic acid and produce *one* molecule of formic acid, and all also give products II, III, IV and V. Thus all these hexosides must be six-membered rings, and also it follows that all " normal " methyl α-pyranosides have the same configuration for C_1; this has already been shown to be VI.



VI

Similarly, all β-compounds, on oxidation with periodic acid, give the stereoisomer of II, *i.e.*, L′-methoxy-D-hydroxymethyldiglycolaldehyde.

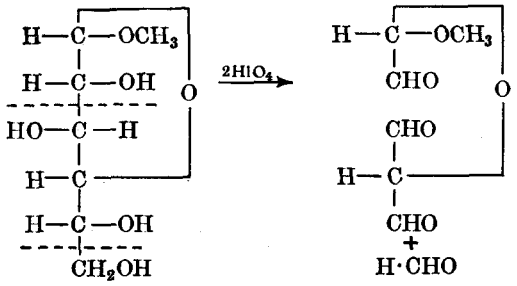
Aldopentopyranosides also give similar products as those obtained from the aldohexopyranosides, *e.g.*, methyl α-D-arabinopyranoside, VII, gives D′-methoxydiglycolaldehyde, VIII. Since all methyl α-D-aldopentopyrano-



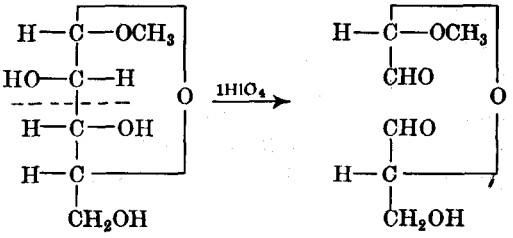
VII VIII

sides give the *same* diglycolaldehyde, they too have the same configuration for C_1, *viz.*, VI.

When hexofuranosides, *i.e.*, the " abnormal " glycosides, are oxidised with periodic acid, *two* molecules of acid are consumed and one molecule of *formaldehyde* is formed. These results are in keeping with the presence of a five-membered ring, *e.g.*, methyl α-D-glucofuranoside.



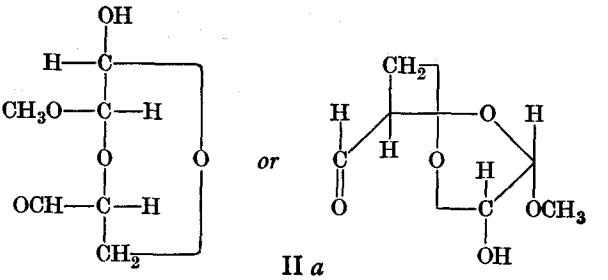
Oxidation of methyl α-D-arabinofuranoside, IX, consumes *one* molecule of periodic acid, and no carbon atom is eliminated (either as formaldehyde or formic acid); thus the ring is five-membered. Furthermore, since the dialdehyde II obtained is the same as that from methyl α-D-glucopyranoside, I, the configuration of C_1 is the same in both I and IX.



IX II

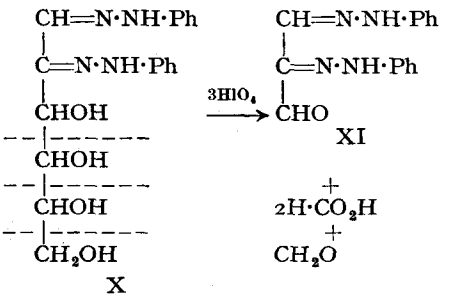
There appears to be some doubt about the structure of II. Various formulæ have been proposed (Hurd *et al.*, 1953; Smith *et al.*, 1955), and

Mester *et al.* (1957) have obtained evidence that of these structures the cyclic hemiacetal (II*a*) is the most likely.



II *a*

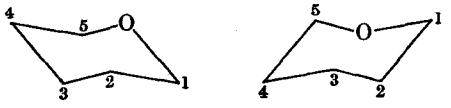
Hough *et al.* (1956) have carried out periodate oxidations on phenylosazones of reducing monosaccharides (X) and obtained formaldehyde, formic acid and mesoxalaldehyde 1 : 2-bisphenylhydrazone (XI). These authors found that XI is obtained from all monosaccharides in which C_3 and C_4 are



free, and 1 molecule of formaldehyde from the terminal CH_2OH group when this is free. They also showed that the osazones of the disaccharides maltose (§15), cellobiose (§16), and lactose (§17) did not give XI but did give formaldehyde. Thus C_3 or C_4 are linked in these disaccharides. On the other hand, the oxidation of the osazone of melibiose (§18) gave XI but no formaldehyde; thus C_6 is linked in this molecule. These oxidations therefore offer a means of differentiating between the two types of disaccharides.

§7h. Conformation of pyranoside rings. Cyclic 1 : 2-glycols form complexes in cuprammonium solutions, a five-membered ring being produced in which the copper atom is linked to two oxygen atoms. Furthermore, the extent of complex formation depends on the spatial arrangement of the two adjacent hydroxyls, the most favoured position being that in which the two groups and the two carbon atoms to which they are attached lie in one plane. Since complex formation changes the molecular rotation, the molecular rotational shift will indicate the extent of complex formation (*cf.* boric acid complexes, §4). Reeves (1950), using this cuprammonium complex formation, has shown that the pyranose sugars assume a chair form in preference to any boat form wherever both are structually possible. Substitution of an oxygen atom for a carbon atom in *cyclo*hexane causes only minor distortions in the ring (Hassel *et al.*, 1947), and consequently the general conformational features are retained in the pyranose sugars. Reeves (1951) proposed the two regular conformations shown, and named them C1 (the normal chair) and 1C (the reverse chair). Reeves (1958) pointed out that there is an infinite number of skew conformations in which angle strain is

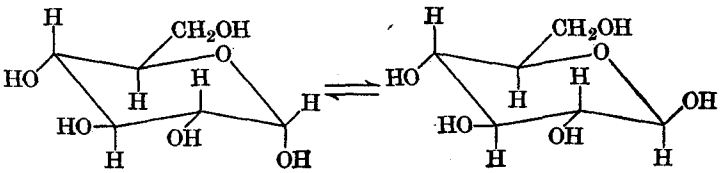
absent. It is still usual, however, to use the regular conformations of Reeves since these are readily related to the Haworth formulæ. Reeves has shown that the C1 conformation is the more stable, and this is supported by Barker *et al.* (1959) who studied the ring structures by periodate oxidations in buffered solutions. Also, according to these authors, the chief exceptions are β-D-altrose, β-D-mannose and β-D-talose, which are considered to be



CI IC

appreciably less stable in the 1C conformation. α-D-Lyxose appears to favour the C1 conformation, and the authors consider that α-D-allose, β-D-ribose, and α-D-xylose favour the 1C rather than C1 conformation.

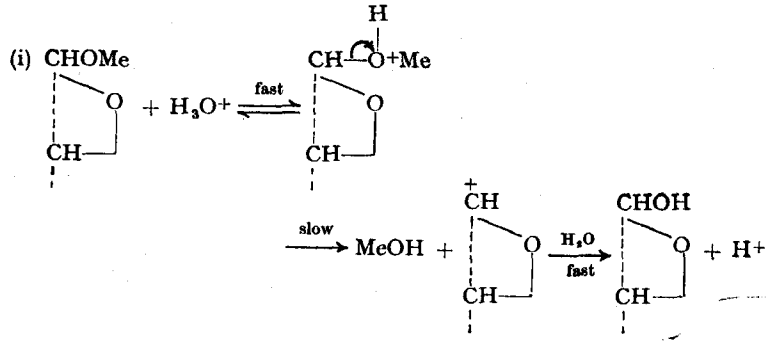
As we have seen (§2), D-glucopyranose is an equilibrium mixture (in solution) of the α- and β-anomers: the conformations of these are:

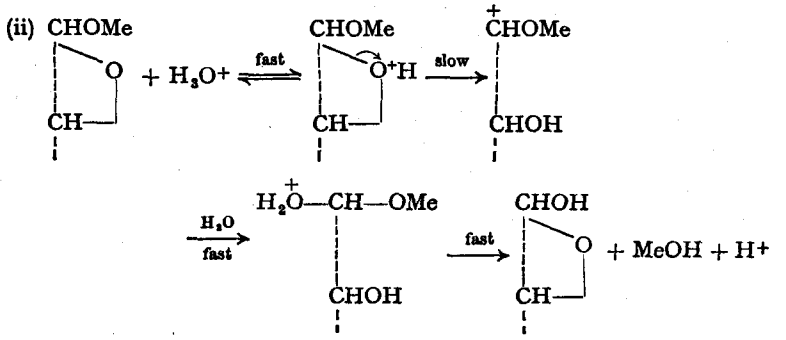


α-(38%) β-(62%)

We have also seen that the more stable isomer is the one with the larger number of equatorial substituents, and so the β-form can be expected to be more stable than the α-. Whiffen *et al.* (1954) have used infra-red spectroscopy to distinguish between α- and β-anomers; the absorption maxima depend on the axial or equatorial conformation of hydroxyl groups.

In general, β-anomers are more reactive than α-, *e.g.*, Bunton *et al.* (1954) have shown that acid-catalysed hydrolysis proceeds more rapidly for β-methyl pyranosides than for the corresponding α-compounds. According to these authors (1955), the hydrolysis proceeds by a unimolecular decomposition of the conjugate acids of the pyranosides. The rate-determining step, however, may be formulated in two ways, both of which are consistent with the evidence available at present.





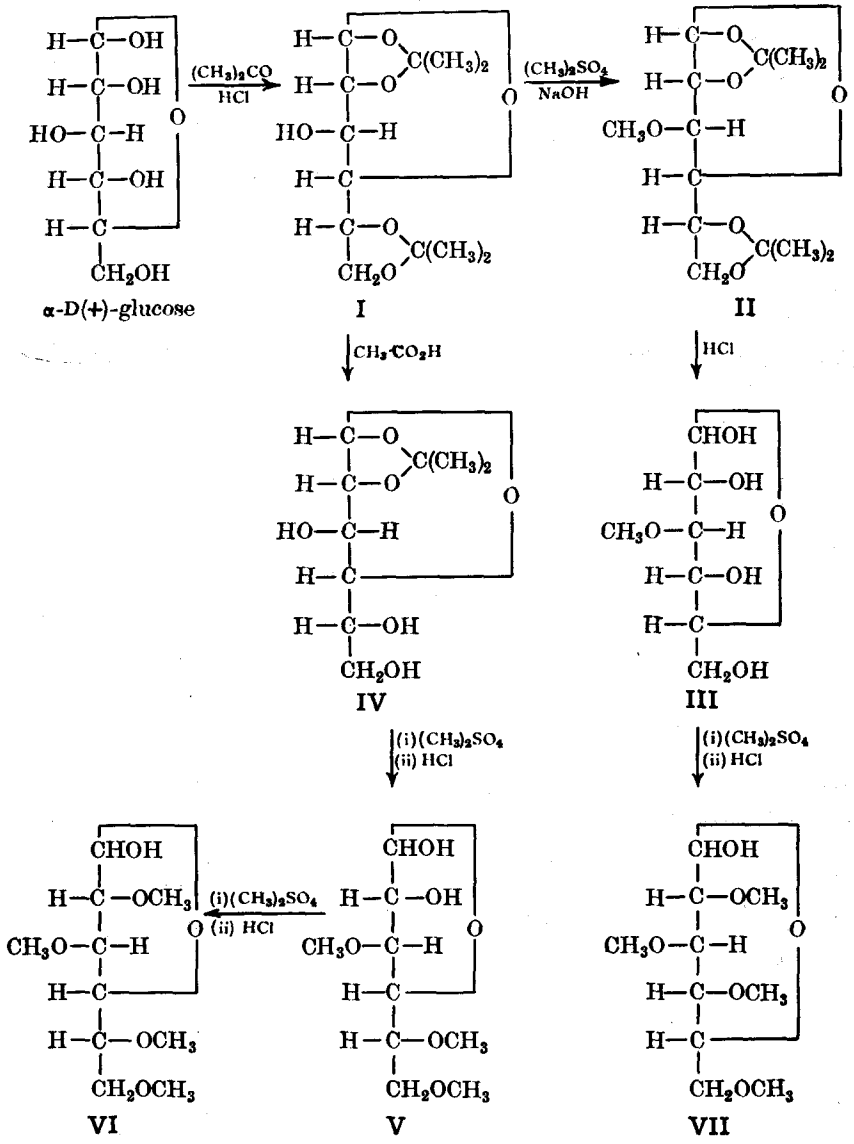
On the other hand, axial hydroxyl groups are less reactive (to esterification and hydrolysis reactions) than equatorial groups (§12. IV). In β-pyranosides, the methoxyl group is equatorial and so mechanism (i) would be more in keeping with the fact that β-anomers are more readily hydrolysed than α- (in which the methoxyl group is axial). However, Bunton *et al.* (1955) also showed that the rate of hydrolysis depends on the nature of the aglycon. In the above example the aglycon is methyl, but when it is phenyl then it is the α-anomer which is hydrolysed faster.

Since the hemiacetal linkage in the ring-form of reducing sugars is very labile, reactions involving the carbonyl group may possibly proceed through the acyclic or the cyclic form (see also mutarotation, §2). Isbell *et al.* (1932) have obtained evidence that the oxidation of an aldose with bromine-water proceeds to the 1,5-lactone by direct oxidation of the pyranose form. Isbell *et al.* (1932–1946) also showed that β-D-anomers (equatorial OH at C_1) are oxidised much faster than the corresponding α-D-anomers (axial OH at C_1). Further experiments on the oxidation of D-glucose by bromine-water appear to show that the α-anomer is first converted into the β-anomer which is then rapidly oxidised directly to δ-gluconolactone (Perlmutter-Hayman *et al.*, 1960). Pentoses (except D-lyxose) are also oxidised in the β-form (Overend *et al.*, 1960). Isbell (1961), however, disagrees with Overend's claim that the rate-determining step is the transformation of α-D-aldopyranoses into the β-anomers.

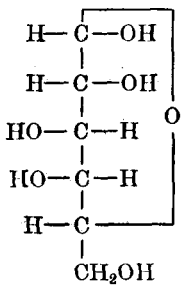
§8. *iso*Propylidene derivatives of the monosaccharides. Sugars condense with anhydrous acetone in the presence of hydrogen chloride, sulphuric acid, etc., at room temperature to form mono- and di-*iso*propylidene (or acetone) derivatives. These are stable towards alkalis, but are readily hydrolysed by acids. In the di-*iso*propylidene derivatives, one *iso*propylidene group is generally removed by hydrolysis more readily than the other, and thus by controlled hydrolysis it is possible to isolate the mono-*iso*propylidene derivative, *e.g.*, di-*iso*propylideneglucose may be hydrolysed by acetic acid to the mono-derivative.

The structures of these *iso*propylidene derivatives have been determined by the methods used for the sugars themselves, *i.e.*, the compound is first methylated, then hydrolysed to remove the acetone groups, and the product finally oxidised in order to ascertain the positions of the methyl groups. Let us consider D-glucose as an example. This forms a di-*iso*propylidene derivative, I, which is non-reducing; therefore C_1 is involved in the formation of I. On methylation, I forms a monomethyldi-*iso*propylideneglucose, II, and this, on hydrolysis with hydrochloric acid, gives a monomethylglucose, III. Hydrolysis of I with acetic acid produces a mono-*iso*propylideneglucose, IV, which is also non-reducing. Thus C_1 in IV must be combined with the *iso*propylidene radical. Methylation of IV, followed by hydrolysis,

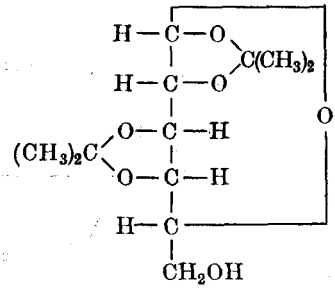
gives a trimethylglucose, V. Methylation of V gives a methyl tetramethylglucoside, and this, on hydrolysis, gives 2 : 3 : 5 : 6-tetra-*O*-methyl-D-glucose, VI, a *known* compound (see §7b). Thus V must be 2 : 3 : 5-, 2 : 3 : 6-, or 3 : 5 : 6-tri-*O*-methyl-D-glucose. Now V forms an osazone without loss of any methyl group; therefore C_2 cannot have a methoxyl group attached to it, and so V must be 3 : 5 : 6-tri-*O*-methyl-D-glucose. Thus one *iso*propylidene radical in di-*iso*propylideneglucose, I, must be 3 : 5-, 3 : 6- or 5 : 6-. Monomethylglucose, III, on methylation followed by hydrolysis, gives 2 : 3 : 4 : 6-tetra-*O*-methyl-D-glucose, VII, a *known* compound (see §7a). Hence III must be 2-, 3-, 4- or 6-*O*-methyl-D-glucose. Since III gives sodium cyanate when subjected to the Weerman test (see §11), it therefore follows that C_2 has a free hydroxyl group. Oxidation of III with nitric acid produces a monomethylsaccharic acid; therefore C_6 cannot have a methoxyl group attached to it. This monomethylsaccharic acid forms a lactone which behaves as a γ-lactone; therefore a methoxyl group cannot be at C_4. Thus, by the process of elimination, this monomethylglucose, III, must be 3-*O*-methyl-D-glucose. It therefore follows that the two *iso*propylidene groups in the di-*iso*propylidene derivative must be 1 : 2- and 5 : 6-, the ring being furanose, and the mono-*iso*propylidene derivative being 1 : 2-. The foregoing reactions can be written as on opposite page:



As a result of much experimental work (of the foregoing type), it has been found that acetone usually condenses with *cis*-hydroxyl groups on *adjacent* carbon atoms, the condensation occurring in such a way as to favour the formation of the di-*iso*propylidene derivative. For this to occur, the ring often changes size, *e.g.*, in α-D-galactopyranose, VIII, the hydroxyl groups on C_1 and C_2 are in the *cis* position, as are also the hydroxyl groups on C_3 and C_4. Thus galactose forms the 1 : 2–3 : 4-di-*O*-*iso*propylidene-D-galactopyranose, IX. On the other hand, in α-D-glucopyranose, only the two hydroxyl groups on C_1 and C_2 are in the *cis* position, and thus, in order to form the *di-iso*propylidene derivative, the ring changes from pyranose to furanose, the latter producing 1 : 2–5 : 6-di-*O*-*iso*propylidene-D-glucofuranose (I). The mono-derivative is 1 : 2-*O*-*iso*propylidene-D-glucofuranose

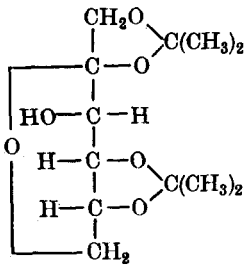


VIII

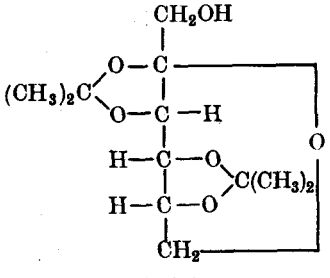


IX

(IV). Fructose can form *two* di-*iso*propylidene derivatives which both contain the pyranose ring.



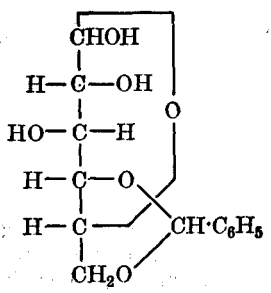
1:2-4:5-



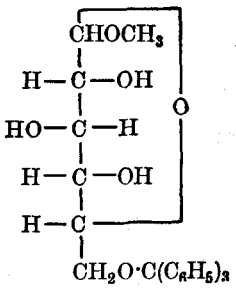
2:3-4:5-

§9. Other condensation products of the sugars. Not only does acetone condense with sugars, but so do other oxo compounds such as formaldehyde, acetaldehyde and benzaldehyde. Benzaldehyde condenses with two *cis* hydroxyl groups on *alternate* carbon atoms, *e.g.*, glucose forms 4 : 6-*O*-benzylidene-D-glucopyranose, I.

Triphenylmethyl chloride reacts with sugars to form triphenylmethyl ethers; these are usually known as *trityl* derivatives. Trityl ethers are



I

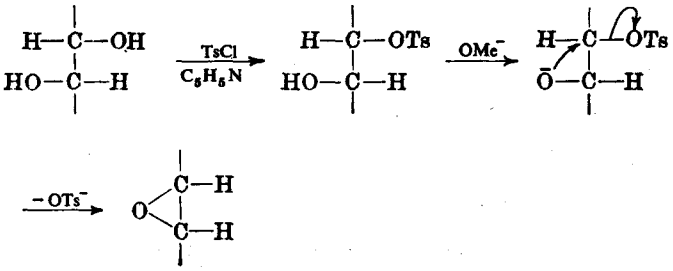


II

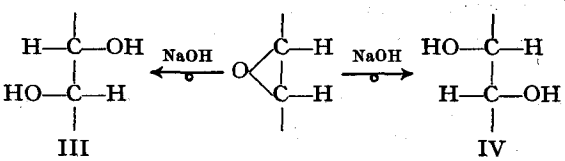
formed much faster with primary alcoholic groups than with secondary, *e.g.*, methyl glucopyranoside reacts with triphenylmethyl chloride in pyridine solution to form methyl 6-tritylglucopyranoside, II.

p-Toluenesulphonyl chloride (represented as TsCl in the following equations) reacts with sugars in the presence of pyridine to form *tosyl* esters. These esters usually produce **epoxy-sugars (anhydro sugars)** when hydrolysed with sodium methoxide in the cold, provided that there is a free

hydroxyl group on an adjacent carbon atom and that this hydroxyl and the tosyl group are *trans* to each other. This is an example of neighbouring hydroxyl group participation (§6c. III), and the mechanism is:

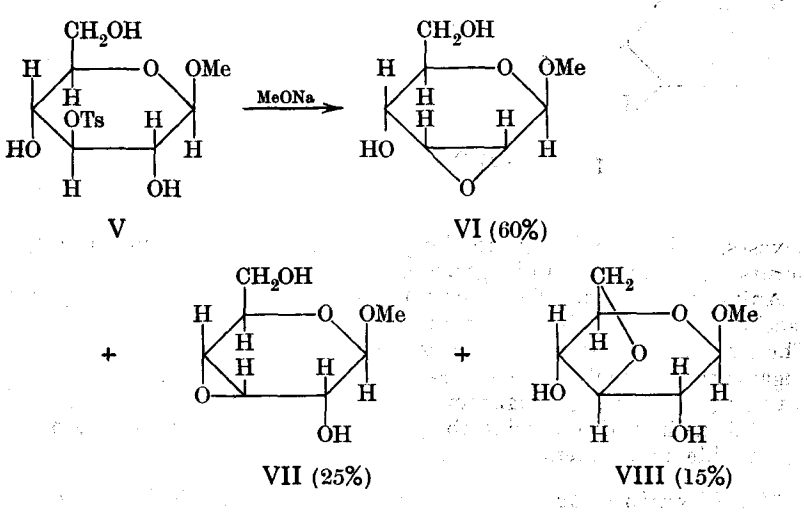


On hydrolysis with alkali, these anhydro sugars form a mixture of *two* sugars, inversion occurring at either carbon when the epoxide ring opens (see §5. IV).



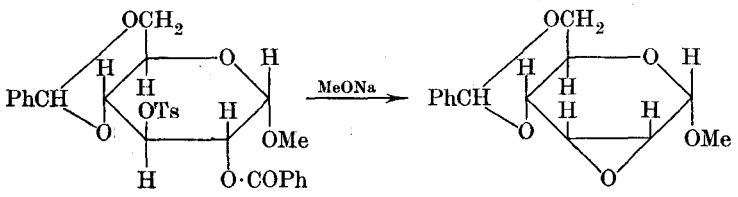
In III the configurations of the two carbon atoms are the same as in the original sugar, but in IV *both* configurations are inverted (to form a new sugar).

When the tosyl group is *trans* to two hydroxyl groups (on adjacent carbon atoms), *two* anhydro sugars are formed. At the same time, however, *larger* epoxide rings may be produced *without* inversion, *e.g.*, Peat *et al.* (1938) treated 3-tosyl methyl β-glucoside (V) with sodium methoxide and obtained a mixture of 2 : 3-anhydroalloside (VI; with inversion), 3 : 4-anhydroalloside (VII; with inversion), and 3 : 6-anhydroglucoside (VIII; no inversion).



It is possible, however, by using suitable derivatives of a tosyl ester to obtain only one anhydro sugar, *e.g.*, 2-benzoyl-3-tosyl 4 : 6-benzylidene

methyl α-glucoside (IX), on treatment with sodium methoxide, forms 2 : 3-anhydro 4 : 6-benzylidene methyl α-alloside (X).

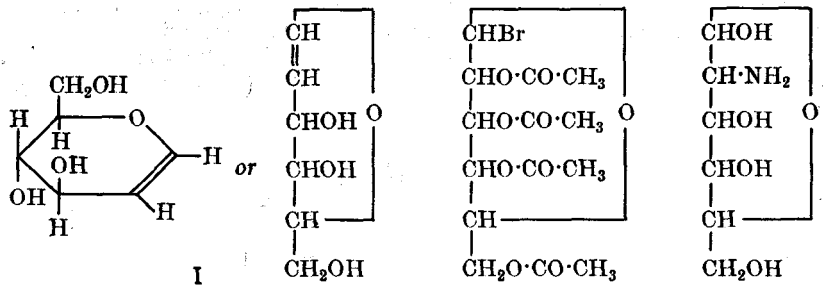


IX X

For the formation of the epoxide to proceed easily, it is necessary that the *trans* OH and Ts groups should be diaxial. In the majority of tosyl derivatives, however, both the tosyl group and the vicinal *trans*-hydroxyl group are equatorial (*cf.* §7h). Nevertheless, these tosyl derivatives are still easily converted into epoxides. This may be explained on the basis that the normal chair form (C1) readily changes into the reverse chair form (1C); consequently both groups are now axial and so epoxide formation proceeds readily (cf. §5b. IV).

§10. Glycals and glycosamines and anhydro sugars. Glycals are sugar derivatives which have a pyranose ring structure and a double bond between C_1 and C_2, *e.g.*, D-glucal is I. Glycals may be prepared by reducing acetobromo compounds (see §24) with zinc dust and acetic acid, *e.g.*, D-glucal from tetra-*O*-acetyl-D-glucopyranosyl bromide, II, followed by hydrolysis of the acetyl groups.

Glycosamines are amino-sugars in which a hydroxyl group has been replaced by an amino-group. All naturally occurring amino-sugars are



I II III

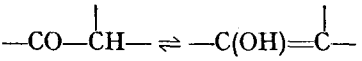
hexoses, and the amino-group always occurs on C_2, *e.g.*, glucosamine, which occurs in chitin, is 2-aminoglucose, III (see also §23).

Anhydro sugars. These may be regarded as being derived from monosaccharides by the elimination of a molecule of water to form an epoxide. The size of the oxiran ring varies from 1 : 2- to 1 : 6-. The 1 : 2-anhydro sugars are commonly known as α-glycosans, and may be prepared in various ways, *e.g.*, by heating a sugar under reduced pressure (Pictet *et al.*, 1920). A general method of producing the ethylene oxide series is by the hydrolysis of suitable tosyl esters (see §9).

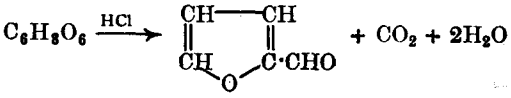
§11. Vitamin C or L-ascorbic acid. Ascorbic acid is very closely related to the monosaccharides, and so is conveniently dealt with here. Hawkins (1593) found that oranges and lemons were effective for treating

scurvy, a disease particularly prevalent among seamen. The first significant step in elucidating the nature of the compound, the absence of which from the diet caused scurvy, was that of Holst and Frölich (1907), who produced experimental scurvy in guinea-pigs. Then Szent-Györgi (1928) isolated a crystalline substance from various sources, *e.g.*, cabbages, paprika, etc., and found that it had antiscorbutic properties. This compound was originally called *hexuronic acid*, and later was shown to be identical with vitamin C, m.p. 192°, $[\alpha]_D$ of +24°.

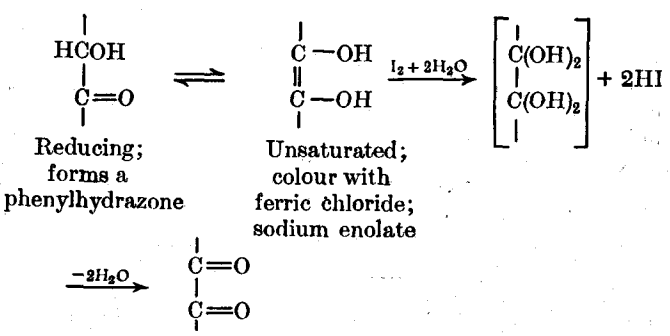
The structure of vitamin C was elucidated by Haworth, Hirst and their co-workers (1932, 1933). The molecular formula was shown to be $C_6H_8O_6$, and since the compound formed a monosodium and monopotassium salt, it was thought that there was a carboxyl group present. Vitamin C behaves as an unsaturated compound and as a strong reducing agent; it also forms a phenylhydrazone and gives a violet colour with ferric chloride. All this suggests that a keto-enol system is present, *i.e.*,



The presence of an aldehyde group was excluded by the fact that vitamin C does not give the Schiff reaction. Now, when boiled with hydrochloric acid, ascorbic acid gives a quantitative yield of furfuraldehyde:



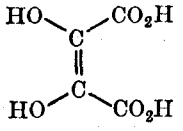
This reaction suggests that ascorbic acid contains at least five carbon atoms in a straight chain, and also that there are a number of hydroxyl groups present (*cf.* the pentoses). Aqueous iodine solution oxidises ascorbic acid to dehydroascorbic acid, two atoms of iodine being used in the process and two molecules of hydrogen iodide are produced; the net result is the removal of two hydrogen atoms from ascorbic acid. Dehydroascorbic acid is neutral and behaves as the lactone of a monobasic hydroxy-acid; and on reduction with hydrogen sulphide, dehydroascorbic acid is reconverted into ascorbic acid. Since this oxidation-reduction process may be carried out with " mild " reagents, it leads to the suggestion that since the oxidation product, dehydroascorbic acid, is a lactone, then ascorbic acid itself is a lactone and *not* an acid as suggested previously. Since, however, ascorbic acid can form salts, this property must still be accounted for. One reasonable possibility is that the salt-forming property is due to the presence of an *enol* group, the presence of which has already been indicated. Thus all the preceding reactions can be explained by the presence of an α-hydroxyketone grouping in ascorbic acid:



The final result is the removal of two hydrogen atoms to form dehydroascorbic acid.

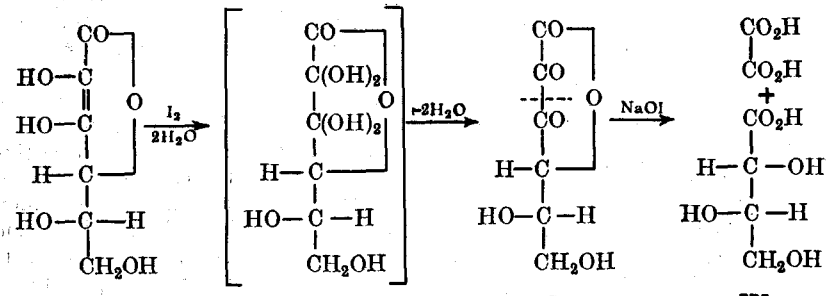
$$C_6H_8O_6 + I_2 \longrightarrow C_6H_6O_6 + 2HI$$

Although all these reactions may appear to be speculative, they are known to occur with dihydroxymaleic acid; hence by analogy with this compound, the explanation offered for the reactions of ascorbic acid is very strongly supported.



Dihydroxymaleic acid

When dehydroascorbic acid is oxidised with sodium hypoiodite, oxalic and L-threonic acids are produced in quantitative yields (Hirst, 1933). L-Threonic acid, IV, was identified by methylation and then conversion into the crystalline amide; this compound was shown to be identical with tri-*O*-methyl-L-threonamide (obtained from L-threose). Further evidence for the nature of product IV is given by the fact that on oxidation with nitric acid it gives D(+)-tartaric acid. The formation of oxalic and L-threonic acids suggests that dehydroascorbic acid is III, the lactone of 2 : 3-diketo-L-gulonic acid. Hence, if we assume that I is the structure of ascorbic acid, the foregoing reactions may be formulated as follows, dehydroascorbic acid being formed *via* II.

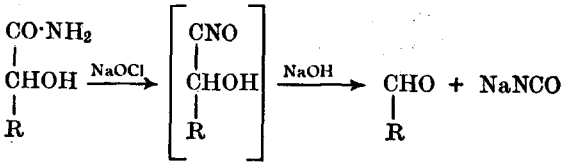


I II III IV

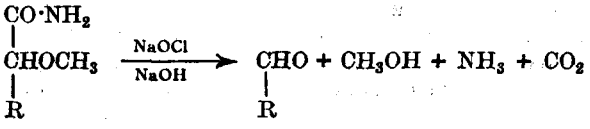
The ring in ascorbic acid has been assumed to be five- and not six-membered, because the lactone (*i.e.*, ascorbic acid) is stable towards alkali (*cf.* §7c). In actual fact, however, the same final products would also have been obtained had the ring been six-membered. It must therefore be admitted that the weakness of the above proof of structure lies in the evidence used for ascertaining the size of the ring. Structure I, however, has been amply confirmed by other analytical evidence. Diazomethane converts ascorbic acid into dimethylascorbic acid (V); these two methoxyl groups are most likely on C_2 and C_3, since diazomethane readily methylates acidic (in this case, enolic) hydroxyl groups. This dimethyl derivative is neutral, and dissolves in aqueous sodium hydroxide to form a sodium salt without the elimination of a methyl group; thus there cannot be a carbomethoxyl group present, and so it is most likely that two enolic hydroxyl groups are present (Hirst, 1933). Furthermore, the formation of the sodium salt from the neutral compound suggests the opening of a lactone ring (the two enolic

groups are now methylated and so cannot form a sodium salt). The similarity in structure between ascorbic acid and its dimethyl derivative is shown by the fact that the absorption spectra of both are similar. When this dimethyl derivative is methylated with methyl iodide in the presence of dry silver oxide (Purdie method; see §7), two further methyl groups are introduced (VI), and since all four methyl groups behave as methyl ethers, it therefore follows that two alcoholic groups are present in dimethylascorbic acid. Ozonolysis of this tetramethyl compound produces *one* neutral substance containing the *same* number of carbon atoms as its precursor. Since ozonolysis of a carbon–carbon double bond results in scission of that bond, there must be a ring system present in the tetramethyl compound to hold together the two fragments (VII). This ozonised product, on hydrolysis with barium hydroxide, gives oxalic acid and dimethyl-L-threonic acid (VIII). These products contain three carboxyl groups in all, and since ozonolysis of a double bond produces only two, the third carboxyl group must have already been present *as a lactone* in order that ascorbic acid should behave as a neutral compound.

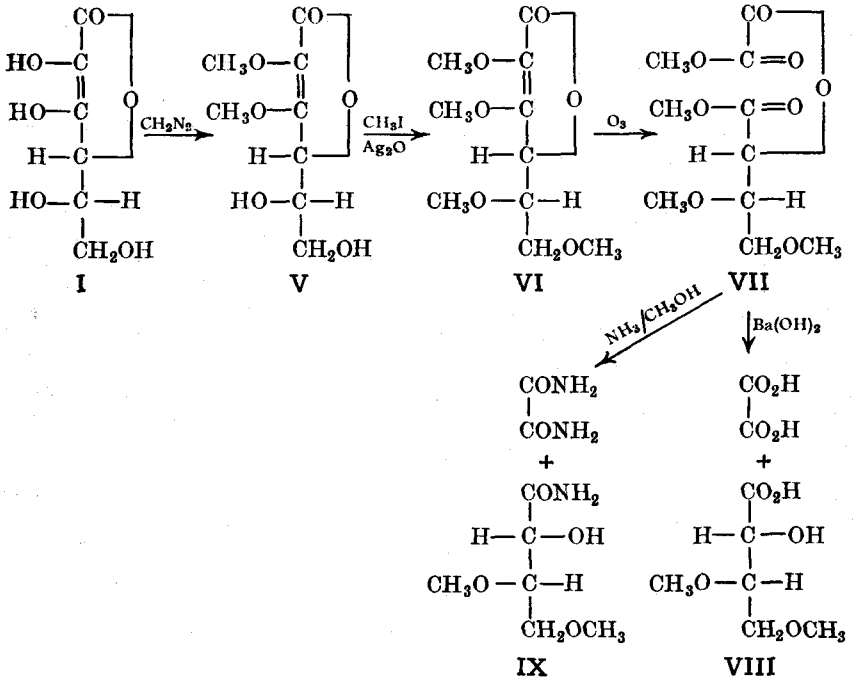
The key to the size of the ring in ascorbic acid is the structure of this dimethyl-L-threonic acid, the nature of which has been ascertained as follows. On methylation, followed by conversion to the amide, dimethyl-L-threonic acid gives trimethyl-L-threonamide. Thus this dimethyl compound, which was unknown when isolated, is a dimethyl-L-threonic acid; but where are the two methoxyl groups? Their positions were ascertained by means of the **Weerman test.** This test is used for showing the presence of a *free* hydroxyl group in the α-position to an amide group, i.e., in an α-hydroxy-amide. Treatment of a methylated hydroxy-amide with alkaline sodium



hypochlorite gives an aldehyde and *sodium cyanate* if there is a *free* hydroxyl group on the α-carbon atom. If there is no free hydroxyl group on the α-carbon atom, *i.e.*, this atom is attached to a methoxyl group, then treatment with alkaline sodium hypochlorite produces an aldehyde, methanol, ammonia and carbon dioxide.

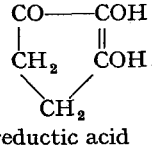


The dimethylthreonic acid obtained from the ozonised product was converted into the amide (IX), and this, when subjected to the Weerman test, gave sodium cyanate as one of the products. Thus this dimethylthreonic acid contains a free α-hydroxyl group, and consequently must be 3 : 4-di-*O*-methyl-L-threonic acid, VIII. Therefore the lactone ring in ascorbic acid must be γ-, since a δ-lactone could not have given VIII (actually, 2 : 4-di-*O*-methyl-L-threonic acid would have been obtained). The amide IX was also obtained, together with oxamide, by the action of ammonia in methanol on the ozonised product, VII. All the foregoing facts can be represented by the following equations:



I V VI VII IX VIII

An interesting point about ascorbic acid is that it is *not* reduced by lithium aluminium hydride (Petuely *et al.*, 1952). Thus ascorbic acid does not contain a " normal " carbonyl group. It has now been shown that all **reductones** are not reduced by lithium aluminium hydride. Reductones are compounds which contain the ene-α-diol-α-carbonyl grouping,

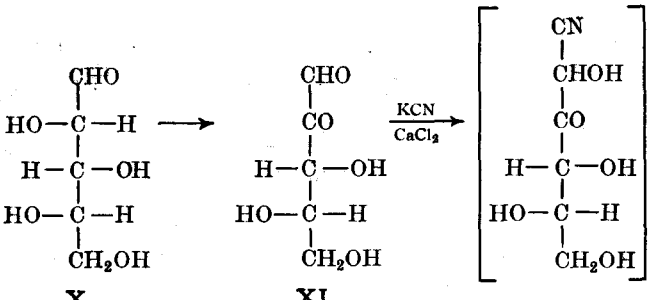


reductic acid

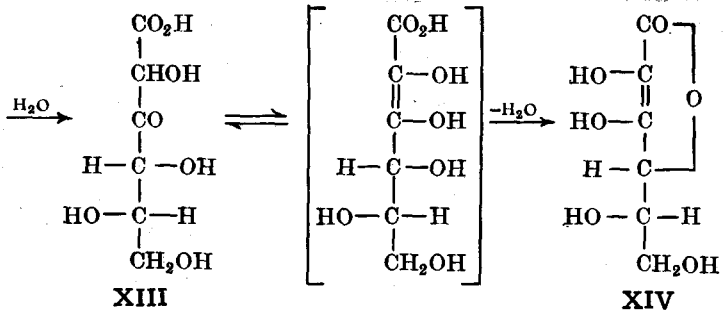
—CO—C(OH)=C(OH)—,

and examples of reductones are ascorbic and reductic acids.

Synthesis of ascorbic acid. Many methods of synthesising ascorbic acid are now available, *e.g.*, that of Haworth and Hirst (1933), L-Lyxose, X, was converted into L(−)-xylosone, XI (treatment with phenylhydrazine and then hydrolysis of the osazone with hydrochloric acid), and XI, on treat

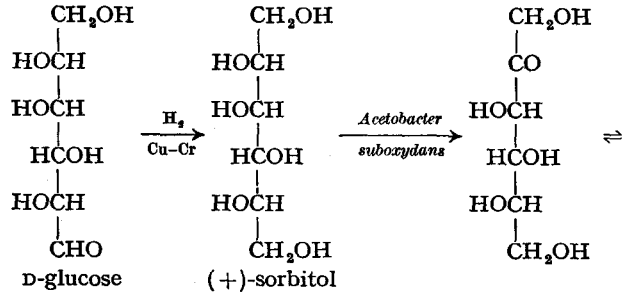


X XI XII

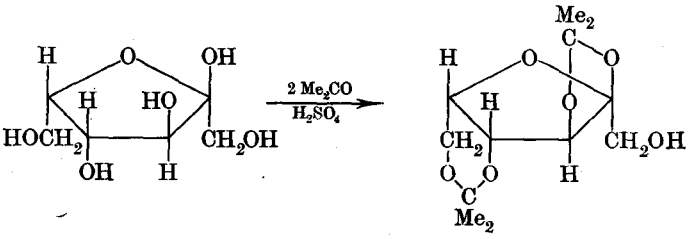


XIII XIV

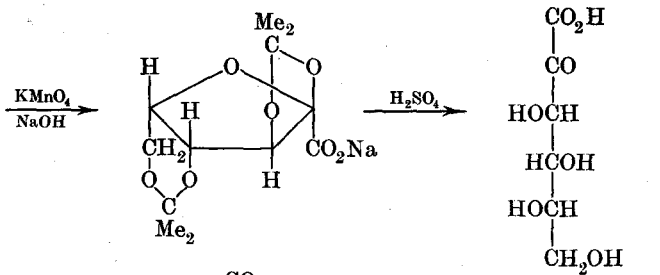
ment in an atmosphere of nitrogen with aqueous potassium cyanide containing calcium chloride, gave the β-keto-cyanide XII, which hydrolyses spontaneously into *pseudo*-L-ascorbic acid, XIII. This, on heating for 26 hours with 8 per cent. hydrochloric acid at 45–50°, gave a quantitative yield of L(+)-ascorbic acid, XIV.



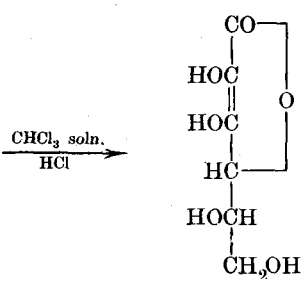
D-glucose (+)-sorbitol



(−)-sorbose diacetone-(−)-sorbose



2-ketogulonic-acid



L-ascorbic acid

Ascorbic acid is now synthesised commercially by several methods, *e.g.*, D-glucose is catalytically hydrogenated to (+)-sorbitol which is then converted into (−)-sorbose by microbiological oxidation (using *Acetobacter suboxydans* or *Acetobacter xylinum*). (−)-Sorbose can be oxidised directly to 2-keto-(−)-gulonic acid with nitric acid, but the yield is less than when the oxidation is carried out as shown above. Nitric acid oxidises other alcohol groups besides the first, but by protecting these by means of 2 : 3–4 : 6-di-*iso*propylidene formation (§8), the yield of the gulonic acid is higher. The gulonic acid is then dissolved in mixed solvents (of which chloroform is the main constituent) and hydrogen chloride passed in. The product, L-ascorbic acid, is then finally purified by charcoaling (see previous page).

Biosynthesis of ascorbic acid (see also §32a. VIII). Horowitz *et al.* (1952) and Burns *et al.* (1956) have shown that rat and plant tissues can convert D-glucose into ascorbic acid. A very interesting observation is that glucose labelled at C_1 (with ^{14}C) produces the vitamin labelled at C_6. In this way, the glucose molecule is " turned upside down " to form the glucose derivative (*cf.* the stereochemistry of glucose and gulose, §1).

DISACCHARIDES

§12. Introduction. The common disaccharides are the dihexoses, and these have the molecular formula $C_{12}H_{22}O_{11}$. Just as methanol forms methyl glycosides with the monosaccharides, so can other hydroxy compounds also form glycosides. The monosaccharides are themselves hydroxy compounds, and so can unite with other monosaccharide molecules to form glycosidic links. Study of the disaccharides (of the dihexose type) has shown that three types of combination occur in the natural compounds:

(i) The two monosaccharide molecules are linked through their reducing groups, *e.g.*, sucrose.
(ii) C_1 of one molecule is linked to C_4 of the other, *e.g.*, maltose.
(iii) C_1 of one molecule is linked to C_6 of the other, *e.g.*, melibiose.

Since the glycosidic link may be α or β, then different stereoisomeric forms become possible for a given pair of hexoses. In group (i), there are four forms possible theoretically: α_1–α_2, α_1–β_2, β_1–α_2 and β_1–β_2. In groups (ii) and (iii), the reducing group of the second molecule is free, and so in these two cases there are only two possibilities: α_1- and β_1-. In group (i), since *both* reducing groups are involved in glycoside formation, the resultant disaccharide will be non-reducing. In groups (ii) and (iii), since *one* reducing group is free, the resultant disaccharide will be reducing, and can exist in two forms, the α- and β-.

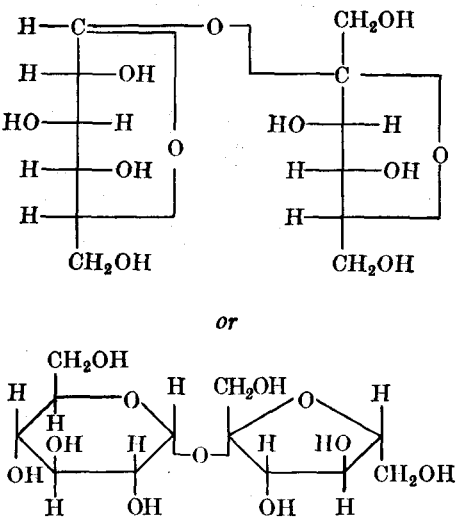
General procedure. The disaccharide is first hydrolysed with dilute acids, and the two monosaccharide molecules then identified. One of the earlier methods of separating sugars in a sugar mixture was by fractional crystallisation; the separation and identification is now carried out by means of partition chromatography. When the constituents have been identified, the next problem is to ascertain which hydroxyl group of the molecule acting as the alcohol (*i.e.*, the aglycon; §3) is involved in forming the glycosidic link. This is done by completely methylating the disaccharide; the methyl glycoside (of a reducing sugar) cannot be prepared by means of methanol and hydrochloric acid, since this will lead to hydrolysis of the disaccharide. Purdie's method cannot be used for reducing disaccharides since these will be oxidised (see §7). The only satisfactory way is Haworth's method, and to ensure complete methylation, this may be *followed* by the Purdie method. The methylated disaccharides are then hydrolysed, and the methylated monosaccharides so obtained are investigated by the oxidation

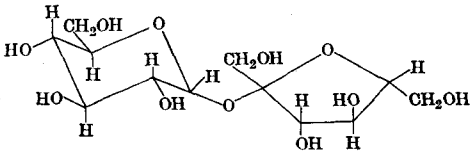
methods described previously (see §§7a, 7b, 7e). Reducing disaccharides are also oxidised to the corresponding bionic acid, this is then fully methylated, hydrolysed, and the methylated monosaccharide molecules examined. By this means the hydroxyl group involved in the glycosidic link and the size of the oxide ring are ascertained.

The final problem is to decide whether the glycosidic link is α or β. This is done by means of enzymes, maltase hydrolysing α-glycosides and emulsin β-glycosides (*cf.* §3). In non-reducing sugars, the problem is far more difficult since the links α_1–α_2, α_1–β_2, β_1–α_2 would *all* be hydrolysed by maltase. Consideration of the optical rotations has given information on the nature of the link (*cf.* §6). Finally, a number of disaccharides have been synthesised, the acetobromo-sugars being the best starting materials (see §24).

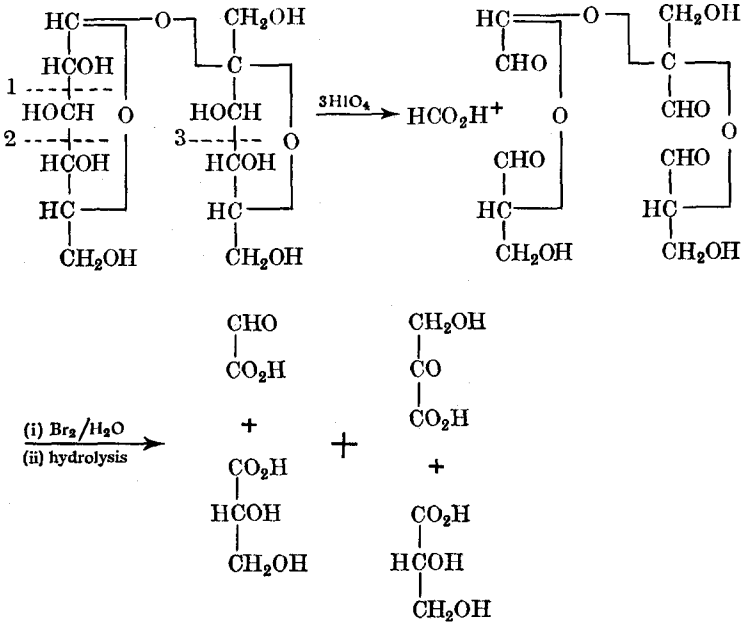
§13. Sucrose. Sucrose has been shown to be α-D-glucopyranosyl-β-D-fructofuranoside. Sucrose is hydrolysed by dilute acids or by the enzyme invertase to an equimolecular mixture of D(+)-glucose and D(−)-fructose. Methylation of sucrose (Haworth method) gives octa-*O*-methylsucrose and this, on hydrolysis with dilute hydrochloric acid, gives 2 : 3 : 4 : 6-tetra-*O*-methyl-D-glucose and 1 : 3 : 4 : 6-tetra-*O*-methyl-D-fructose. The structures of these compounds were determined by the oxidation methods previously described (see §§7a, 7e). Thus glucose is present in the pyranose form, and fructose as the furanose.

Since sucrose is a non-reducing sugar, both glucose and fructose must be linked *via* their respective reducing groups. The stereochemical nature of the glycosidic link may be any one of the four possibilities discussed (see §12), but the evidence indicates that it is α-glucose linked to β-fructose. Maltase hydrolyses sucrose; therefore an α-link is present. Furthermore, since the mutarotation of the glucose produced is in a downward direction, it therefore follows that α-glucose is liberated at first. The mutarotation of fructose is too rapid to be followed experimentally, and hence the nature of the link in this component remains to be determined. There is, however, an enzyme which hydrolyses methyl β-fructofuranosides, and it has been found that it also hydrolyses sucrose. This suggests that fructose is present in sucrose in the β-form, and is supported by calculations of the optical rotation of the fructose component. The following structure for sucrose accounts for all of the above facts:



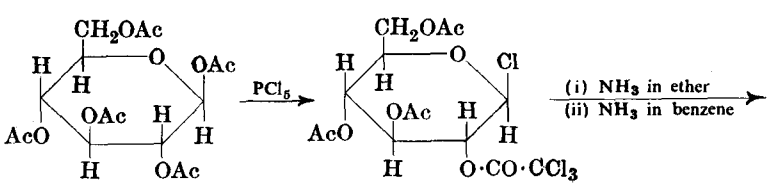


Oxidation of sucrose with periodic acid confirms this structure (but not the nature of the glycosidic link). Three molecules of periodic acid are consumed, and one molecule of formic acid is produced. Subsequent oxidation with bromine water, followed by hydrolysis, gives glyoxylic, glyceric and hydroxypyruvic acids (Fleury *et al.*, 1942).

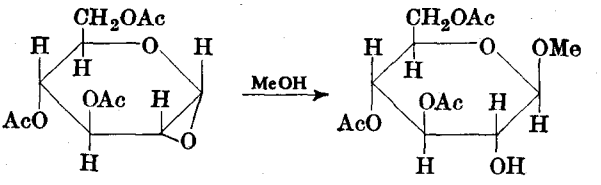


Beevers *et al.* (1947) examined sucrose sodium bromide dihydrate by X-ray analysis, and confirmed the stereochemical configuration found chemically, and also showed that the fructose ring is five-membered.

Sucrose has now been synthesised by Lemieux *et al.* (1953, 1956). Brigl (1921) prepared the sugar epoxide, 3 : 4 : 6-tri-*O*-acetyl-1 : 2-anhydro-α-D-glucose, II, from tetra-*O*-acetyl-β-D-glucose, I (*cf.* §9; see also §24).

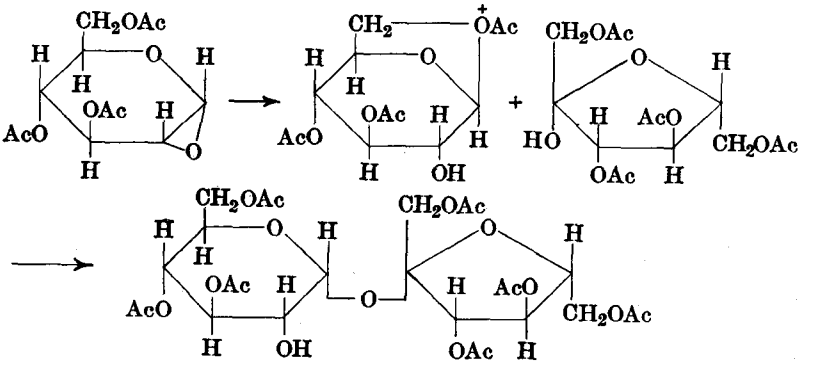


I



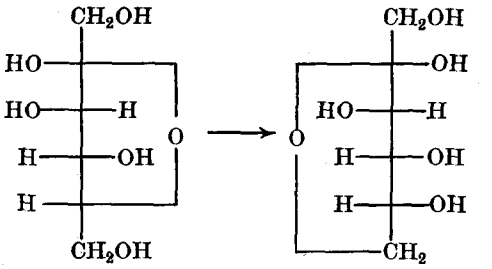
II III

Brigl also showed that II reacted with methanol to give methyl β-D-glucopyranoside triacetate, III, whereas with phenol, the α-glucopyranoside was the main product. Other workers showed that secondary alcohols gave α,β-mixtures. Lemieux was therefore led to believe that fructofuranose, a hindered secondary alcohol, would react with anhydroglucopyranose to form an α-glucose linkage. 1 : 2-Anhydro-α-D-glucopyranose triacetate and 1 : 3 : 4 : 6-tetra-*O*-acetyl-D-fructofuranose were heated in a sealed tube at 100° for 104 hours. The product, sucrose octa-acetate, on deacetylation, gave sucrose (yield about 5 per cent.). According to Lemieux, the reaction proceeds as follows:

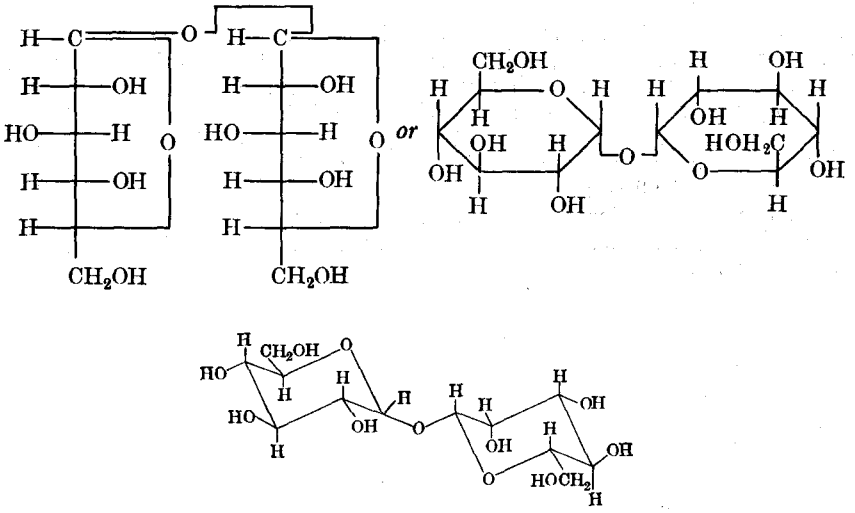


The CH_2OAc group at position 6 in the glucopyranose molecule enters into neighbouring group participation in the opening of the oxide ring, and consequently shields this side from attack. Thus the fructofuranose molecule is forced to attack from the other side and this produces the desired α-glucopyranose linkage.

One other point that is of interest is the "inversion" of sucrose on hydrolysis. Hydrolysis of sucrose gives first of all α-D(+)-glucopyranose and β-D(+)-fructofuranose (this is believed to be dextrorotatory), but the latter is unstable and immediately changes into the stable form, D(−)-fructopyranose (the rotation of (−)-fructose is much greater than that of (+)-glucose).

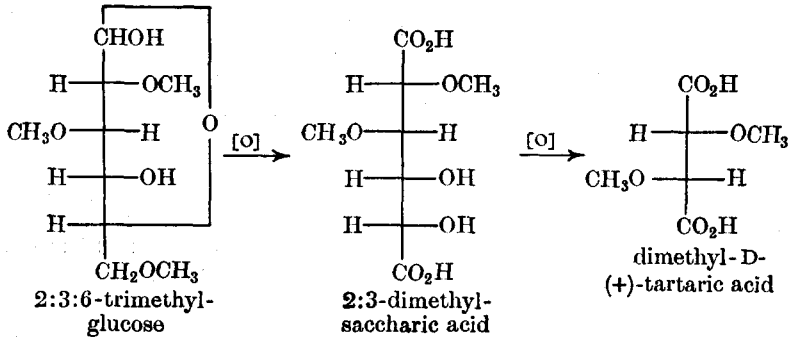


§14. Trehalose. This is believed to be α-D-glucopyranosyl-α-D-glucopyranoside. It is a non-reducing sugar which occurs in yeasts and fungi. It is hydrolysed by hydrochloric acid to two molecules of D-glucose; methylation of trehalose gives octa-*O*-methyltrehalose which, on hydrolysis, produces two molecules of 2 : 3 : 4 : 6-tetra-*O*-methyl-D-glucose (see §7a). The nature of the glycosidic link is uncertain, but there is some evidence to show that it is α : α, *e.g.*, the high positive rotation. Thus trehalose may be written.



§15. Maltose. This is 4-*O*-α-D-glucopyranosyl-D-glucopyranose. Maltose is hydrolysed by dilute acids to two molecules of D-glucose; it is a reducing sugar, undergoes mutarotation, and forms an osazone. Thus there is one free reducing group present, and since maltose is hydrolysed by maltase, the glycosidic link of the non-reducing half of the molecule is therefore α-. Complete methylation of maltose gives an octamethyl derivative which is non-reducing, and this, on hydrolysis with very dilute cold hydrochloric acid, is converted into heptamethylmaltose, which has reducing properties. Thus the original octamethyl derivative must be methyl hepta-*O*-methyl-D-maltoside; this is further evidence that only *one* free reducing group is present in maltose. Hydrolysis of hepta-*O*-methylmaltose with moderately concentrated hydrochloric acid produces 2 : 3 : 6-tri-*O*-methyl-D-glucose and 2 : 3 : 4 : 6-tetra-*O*-methyl-D-glucose. The structure of the latter is known (see §7a), but that of the former was elucidated as follows. Analysis of the compound showed that it was a trimethyl derivative, and since it formed a phenylhydrazone but not an osazone, C_2 must therefore be attached to a methoxyl group. On further methylation, this trimethylglucose gave 2 : 3 : 4 : 6-tetra-*O*-methyl-D-glucose, and so the trimethyl compound must be one of the following: 2 : 3 : 4-, 2 : 3 : 6- or 2 : 4 : 6-tri-*O*-methyl-D-glucose. Now, on careful oxidation with nitric acid, the trimethylglucose forms a dimethylsaccharic acid. This acid contains two terminal carboxyl groups; one has been derived from the free " aldehyde " group, and the other by oxidation at C_6, and since in its formation one methyl group is lost, this dimethylsaccharic acid must have been derived from a trimethylglucose having a methoxyl group at C_6. Thus the trimethylglucose must be either 2 : 3 : 6- or 2 : 4 : 6-tri-*O*-methyl-D-glucose. On further oxidation, the dimethylsaccharic acid forms dimethyl-D-tartaric acid; this can only arise from a precursor with two methoxyl groups on *adjacent* carbon atoms, and so it

follows that the trimethylglucose must be 2 : 3 : 6-tri-*O*-methyl-D-glucose. This is confirmed by the fact that the other two possible compounds, *viz.*, 2 : 3 : 4- and 2 : 4 : 6-tri-*O*-methyl-D-glucose, have been synthesised, and were shown to be different from the trimethylglucose obtained from maltose. The foregoing reactions may thus be written:

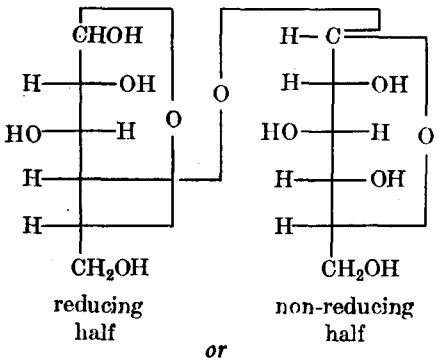


2:3:6-trimethyl-glucose

2:3-dimethyl-saccharic acid

dimethyl-D-(+)-tartaric acid

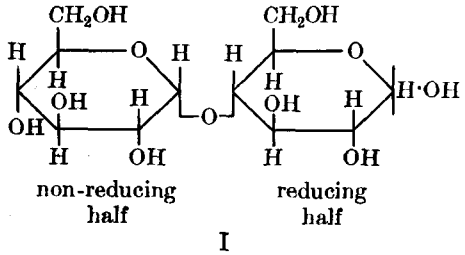
From this it can be seen that structure I for maltose satisfies all the above facts. This structure, however, is not the only one that satisfies all the facts. The structure of the non-reducing half is certain, but that of the reducing half need not necessarily be pyranose as shown in I, since a furanose structure, II, would also give 2 : 3 : 6-tri-*O*-methyl-D-glucose. To decide whether C_4 (as in I) or C_5 (as in II) was involved in the glycosidic link,



reducing half

non-reducing half

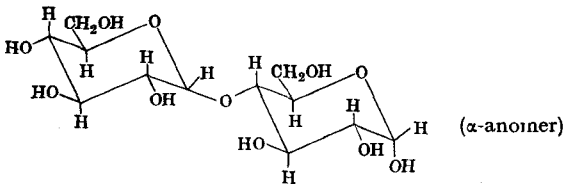
or



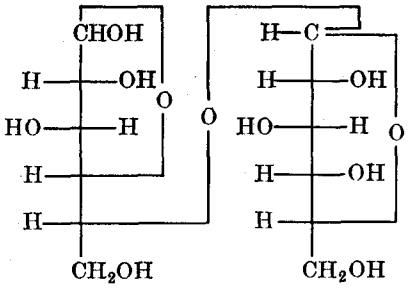
non-reducing half

reducing half

I

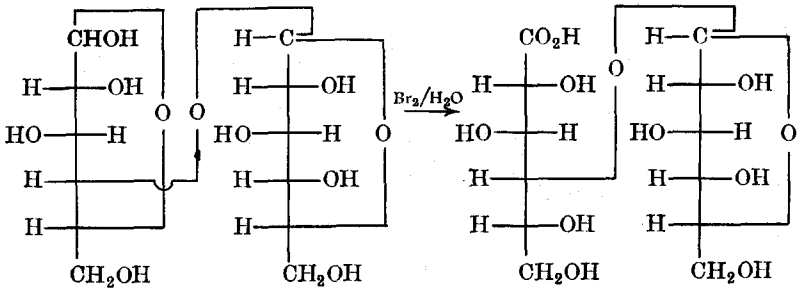


(α-anomer)

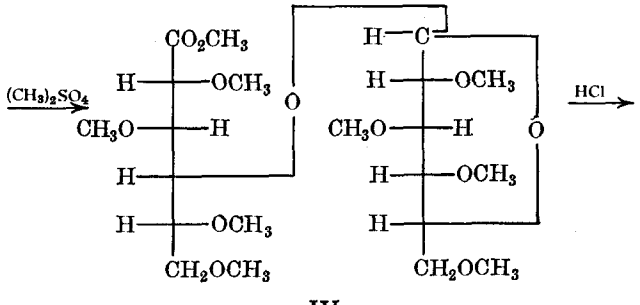


II

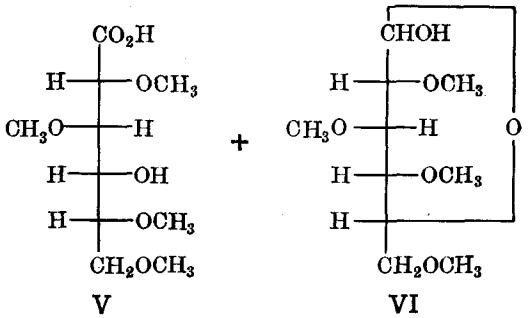
Haworth *et al.* (1926) oxidised maltose with bromine water to maltobionic acid, III, and this, on methylation, gave the methyl ester of octamethyl-maltobionic acid, IV, which, on vigorous hydrolysis, gave 2 : 3 : 5 : 6-tetra-*O*-methyl-D-gluconic acid, V (as lactone), and 2 : 3 : 4 : 6-tetra-*O*-methyl-D-



I III



IV

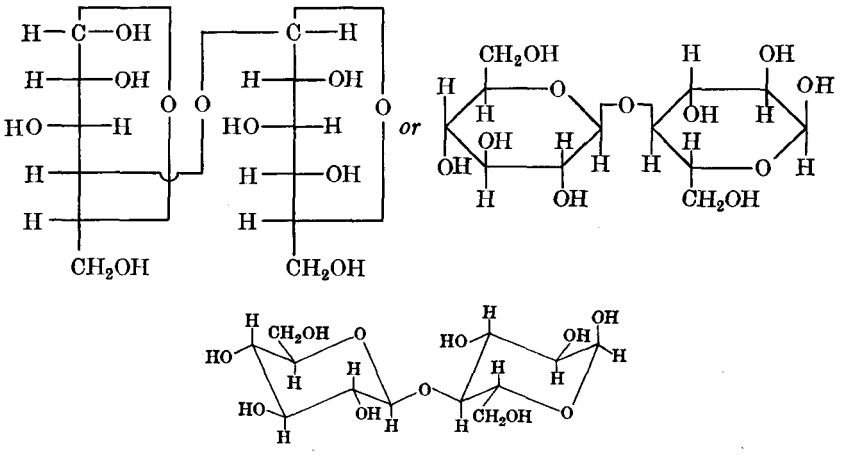


V VI

glucose, VI. V can be obtained only if maltose has structure I; structure II would have given 2 : 3 : 4 : 6-tetra-*O*-methyl-D-gluconic acid. Thus maltose is I and not II. Confirmation of the α-glycosidic linkage is afforded by the agreement of the specific rotation of maltose with that calculated for structure I, and further evidence for the linkage at C_4 is as follows. Since maltose is a reducing sugar, C_1 (of the reducing half) is free, and since maltose forms an osazone, C_2 is also free, *i.e.*, not combined with an alkoxyl group. Zemplen (1927) degraded maltose by one carbon atom (see Vol. I), and obtained a compound which still formed an osazone; therefore C_3 is free. On further degrading by one carbon atom, a compound was obtained which did *not* form an osazone; therefore C_4 in maltose is not free (see also §7g).

Maltose has been synthesised by the action of yeast on D-glucose (Pringsheim *et al.*, 1924), and maltose octa-acetate has been synthesised by heating a mixture of equimolecular amounts of α- and β-D-glucose at 160°, and then acetylating the product (Pictet *et al.*, 1927).

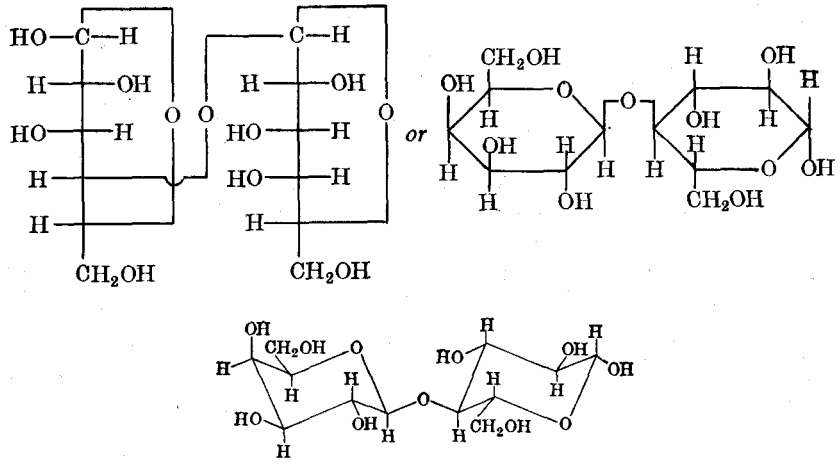
§16. Cellobiose (4-*O*-β-D-glucopyranosyl-D-glucopyranose). Cellobiose is hydrolysed by dilute acids to two molecules of D(+)-glucose; since this hydrolysis is also effected by emulsin, the glycosidic link must be β. Cellobiose is a reducing sugar, and so one reducing group is free. Methylation, followed by hydrolysis, gives 2 : 3 : 6-trimethyl-D-glucose and 2 : 3 : 4 : 6-tetramethyl-D-glucose (these are the same products obtained from maltose, §15). Oxidation with bromine water converts cellobiose into cellobionic acid, and this, on methylation followed by hydrolysis, gives 2 : 3 : 5 : 6-tetramethylgluconic acid and 2 : 3 : 4 : 6-tetramethylglucose (again the same products as for maltose). Thus cellobiose and maltose differ only in that the former has a β-glycosidic link, whereas the latter has an α-. Thus cellobiose is (α-form):



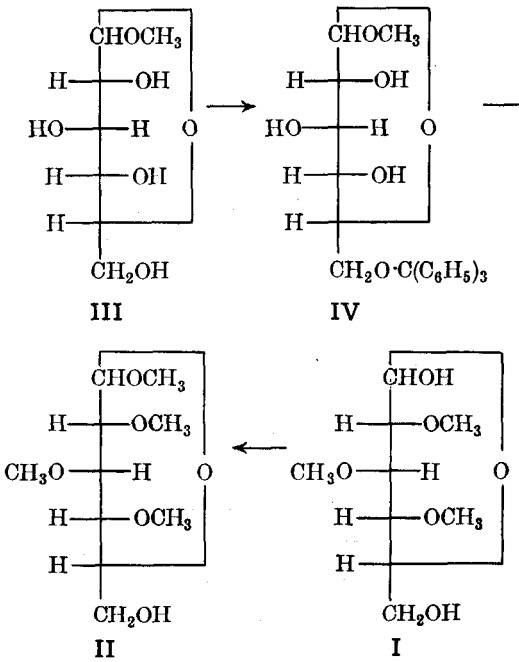
Degradation experiments confirm the C_4 linkage (see also §7g), and the structure has also been confirmed by synthesis (*e.g.*, Stacey, 1946).

§17. Lactose (4-*O*-β-D-galactopyranosyl-D-glucopyranose). Lactose is a reducing sugar, and is hydrolysed by dilute acids to one molecule of D(+)-glucose and one molecule of D(+)-galactose. Since lactose is hydrolysed by lactase (which has been shown to be identical with the β-glycosidase in emulsin), the two monosaccharide molecules are linked by a β-glycosidic link. The evidence, given so far, does not indicate which molecule is the reducing half. On methylation, lactose forms methyl heptamethyl-lactoside, and this, on vigorous hydrolysis, gives 2 : 3 : 6-tri-*O*-methyl-D-glucose

(see §15) and 2 : 3 : 4 : 6-tetra-*O*-methyl-D-galactose; thus glucose is the reducing half. Oxidation with bromine water converts lactose into lactobionic acid, and this, on methylation followed by hydrolysis, gives 2 : 3 : 5 : 6-tetra-*O*-methyl-D-gluconic acid and 2 : 3 : 4 : 6-tetra-*O*-methyl-D-galactose. Lactose is therefore (β-form) [see also §7g]:

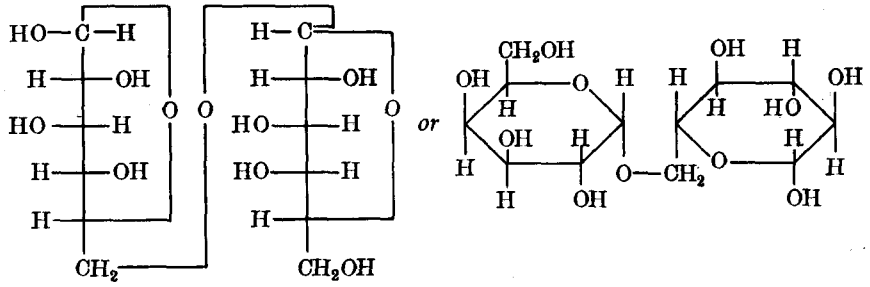


§18. Melibiose (6-*O*-α-D-galactopyranosyl-D-glucopyranose). This disaccharide is obtained from the trisaccharide, raffinose (§20); it is a reducing



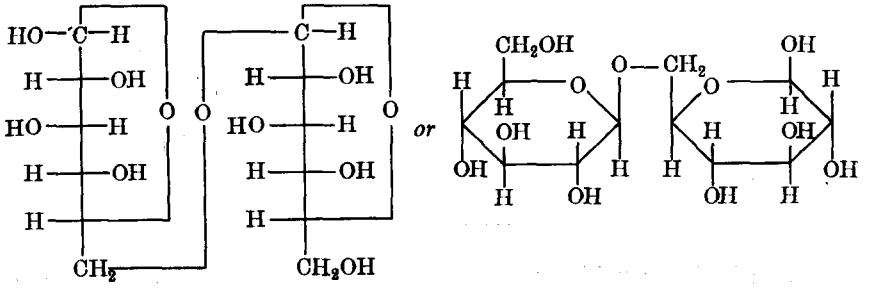
sugar, forms an osazone, and undergoes mutarotation. When hydrolysed by dilute acids, melibiose gives D-glucose and D-galactose. Methylation converts melibiose into methyl heptamethylmelibioside, and this, on hydrolysis, forms 2 : 3 : 4-trimethyl-D-glucose and 2 : 3 : 4 : 6-tetramethyl-D-galac-

tose. The structure of the former has been established as follows. The trimethylglucose, I, readily forms a crystalline methyl trimethylglucoside, II. Now methyl glucopyranoside, III, can be converted into the 6-trityl derivative, IV (see §9), and this, on methylation followed by removal of the trityl group, gives II. Thus II must be methyl 2 : 3 : 4-tri-*O*-methyl-D-glucopyranoside, and consequently I is 2 : 3 : 4-tri-*O*-methyl-D-glucose. From the foregoing facts, it can be seen that galactose is the non-reducing half of melibiose, and that its reducing group is linked to C_6 of glucose, the reducing half. This has been confirmed by oxidation of melibiose with bromine water to melibionic acid, and this, on methylation followed by hydrolysis, gives 2 : 3 : 4 : 5-tetra-*O*-methyl-D-gluconic acid and 2 : 3 : 4 : 6-tetra-*O*-methyl-D-galactose; the structure of the former is shown by the fact that, on oxidation with nitric acid, it forms tetramethylsaccharic acid. There has been some doubt about the nature of the glycosidic link, but the evidence appears to be strongly in favour of α-. Thus the structure of melibiose is (β-form) [see also §7g]:



Melibiose has been synthesised chemically.

§19. Gentiobiose (6-*O*-β-D-glucopyranosyl-D-glucopyranose). This was originally obtained from the trisaccharide, gentianose (§20), but it also occurs in some glycosides, *e.g.*, amygdalin (§27). Gentiobiose is a reducing sugar, forms an osazone and undergoes mutarotation; hydrolysis with dilute acids produces two molecules of D-glucose. Since this hydrolysis is also effected by emulsin, the glycosidic link must be β-. Methylation, followed by hydrolysis, gives 2 : 3 : 4-trimethyl-D-glucose and 2 : 3 : 4 : 6-tetramethyl-D-glucose. Oxidation to gentiobionic acid, this then methylated and followed by hydrolysis, gives 2 : 3 : 4 : 5-tetramethyl-D-gluconic acid and 2 : 3 : 4 : 6-tetramethyl-D-glucose. Thus gentiobiose is (β-form):



Gentiobiose has been synthesised chemically.

Another disaccharide containing the 1 : 6-glycosidic link is primeverose (§26).

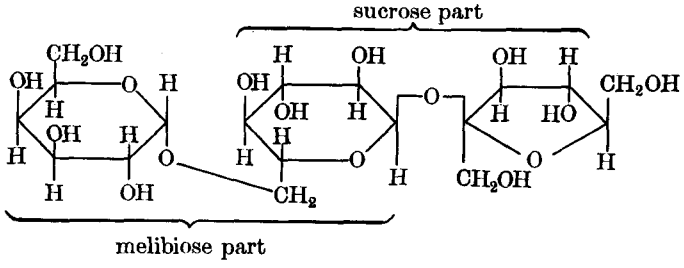
§20. Trisaccharides. The trihexose trisaccharides have the molecular formula $C_{18}H_{32}O_{16}$. They are of two types, reducing and non-reducing.

Manninotriose is the only reducing trisaccharide that has been isolated from natural sources. All the others of this group have been obtained by degrading polysaccharides or by synthesis, *e.g.*, **cellotriose** from cellulose. Two important non-reducing trisaccharides are raffinose and gentianose.

Raffinose occurs in many plants, particularly beet. Controlled hydrolysis with dilute acids gives D-fructose and melibiose; vigorous hydrolysis gives D-fructose, D-glucose and D-galactose. It is also hydrolysed by the enzyme invertase to fructose and melibiose, and by an α-glycosidase to galactose and sucrose. These facts show that the three monosaccharide molecules are linked in the following order:

galactose—glucose—fructose

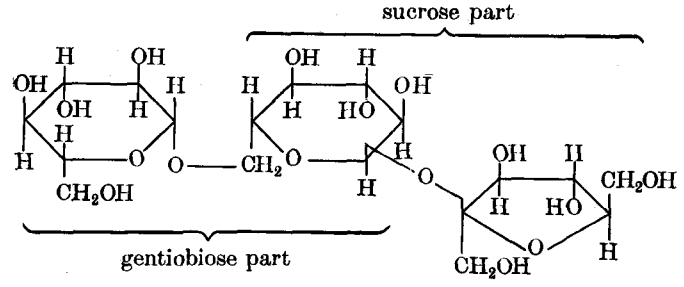
This arrangement is confirmed by the products obtained by methylation of raffinose, followed by hydrolysis, *viz.*, 2 : 3 : 4 : 6-tetramethylgalactose, 2 : 3 : 4-trimethylglucose and 1 : 3 : 4 : 6-tetramethylfructose. Furthermore, since the structures of sucrose (§13) and melibiose (§18) are known, the structure of raffinose must therefore be:



Gentianose occurs in gentian roots. Controlled hydrolysis with dilute acids gives D-fructose and gentiobiose; this hydrolysis is also effected by the enzyme invertase. Emulsin also hydrolyses gentianose to D-glucose and sucrose. Thus the arrangement of the three monosaccharide molecules is:

glucose—glucose—fructose

Hence the structure of gentianose is:

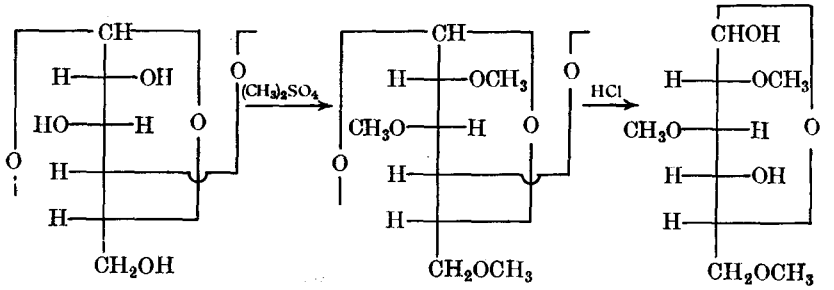


POLYSACCHARIDES

Polysaccharides are high polymers of the monosaccharides, and may be roughly divided into two groups: those which serve as "structures" in plants and animals, *e.g.*, cellulose, and those which act as a metabolic reserve in plants and animals, *e.g.*, starch.

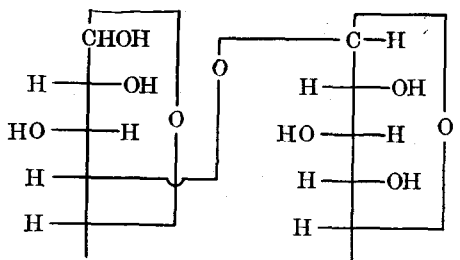
§21. Cellulose. The molecular formula of cellulose is $(C_6H_{10}O_5)_n$. When hydrolysed with fuming hydrochloric acid, cellulose gives D-glucose in 95–96

per cent. yield (Irvine *et al.*, 1922); therefore the structure of cellulose is based on the D-glucose unit. Methylation, acetylation, or " nitration " of cellulose produces a trisubstitution product as a maximum substitution product, and it therefore follows from this that each glucose unit present has *three* hydroxyl groups in an uncombined state. When fully methylated cellulose is hydrolysed, the main product is 2 : 3 : 6-tri-*O*-methyl-D-glucose (90 per cent.). Thus the three free hydroxyl groups in each glucose unit must be in the 2, 3 and 6 positions, and positions 4 and 5 are therefore occupied. Now, if we assume that the ring structure is present in each unit, then this would account for position 5 (or alternatively, 4) being occupied. Furthermore, if we also assume that the glucose units are linked by C_1 of one unit to C_4 of the next (or alternatively, C_5), then the following tentative structure for cellulose would account for the facts:



It should be noted, however, that if the linkages at 4 and 5 were interchanged, the same trimethylglucose would still be obtained on hydrolysis (*cf.* maltose, etc.).

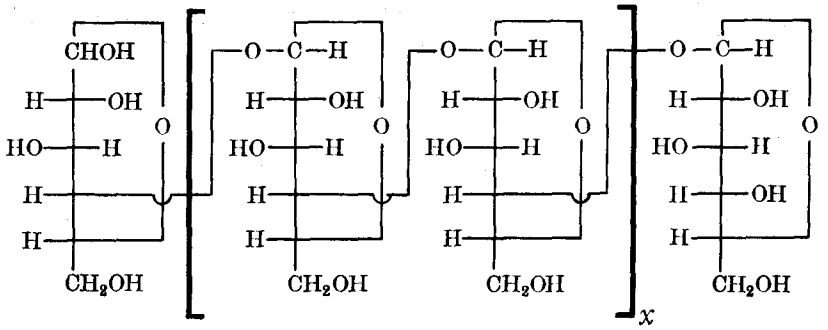
When subjected to acetolysis, *i.e.*, simultaneous acetylation and hydrolysis (this is carried out with a mixture of acetic anhydride and concentrated sulphuric acid), cellulose forms cellobiose octa-acetate. Thus the cellobiose unit is present in cellulose, and since the structure of cellobiose is known (see §16), it therefore follows that the glucose units are present in the pyranose form, *i.e.*, C_5 is involved in ring formation, and so the glucose units are linked C_1—C_4. The isolation of cellobiose indicates also that *pairs* of glucose units are joined by β-links, but it does indicate whether the links between the



glucose units are the same (all β-) or alternate (α and β), since all the links could be β-, or each pair of cellobiose units could be joined by α-links; the latter possibility is not likely, but it is not definitely excluded. Very careful acetolysis of cellulose, however, has produced a cellotriose, cellotetraose and

a cellopentaose, and in all of these the C_1—C_4 links have been shown to be β- (from calculations of the optical rotations), and so we may conclude that *all* the links in cellulose are β-. This conclusion is supported by other evidence, *e.g.*, the kinetics of hydrolysis of cellulose.

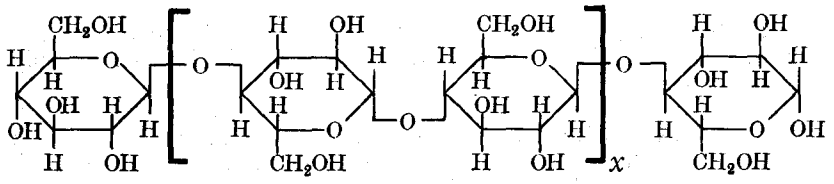
Cellulose forms colloidal solutions in solvents in which glucose is soluble, and so it is inferred that cellulose is a very large molecule. Moreover, since cellulose forms fibres, *e.g.*, rayon, it appears likely that the molecule is linear; X-ray analysis also indicates the linear nature of the molecule, and that the cellulose molecule has a long length. Hence a possible structure for cellulose is:



cellobiose units

I*a*

or



I*b*

It should be noted that in the structure given for cellulose, the first glucose unit in I*a* (*i.e.*, the one on the left-hand side; this unit is on the right-hand side in I*b*) has a free reducing group, but since this group is at the end of a very long chain, its properties tend to be masked; thus cellulose does not exhibit the strong reducing properties of the sugars.

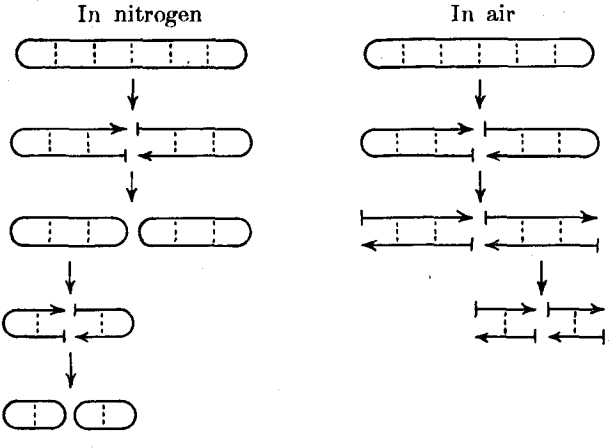
The cellulose molecule is not planar, but has a screw-axis, each glucose unit being at right angles to the previous one. Although free rotation about the C—O—C link might appear possible at first sight, it apparently does not occur owing to the steric effect. This and the close packing of the atoms give rise to a rigid chain molecule. The long chains are held together by hydrogen bonding, and thus cellulose has a three-dimensional brickwork. This would produce strong fibres with great rigidity but no flexibility, and consequently, although the fibres would have great tensile strength, they could not be knotted without snapping. Since the fibres can be knotted without snapping, they must possess flexibility, and the presence of the latter appears to be due to the partly amorphous character of cellulose.

The chemical structure of cellulose appears to be more complicated than the one given above. Schmidt *et al.* (1932) showed that carboxyl groups are present in carefully purified cotton fibres. Kleinert *et al.* (1944) have suggested that various other groups, which are not necessarily carbohydrate in

nature, may bind the glucose chains together. It should be remembered, in this connection, that 100 per cent. glucose is never obtained from the hydrolysis of cellulose.

The molecular weight of cellulose. Owing to its insolubility, simple methods of molecular weight determination (depression of freezing point and elevation of boiling point) cannot be applied to cellulose.

Chemical methods. Examination of the formula of cellulose shows that on methylation, followed by hydrolysis, the end unit (the non-reducing end) would contain four methoxyl groups, and all the other units three. Hence, by the determination of the percentage of the tetramethyl derivative (2 : 3 : 4 : 6-) it is therefore possible to estimate the length of the chain. Haworth (1932) separated the methylated glucoses by vacuum distillation; Hibbert (1942) used fractional distillation; Bell (1944), using silica, and Jones (1944), using alumina, effected separation by means of chromatography. The value for the molecular weight of cellulose was found to be between 20,000 and 40,000 (Haworth, 1932); this corresponds approximately to 100 to 200 glucose units. This " end-group assay ", however, gives rise to the following difficulty. When cellulose is very carefully prepared from cotton, and then methylated in an atmosphere of nitrogen, *i.e.*, in the *absence* of oxygen, no 2 : 3 : 4 : 6-tetramethylglucose was obtained after hydrolysis (Haworth *et al.*, 1939). One explanation that has been offered is that during methylation under ordinary conditions, *i.e.*, in air, cellulose is partially degraded, *e.g.*, osmotic pressure measurements carried out on methylated cellulose, produced by two methylations in air, gave a value of 190 glucose units; sixteen successive methylations in air gave a methylated cellulose of 45 glucose units, as estimated by osmotic pressure measurements (Haworth *et al.*, 1939). Haworth explained these results by suggesting that the cellulose molecule consists of a very large loop, which undergoes progressive shortening on methylation. When the methylation is carried out in an atmosphere of nitrogen, the exposed ends of the shortened loop recombine, but cannot do so when methylated in the presence of air. Haworth also suggested that in order that the two chains should be held parallel to form a loop, it is necessary to have cross-linkages holding the two sides together. The nature of these suggested cross-links is unknown. If primary valencies were involved, then some *dimethylglucose* should be obtained from the hydrolysate. Some of this compound has in fact been isolated, but it is not certain that it is actually present in the methylated cellulose, since it may arise by demethylation during the degradation of the methylated cellulose. The following is a pictorial representation of Haworth's suggestion.



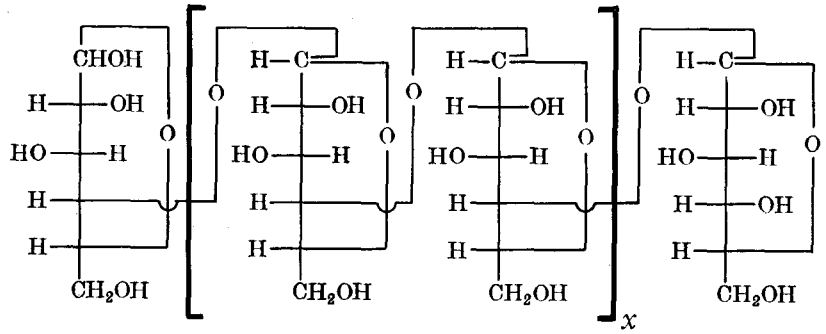
By means of chromatography, McGilvray (1953) has detected 2 : 3 : 4 : 6-tetra-*O*-methyl-D-glucose in the hydrolysate after the methylation of cellulose *in an atmosphere of nitrogen*. Thus degradation of the chain has occurred under these conditions, and so there is no evidence for the linking of the end groups in the absence of oxygen. Furthermore, McGilvray determined the degree of polymerisation from viscosity and osmotic pressure measurements, and also from the end-group assay. The values obtained from the first two methods were greater than that obtained from the third method, and McGilvray suggests these results may be accounted for by assuming a slightly branched structure for the soluble methylcelluloses.

A number of other chemical methods have been used for estimating the molecular weight of cellulose, *e.g.*, that of Hirst *et al.* (1945); this is based on the periodate oxidation (§7g). Examination of the formula of cellulose shows that the terminal reducing unit would give two molecules of formic acid and one of formaldehyde (this reducing unit, which is left in I*a*, behaves as the open-chain molecule, since it is *not* a glycoside), whereas the other terminal unit (right in I*a*) would give one molecule of formic acid; *i.e.*, one cellulose molecule gives three molecules of formic acid and one of formaldehyde. Estimation of the formic acid produced gives the value of the chain-length as approximately 1000 glucose units. There appears, however, to be some uncertainty with these results, since " over-oxidation " as well as normal oxidation with periodic acid results, the former possibly being due to the progressive attack on the chain-molecules from their reducing ends (Head, 1953).

Physical methods. Ultracentrifuge measurements have given a value of 3600 glucose units for *native* cellulose; lower values were obtained for purified cellulose and its derivatives (Kraemer, 1935). These differences are probably due to the degradation of the chains during the process of purification and preparation of the derivatives. Viscosity measurements on cellulose in Schweitzer's solution give a value of 2000–3000 glucose units; lower values were obtained for viscosity measurements on derivatives of cellulose in organic solvents (Staudinger *et al.*, 1935–1937). Osmotic pressure measurements on derivatives of cellulose have given values of approximately 1000 glucose units (Meyer, 1939). Schulz *et al.* (1954, 1958) have determined the molecular weight of cellulose nitrate by measurements of viscosity, etc., and obtained results varying from 1400 to 7800 glucose units, the value depending on the source of the cellulose.

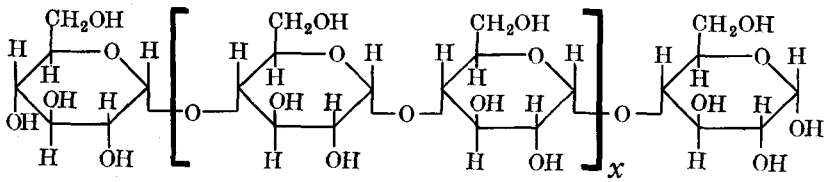
From the foregoing account, it can be seen that the values obtained chemically and physically are not in agreement. This indicates the uncertainty of the value of n, and also that the value of n depends on the source and treatment of cellulose. However, the more recent work of Schulz (see above) is reliable in that evidence was obtained that no degradation occurred in the course of purification and conversion into the nitrates.

§22. Starch. The molecular formula of starch is $(C_6H_{10}O_5)_n$. Hydrolysis of starch with acids produces a quantitative yield of D-glucose (*cf.* cellulose); thus the structure of starch is based on the glucose unit. Methylation of starch gives the trimethylated compound (maximum substitution), and this, on hydrolysis, produces 2 : 3 : 6-tri-*O*-methyl-D-glucose as the main product, and a small amount (about 4·5 per cent.) of 2 : 3 : 4 : 6-tetra-*O*-methyl-D-glucose. Oxidation studies (periodic acid) have also shown the presence of 1 : 4-linked D-glucopyranose residues. Starch is hydrolysed by the enzyme diastase (β-amylase) to maltose (see also below). Thus the maltose unit is present in starch, and so we may conclude that all the glucose units are joined by α-links (*cf.* cellulose). The following structure for starch fits these facts:



maltose units

or



The Haworth end-group assay (1932) showed that starch is composed of approximately 24–30 glucose units. Thus starch is a linear molecule, at least as far as 24–30 units. Haworth, however, pointed out that this was a *minimum* chain-length, and that starches may differ by having different numbers of this repeating unit (see also below). Viscosity measurements, however, showed the presence of a highly branched structure. Now, it has long been known that starch can be separated into two fractions, but it is only fairly recently that this separation has been satisfactorily carried out; the two fractions are α-amylose (the A-fraction; 17–34 per cent.) and β-amylose (amylopectin, or the B-fraction). The fractionation has been carried out in several ways, *e.g.*, *n*-butanol is added to a hot colloidal solution (aqueous) of starch, and the mixture allowed to cool to room temperature. The A-fraction is precipitated, and the B-fraction is obtained from the mother liquors by the addition of methanol (Schoch, 1942). Haworth *et al.* (1946) have used thymol to bring about selective precipitation.

α-Amylose is soluble in water, and the solution gives a blue colour with iodine. β-Amylose is insoluble in water, and gives a violet colour with iodine. Both amyloses are mixtures of polymers, and the average molecular weight depends on the method of preparation of the starch used.

α-Amylose (A-fraction). Meyer *et al.* (1940) measured the osmotic pressure of solutions of α-amylose acetate, and obtained values of 10,000–60,000 for the molecular weight; values up to 1,000,000 have been reported. When α-amylose with a chain-length of about 300 glucose units (as shown by osmotic pressure measurements) was methylated and then hydrolysed, about 0·3 per cent. of 2 : 3 : 4 : 6-tetra-*O*-methyl-D-glucose was obtained. This value is to be expected from a straight chain composed of approximately 300 glucose units. From this evidence it would therefore appear that α-amylose is a *linear* polymer, and this is supported by the early work with soya-bean β-amylase (diastase). This enzyme converts α-amylose into maltose in about 100 per cent. yield; this indicates that a large number of maltose units are joined by α-links, *i.e.*, α-amylose is a linear molecule. Peat *et al.* (1952), however, showed that highly purified soya-bean β-amylase

gives only about 70 per cent. of maltose, and this has been confirmed by other workers. Since β-amylase only attacks α-1 : 4-glucosidic linkages, it thus appears that α-amylose contains a small number of other linkages. Careful purification of " crude " soya-bean β-amylase showed the presence of two enzymes, β-amylase and another which was named Z-enzyme; it is the latter which was shown to hydrolyse the non α-1 : 4-linkages. Thus unpurified β-amylase (which contains both enzymes) degrades α-amylose completely to maltose. It has also been shown that Z-enzyme has β-glucosidase activity and that emulsin can hydrolyse these " anomalous " linkages. These observations suggest that α-amylose contains a small number of β-glucosidic linkages.

Another difficulty arises from the fact that the structure of potato amylose depends on its method of preparation, *e.g.*, one sample is completely degraded by purified β-amylase, whereas other samples are not. The first sample represents about 40 per cent. (by weight) of the total amylose in potato starch, and thus it follows that potato amylose is heterogeneous both in structure and in size. A large proportion is completely linear (and contains about 2000 glucose units), and the remainder (which contains about 6000 units) contains a small number of these anomalous linkages. The nature of these anomalous linkages is still uncertain.

Amylopectin (B-fraction). Molecular weight determinations of amylopectin by means of osmotic pressure measurements indicate values of 50,000 to 1,000,000 (Meyer *et al.*, 1940). Larger values have also been reported, *e.g.*, Witnauer *et al.* (1952) have determined the molecular weight of potato amylopectin by the method of light scattering, and report an average value of 10,000,000 or more. Let us consider an amylopectin having an average molecular weight of 550,000; this corresponds to about 3000 glucose units. The end-group assay by methylation shows the presence of one unit with four free hydroxyl groups per 24–30 glucose units; the same results are also obtained by the periodate method. Thus the 3000 units are joined in such a manner as to give about 100 end units; it therefore follows that the chain must be *branched*. The problem is further complicated by the fact that Hirst (1940), after methylating amylopectin and hydrolysing the product, obtained, in addition to tri- and tetra-*O*-methyl-D-glucose, about 3 per cent. of 2 : 3-di-*O*-methyl-D-glucose. This has been taken to mean that some glucose units are also joined by C_1 and C_6 atoms. Furthermore, in certain experiments, enzymic hydrolysis has given a small amount of 1 : 6 α-linked diglucose, *i.e.*, *iso*maltose is also present in amylopectin (Montgomery *et al.*, 1947, 1949). Wolfrom *et al.* (1955, 1956) have obtained evidence that there is also an α-D-1 : 3-bond in amylopectin; the principal bond is α-D-1 : 4, and branching occurs through α-D-1 : 6-bonds.

The branching of the chains in amylopectin is supported by the following evidence:

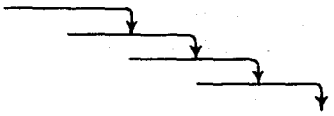
(i) Amylopectin acetate does not form fibres; fibre formation is characteristic of *linear* molecules.

(ii) β-Amylase hydrolyses amylopectin to give only about 50 per cent. of maltose. Thus there are " blocked " points, and these will occur at the branch points.

(iii) Amylopectin solutions do not show an orientation of the molecules in the direction of flow in the concentric cylinder technique; the molecules are therefore not linear.

The detailed structure of amylopectin is still not settled. Haworth and Hirst (1937) suggested a laminated formula for amylopectin; each line represents a basal chain of 24–30 glucose units joined by α C_1—C_4 links, and each arrow head represents the joining of the terminal reducing group (C_1) of each chain to the central glucose member (at C_6) of the next chain.

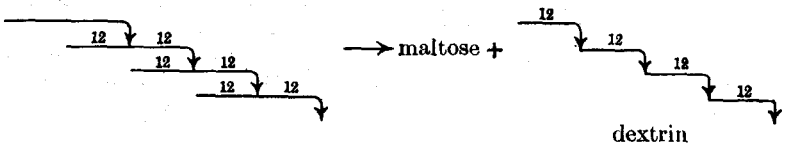
If it is branched in the fashion shown, then methylated amylopectin should give some dimethylglucose on hydrolysis. Since 2 : 3-di-*O*-methyl-D-glucose is actually obtained, the link must be C_1 of one chain to C_6 of the next. If



the unions are as regular as this, then there will be one C_1—C_6 link for every one end group. Hirst *et al.* (1945), however, showed by the end-group assay by the periodic acid method that amylopectin contains only traces of glucose residues joined solely by C_1—C_6 links.

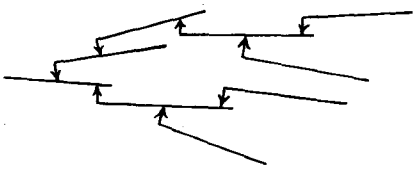
Prolonged methylation of amylopectin produces a diminution of the molecular size (as determined by physical methods); *e.g.*, methylation of starch seventeen times changed the particle size from 3000 glucose units to 190 units (Averill, 1939). This diminution in particle size cannot be due to the break-down of the basal chains, since the end-group assay always gives the same basal chain-length, whether the methylation is carried out in air or in an atmosphere of nitrogen. Haworth therefore suggested that this diminution in particle size is due to the "disaggregation" of the basal chains.

As pointed out previously, β-amylase gives only 50 per cent. of maltose with amylopectin. The high molecular weight residue is known as **dextrin,** and this is not degraded because of the presence of the C_1—C_6 link (β-amylase breaks only α C_1—C_4 links). According to Haworth (1946), β-amylase attacks the chain, breaking them into units of two, the attack stopping at the cross-links. Thus:



In support of this explanation, it has been found that dextrin has a unit chain-length of 11–12 glucose units.

Further work has shown that the Haworth laminated formula does not satisfy the behaviour of amylases on amylopectin; the formula is far too regular (*cf.* Hirst's work, above). Meyer (1940) proposed a highly branched structure; this fits the behaviour of the amylases better. Furthermore,



mathematical calculations have shown that the regular form is unlikely. A difficulty of the Meyer structure, however, is that amylopectin would be globular; this is not in keeping with all the evidence.

§23. Some other polysaccharides. A number of other polysaccharides besides cellulose and starch also occur naturally, and some of these are described briefly below.

Glycogen. This is the principal reserve carbohydrate in animals. It is

hydrolysed by β-amylase to maltose, and molecular weight determinations by physical methods give values between 1 and 2,000,000. The molecular structure of glycogen appears to be similar to that of amylopectin; both polysaccharides have many features in common. One main difference is their degree of branching, the average chain-length in amylopectin being about 24 glucose units and in glycogen about 12.

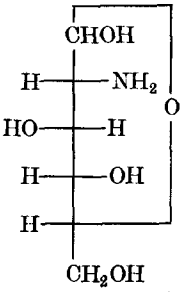
Inulin. This is a fructosan, and occurs in dahlia tubers, dandelion roots, etc. Acid hydrolysis gives D-fructose, but if inulin is first methylated and then hydrolysed, 3 : 4 : 6-tri-*O*-methyl-D-fructose is the main product, thus indicating that inulin is composed of fructofuranose units.

Mannans are polysaccharides which yield only mannose on hydrolysis; they are found in ivory nut, seaweeds, bakers' yeast, etc. Similarly, **galactans** yield only galactose on hydrolysis; they occur in seeds, wood, etc. There are also polysaccharides which contain pentose residues only, *viz.* **pentosans,** *e.g.*, **xylans** give D-xylose; *arabans* give L-arabinose. Some pentosans are composed of both xylose and arabinose, and other polysaccharides are composed of pentose and hexose units, *e.g.*, *xylo-glucans* (xylose and glucose), *arabo-galactans*, etc. In addition to these neutral polysaccharides, there are also the acid polysaccharides. These are gums and mucilages, and owe their acidity to the presence of uronic acids. Gums are substances which swell in water to form gels (or viscous solutions), *e.g.*, gum arabic and gum tragacanth; on hydrolysis, the former gives arabinose, galactose, rhamnose and glucuronic acid, and the latter xylose, L-fructose and galacturonic acid. Mucilages are polysaccharides which swell in water to form viscous solutions; on hydrolysis, they give galacturonic acid, arabinose, xylose, etc. The **hemi-celluloses** (which are widely distributed in the cell-wall of plants) also contain both uronic acids (glucuronic or galacturonic) and pentoses (xylose, arabinose).

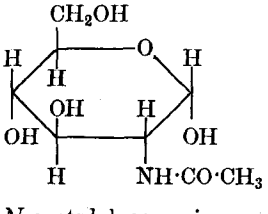
Pectin. This occurs in plants, particularly fruit juices. It is composed of D-galacturonic acid residues and the methyl ester.

Alginic acid. This occurs in the free state and as the calcium salt in various seaweeds. Hydrolysis of alginic acid produces D-mannuronic acid.

Chitin. This is the polysaccharide that is found in the shells of crustaceans. Hydrolysis of chitin by acids produces acetic acid and D-glucosamine (chitosamine; 2-aminoglucose). Chitin is also hydrolysed by an enzyme (which occurs in the intestine of snails) to *N*-acetylglucosamine. X-ray analysis has shown that the structure of chitin is similar to that of cellulose (*N*-acetylglucosamine replaces glucose).



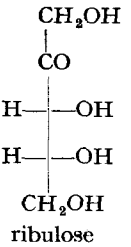
D-glucosamine



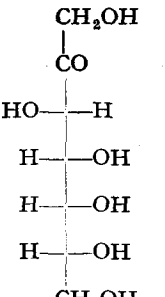
N-acetylglucosamine

N-Methyl-L-glucosamine is a component of streptomycin (see §7. XVIII).

§23a. **Photosynthesis of carbohydrates.** The scheme outlined below is largely that proposed by Calvin *et al.* (1954). These authors exposed certain algæ to carbon dioxide (labelled with ^{14}C) and light, then killed the



ribulose



sedoheptulose

algæ and extracted with ethanol and chromatographed (on paper) the extract. Two monosaccharides, ribulose and sedoheptulose, play an essential part in the photosynthesis of carbohydrates, and the steps involved are as follows :

(i) Ribulose diphosphate accepts one molecule of carbon dioxide and one of water.

(ii) The product now splits into two molecules of phosphoglyceric acid ($CH_2O{\cdot}PO_3H_2{\cdot}CHOH{\cdot}CO_2H$).

(iii) Phosphoglyceric acid undergoes reduction to phosphoglyceraldehyde.

(iv) Two molecules of phosphoglyceraldehyde combine to form hexose phosphate.

(v*a*) Hexose phosphate forms disaccharides and polysaccharides.

(v*b*) A molecule of hexose phosphate reacts with a molecule of phosphoglyceraldehyde to form ribulose phosphate and a tetrose phosphate. The latter reacts with a molecule of phosphoglyceraldehyde to produce sedoheptulose phosphate which, in turn, also reacts with a molecule of phosphoglyceraldehyde to produce one molecule of ribose phosphate and one molecule of ribulose phosphate. The ribose phosphate is then converted into ribulose phosphate, thus completing the cycle.

All the aldohexoses and all the aldopentoses are interconvertible by inversion of one asymmetric carbon atom, but how this occurs in the plant is not certain. Furthermore, aldohexoses may be stepped down to aldopentoses, and again how this occurs is not certain; one suggestion is (see also §32a. VIII):

$$CHO{\cdot}(CHOH)_3{\cdot}CHOH{\cdot}CH_2OH \xrightarrow{\text{oxidation}} CHO{\cdot}(CHOH)_3{\cdot}CHOH{\cdot}CO_2H \xrightarrow{\text{decarboxylation}} CHO{\cdot}(CHOH)_3{\cdot}CH_2OH$$

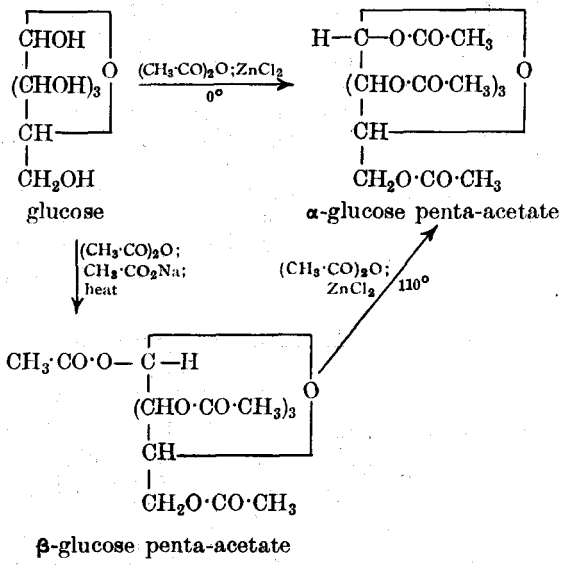
The foregoing account of photosynthesis describes the various intermediates produced. In green plants the presence of chlorophyll (§6. XIX) is necessary for photosynthesis, but its exact function is not certain. It appears that all the light energy is used in the " light phase " to raise chlorophyll *a* from its ground state to an excited state, and then this energy of the excited state is used in the " dark phase " to convert carbon dioxide into carbohydrates (Trebst *et al.*, 1958–1960). Furthermore, the same series of dark-phase reactions has also been shown to occur in non-chlorophyllous cells (*inter alia*, McFadden *et al.*, 1957, 1959). What is peculiar to photosynthesis is its light phase.

GLYCOSIDES

§24. Introduction. A great variety of glycosides occur in plants. The simple glycosides are colourless, soluble in water and are optically active; they do not reduce Fehling's solution. On hydrolysis with inorganic acids, glycosides give a sugar and a hydroxylic compound, the aglycon (§3), which may be an alcohol or a phenol. Most glycosides are hydrolysed by emulsin; therefore they are β-glycosides. Actually, in the natural state, each glycoside is usually associated with an enzyme which occurs in different cells of the plant. Maceration of the plant thus produces hydrolysis of the glycoside by bringing the enzyme in contact with the glycoside. Glucose has been found to be the most common sugar component; when methylated and then hydrolysed, most glycosides give 2 : 3 : 4 : 6-tetra-*O*-methyl-D-glucose. Thus most glycosides are β-D-glucopyranosides.

Synthesis of glycosides. The synthesis of a glycoside uses an **acetobromohexose** as the starting material; this compound is now named systematically as a tetra-*O*-acetyl-D-hexopyranosyl 1-bromide, *e.g.*, if the hexose is glucose, then the α-form will be tetra-*O*-acetyl-α-D-glucopyranosyl 1-bromide.

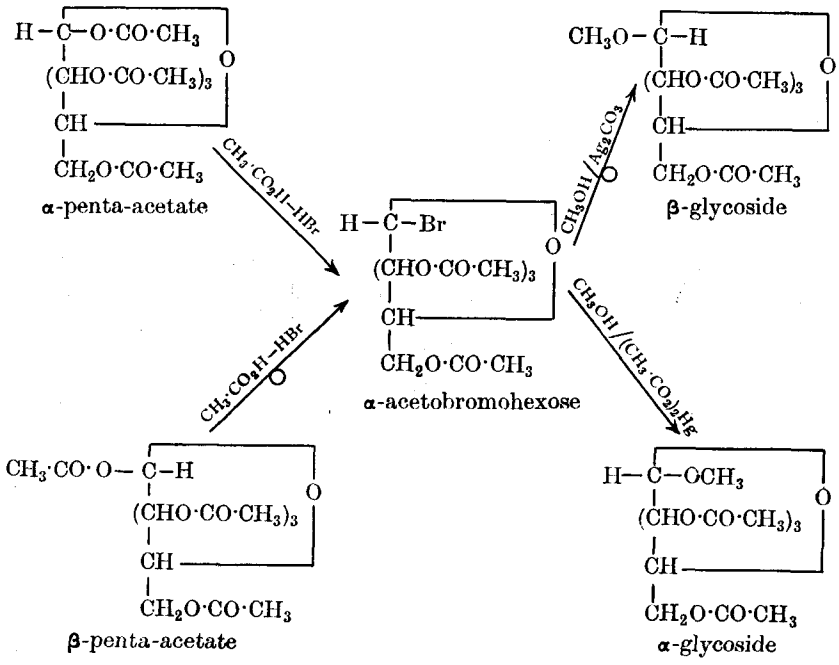
When glucose is treated with acetic anhydride at 0° in the presence of zinc chloride, the product is 1 : 2 : 3 : 4 : 6-penta-*O*-acetyl-α-D-glucose (α-D-glucose penta-acetate). If, however, glucose is heated with acetic anhydride in the presence of sodium acetate, the product is 1 : 2 : 3 : 4 : 6-penta-*O*-acetyl-β-D-glucose. Furthermore, the β-isomer may be converted into the α- by heating with acetic anhydride at 110° in the presence of zinc chloride.



These penta-acetates are readily hydrolysed to glucose by means of dilute aqueous sodium hydroxide, ethanolic ammonia at 0°, or by methanol containing a small amount of sodium methoxide. When dissolved in glacial acetic acid saturated with hydrogen bromide, the glycosidic acetoxyl group of a hexose penta-acetate is replaced by bromine to give an α-acetobromohexose; the α-isomer is obtained whether the penta-acetate used is the α- or β-compound (Fischer, 1911). Thus a Walden inversion occurs with the β-compound (§1. III).

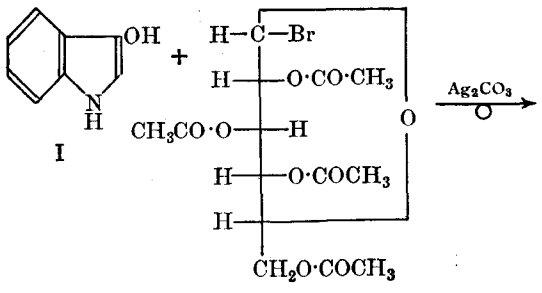
Scheurer *et al.* (1954) have synthesised acetobromo sugars in good yield as follows. Bromine is added to a suspension of red phosphorus in glacial acetic acid, and to this solution (which now contains acetyl bromide) is added the sugar or acetylated sugar, the latter giving the better yields.

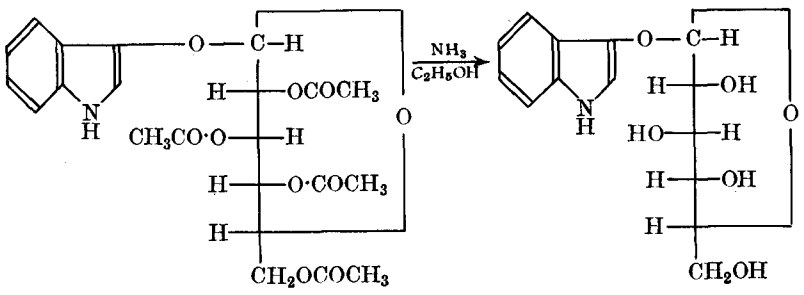
The bromine atom in these acetobromohexoses is very active. Thus it may be replaced by a hydroxyl group when the acetobromohexose is treated with silver carbonate in moist ether (Fischer *et al.*, 1909), or by an alkoxyl group when treated with an alcohol in the presence of silver carbonate (Königs and Knorr, 1901). In the latter reaction the yields are improved if anhydrous calcium sulphate and a small amount of iodine are used instead of silver carbonate (Evans *et al.*, 1938). In either case, the α-acetobromohexose gives the β-glycoside. On the other hand, if mercuric acetate is used instead of silver carbonate, then the α-glycoside is obtained. The foregoing reactions may thus be written (using the symbol $\xrightarrow{\sigma}$ to represent a Walden inversion; see §3. III).



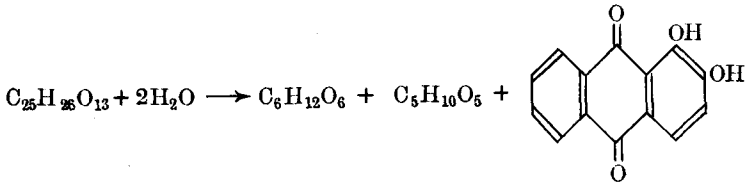
§25. Indican. This glycoside occurs in the leaves of the indigo plant and in the woad plant. When the leaves are macerated with water, the enzyme present hydrolyses indican to glucose and indoxyl, and the latter, on exposure to air, is converted into indigotin (see Vol. I).

The molecular formula of indican is $C_{14}H_{17}O_6N$, and since it gives D-glucose and indoxyl on hydrolysis, it is therefore indoxyl D-glucoside. When indican is methylated (with methyl iodide in the presence of dry silver oxide), tetramethylindican is obtained, and this, on hydrolysis with methanol containing 1 per cent. hydrogen chloride, gives indoxyl and methyl 2 : 3 : 4 : 6-tetra-*O*-methyl-D-glucoside. Thus the glucose molecule is present in the pyranose form, and since indican is hydrolysed by emulsin, the glycosidic link must be β. Thus the structure of indican is III, and this has been confirmed by synthesis from indoxyl, I, and tetra-*O*-acetyl-α-D-glucopyranosyl 1-bromide, II, as follows:

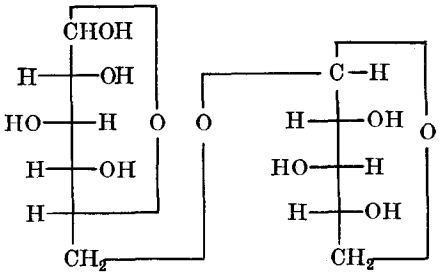




§26. Ruberythric acid. This occurs in the madder root, and on hydrolysis, it was originally believed to give one molecule of alizarin and two molecules of D-glucose. Jones and Robertson (1933), however, showed that two molecules of D-glucose were not present in the hydrolysate; a mixture of two sugars was actually present, D-glucose and D-xylose. Thus the molecular formula of ruberythric acid is $C_{25}H_{26}O_{13}$, and not, as was originally believed, $C_{26}H_{28}O_{14}$. Thus the hydrolysis is:

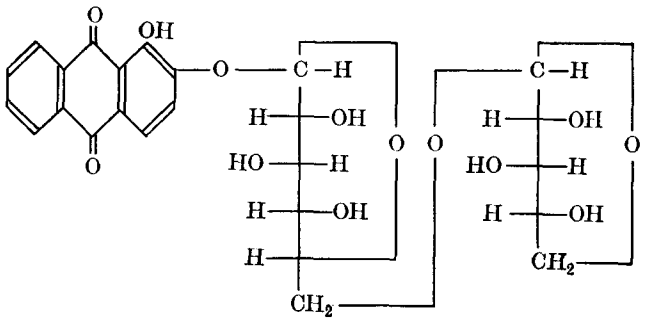


Jones and Robertson also showed that the two monosaccharide molecules were present in the form of the disaccharide **primeverose.** Now, this disaccharide is 6-*O*-β-D-xylopyranosyl-D-glucopyranose (Helferich, 1927), and



primeverose

it therefore follows that alizarin is linked to the glucose half of the primeverose molecule. Further work has shown that the glucosidic link is β, and that it is the 2-hydroxyl group of alizarin that is involved. Thus the structure of ruberythric acid is:



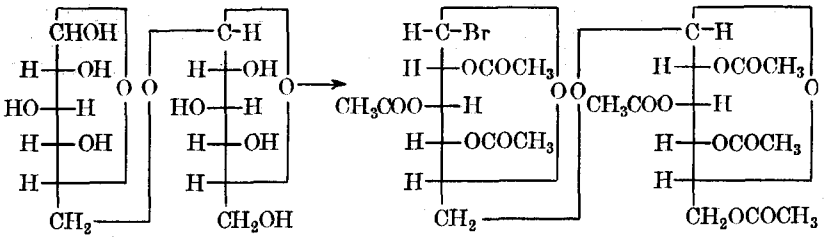
§27. Amygdalin. This occurs in bitter almonds. The molecular formula is $C_{20}H_{27}O_{11}N$, and it is hydrolysed by acids to one molecule of benzaldehyde, two molecules of D-glucose, and one of hydrogen cyanide.

$$C_{20}H_{27}O_{11}N + 2H_2O \longrightarrow C_6H_5{\cdot}CHO + 2C_6H_{12}O_6 + HCN$$

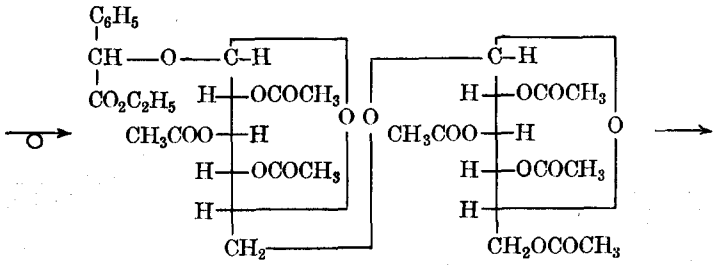
Since emulsin also brings about this hydrolysis, amygdalin must contain a β-glycosidic link. On the other hand, the enzyme zymase hydrolyses amygdalin into one molecule of glucose and a glucoside of (+)-mandelonitrile (this compound is

$$C_{20}H_{27}O_{11}N + H_2O \longrightarrow C_6H_{12}O_6 + C_6H_5{\cdot}CH(CN){\cdot}O{\cdot}C_6H_{11}O_5$$

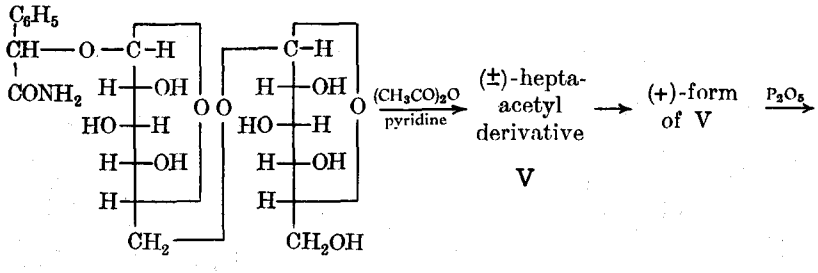
identical with *prunasin*, a naturally occurring glucoside). Thus the aglycon of amygdalin is (+)-mandelonitrile, and the sugar is a disaccharide. Haworth *et al.* (1922, 1923) have shown that this disaccharide is gentiobiose (§19), and have synthesised amygdalin (in 1924) as follows. Gentiobiose, I, was converted into hepta-acetyl-bromogentiobiose, II, by means of acetic anhydride saturated with hydrogen bromide, and then II was condensed with racemic ethyl mandelate in the presence of silver oxide, whereby the β-glycoside, III, was obtained. Treatment of this with ethanolic ammonia hydrolysed the acetyl groups, and at the same time converted the ester group into the corresponding amide; thus the ($\pm$)-amido-glycoside, IV, was obtained. IV was then treated with acetic anhydride in pyridine solution, and the ($\pm$)-hepta-acetyl derivative of the amide, V, was then separated into its diastereoisomers by fractional crystallisation (the mandelic acid portion is + and −, the gentiobiose portion is +; hence the two forms present are ++ and −+, *i.e.*, they are diastereoisomers). The (+)-form was then dehydrated with phosphorus pentoxide to give the (+)-nitrile, VI, and this, on de-acetylation with ethanolic ammonia, gave (+)-amygdalin, VII, which was shown to be identical with the natural compound. (See overleaf.)



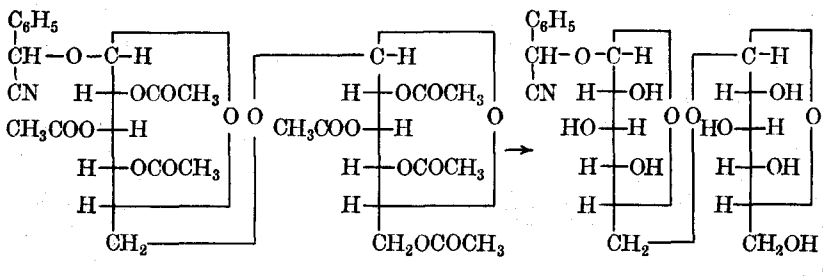
I II



III

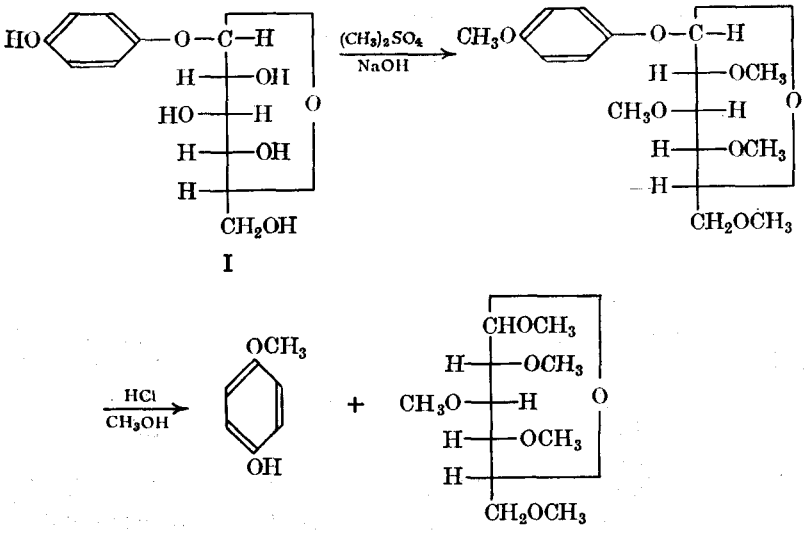


IV V



VI VII

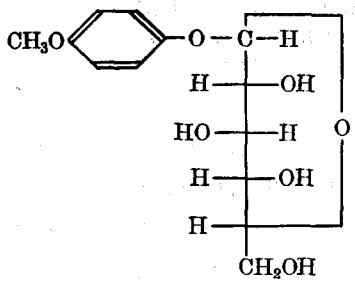
§28. Arbutin and Methylarbutin. **Arbutin** is hydrolysed by emulsin to give one molecule of D-glucose and one of quinol; thus arbutin is a β-glucoside. When methylated (with methyl sulphate in the presence of sodium hydroxide), arbutin forms pentamethylarbutin, and this on hydrolysis with methanolic hydrogen chloride, gives methyl 2 : 3 : 4 : 6-tetra-*O*-methyl-D-glucoside and monomethylquinol (Macbeth *et al.*, 1923); structure I for arbutin accounts for all these facts.



I

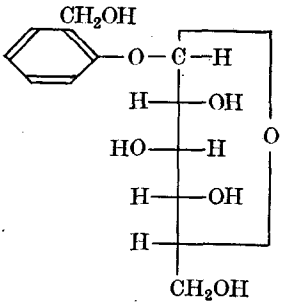
Pentamethylarbutin has been synthesised by converting 2 : 3 : 4 : 6-tetra-*O*-methyl-D-glucose into tetra-*O*-methyl-α-D-glucopyranosyl 1-bromide, and condensing this with monomethylquinol; the product is identical with the methylated natural compound.

Methylarbutin. This is hydrolysed by emulsin to one molecule of D-glucose and one molecule of monomethylquinol; thus methylarbutin is a β-glucoside, and its structure is:



Methylarbutin has been synthesised by condensing tetra-*O*-acetyl-α-D-glucopyranosyl 1-bromide with monomethylquinol in the presence of silver carbonate, followed by de-acetylation.

§29. Salicin. This is hydrolysed by emulsin to one molecule of D-glucose and one of salicyl alcohol (saligenin). Thus salicin is a β-glucoside, but it is not possible to tell from the hydrolytic products whether it is the phenolic or alcoholic group of the salicyl alcohol which forms the glycosidic link. Which group is involved is readily shown as follows (Irvine *et al.*, 1906). Oxidation of salicin with nitric acid forms *helicin*, and this, on hydrolysis, gives glucose and salicylaldehyde. Thus the phenolic group in salicyl alcohol must form the glucoside. Methylation of salicin produces pentamethyl-salicin, and this, on hydrolysis, gives 2 : 3 : 4 : 6-tetra-*O*-methyl-D-glucose. Hence the glucose residue is in the pyranose form; the structure given for salicin fits the foregoing facts. This structure has been confirmed by condensing tetra-*O*-methyl-α-D-glucopyranosyl 1-bromide with salicyl alcohol,

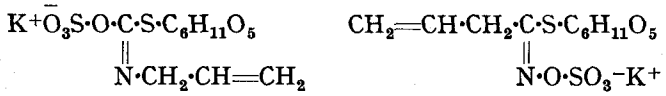


and then methylating the product. The pentamethylsalicin so obtained was identical with the methylated natural product (Irvine *et al.*, 1906).

§**30. Sinigrin.** This glycoside occurs in black mustard seed, and on hydrolysis with the enzyme myrosin, D-glucose, allyl *iso*thiocyanate and potassium hydrogen sulphate are obtained.

$$C_{10}H_{16}O_9NS_2K + H_2O \longrightarrow C_6H_{12}O_6 + CH_2{=}CH{\cdot}CH_2{\cdot}NCS + KHSO_4$$

Sodium methoxide degrades sinigrin, and one of the products obtained is thioglucose, $C_6H_{11}O_5{\cdot}SH$. From this it is inferred that the glucose residue is linked to a sulphur atom in sinigrin. Gadamer (1897) proposed I for the



I II

structure of sinigrin, but Ettlinger *et al.* (1956) have proposed II, since these authors have shown that allyl *iso*thiocyanate is produced by rearrangement when the glycoside is hydrolysed by myrosin (*cf.* the Lossen rearrangement; see Vol. I).

READING REFERENCES

Handbook for Chemical Society Authors, Chemical Society (1960). Ch. 5. Nomenclature of Carbohydrates.
Rosanoff, On Fischer's Classification of Stereoisomers, *J. Amer. Chem. Soc.*, 1906, **28**, 114.
Haworth, *The Constitution of Sugars*, Arnold (1929).
Honeyman, *Chemistry of the Carbohydrates*, Oxford Press (1948).
Percival, *Structural Carbohydrate Chemistry*, Muller (2nd ed., 1962).
Pigman and Goepp, *Chemistry of the Carbohydrates*, Academic Press (1948).
Gilman (Ed.), *Advanced Organic Chemistry*, Wiley. (i) Vol. II (1943, 2nd ed.). Ch. 20, 21. Carbohydrates. Ch. 22. Cellulose. (ii) Vol. IV (1953). Ch. 9. Starch.
Percival, Carbohydrate Sulphates, *Quart. Reviews* (*Chem. Soc.*), 1949, **3**, 369.
Barker and Bourne, Enzymic Synthesis of Polysaccharides, *Quart. Reviews* (*Chem. Soc.*), 1953, **7**, 56.
Hudson, Emil Fischer's Discovery of the Configuration of Glucose, *J. Chem. Educ.*, 1941, **18**, 353.
Advances in Carbohydrate Chemistry, Academic Press (1945–).
Manners, The Enzymic Degradation of Polysaccharides, *Quart. Reviews* (*Chem. Soc.*), 1955, **9**, 73.
Sir Robert Robinson, *The Structural Relationships of Natural Products*, Oxford Press (1955).
Downes, *The Chemistry of Living Cells*, Longmans, Green (2nd ed., 1963).
Newth, Sugar Epoxides, *Quart. Reviews* (*Chem. Soc.*), 1959, **13**, 30.
Ferrier and Overend, Newer Aspects of the Stereochemistry of Carbohydrates, *Quart. Reviews* (*Chem. Soc.*), 1959, **13**, 265.
Sunderwirth and Olson, Conformational Analysis of the Pyranoside Ring, *J. Chem. Educ.*, 1962, **39**, 410.

Manners, Structural Analysis of Polysaccharides, *Roy. Inst. Chem.*, Lectures, Monographs and Reports, 1959, No. 2.
Wiggins, Sugar and its Industrial Applications, *Roy. Inst. Chem.*, Lectures, Monographs and Reports, 1960, No. 5.
Bassham, Photosynthesis, *J. Chem. Educ.*, 1959, **36**, 548.
Park, Advances in Photosynthesis, *J. Chem. Educ.*, 1962, **39**, 424.
Arnon *et al.*, Photoproduction of Hydrogen, Photofixation of Nitrogen and a Unified Concept of Photosynthesis, *Nature*, 1961, **192**, 601.
Roderick, Structural Variety of Natural Products, *J. Chem. Educ.*, 1962, **39**, 2.

CHAPTER VIII

TERPENES

§1. Introduction. The terpenes form a group of compounds the majority of which occur in the plant kingdom; a few terpenes have been obtained from other sources. The simpler mono- and sesqui-terpenes are the chief constituents of the essential oils; these are the volatile oils obtained from the sap and tissues of certain plants and trees. The essential oils have been used in perfumery from the earliest times. The di- and tri-terpenes, which are not steam volatile, are obtained from plant and tree gums and resins. The tetraterpenes form a group of compounds known as the *carotenoids*, and it is usual to treat these as a separate group (see Ch. IX). Rubber is the most important polyterpene.

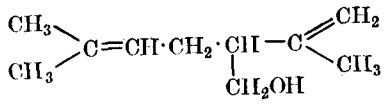
Most natural terpene hydrocarbons have the molecular formula $(C_5H_8)_n$, and the value of n is used as a basis of classification. Thus we have the following classes (these have already been mentioned above):

(i) Monoterpenes, $C_{10}H_{16}$. (ii) Sesquiterpenes, $C_{15}H_{24}$.
(iii) Diterpenes, $C_{20}H_{32}$. (iv) Triterpenes, $C_{30}H_{48}$.
(v) Tetraterpenes, $C_{40}H_{64}$ (these are the carotenoids).
(vi) Polyterpenes, $(C_5H_8)_n$.

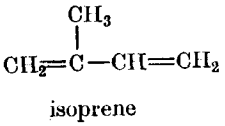
In addition to the terpene hydrocarbons, there are the oxygenated derivatives of each class which also occur naturally, and these are mainly alcohols, aldehydes or ketones.

The term *terpene* was originally reserved for those hydrocarbons of molecular formula $C_{10}H_{16}$, but by common usage, the term now includes all compounds of the formula $(C_5H_8)_n$. There is, however, a tendency to call the whole group *terpenoids* instead of *terpenes*, and to restrict the name *terpene* to the compounds $C_{10}H_{16}$.

The thermal decomposition of almost all terpenes gives isoprene as one of the products, and this led to the suggestion that the skeleton structures of all naturally occurring terpenes can be built up of isoprene units; this is known as the **isoprene rule**, and was first pointed out by Wallach (1887). Thus the divisibility into isoprene units may be regarded as a necessary condition to be satisfied by the structure of any plant-synthesised terpene. Furthermore, Ingold (1925) pointed out that the isoprene units in natural terpenes were joined "head to tail" (the head being the branched end of isoprene). This divisibility into isoprene units, and their head to tail union, may conveniently be referred to as the *special isoprene rule*. It should be noted, however, that this rule, which has proved very useful, can only be used as a guiding principle and not as a fixed rule. Several exceptions to it occur among the simpler terpenes, *e.g.*, lavandulol is composed of two isoprene units which are not joined head to tail; also, the carotenoids are joined tail to tail at their centre (see Ch. IX).

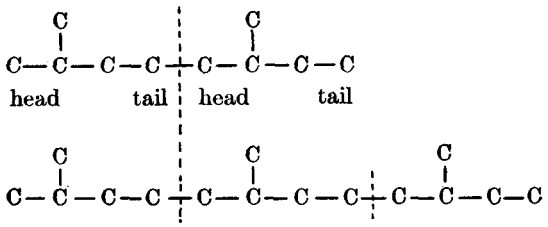


lavandulol

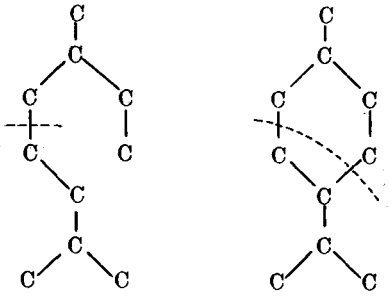


isoprene

The carbon skeletons of open-chain monoterpenes and sesquiterpenes are:



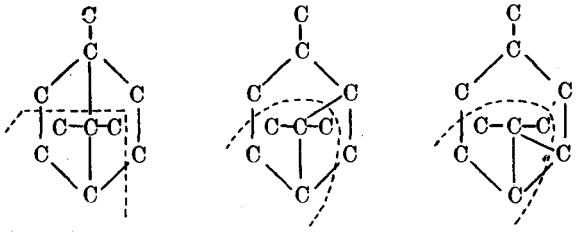
Monocyclic terpenes contain a six-membered ring, and in this connection Ingold (1921) pointed out that a *gem*-dialkyl group tends to render the *cyclo*hexane ring unstable. Hence, in closing the open chain to a *cyclo*hexane ring, use of this "*gem*-dialkyl rule" limits the number of possible structures (but see, *e.g.*, abietic acid, §31). Thus the monoterpene open chain can give rise to only *one* possibility for a monocyclic monoterpene, *viz.*, the *p*-cymene structure. This is shown in the following structures, the acyclic structure being written in the conventional "ring shape".



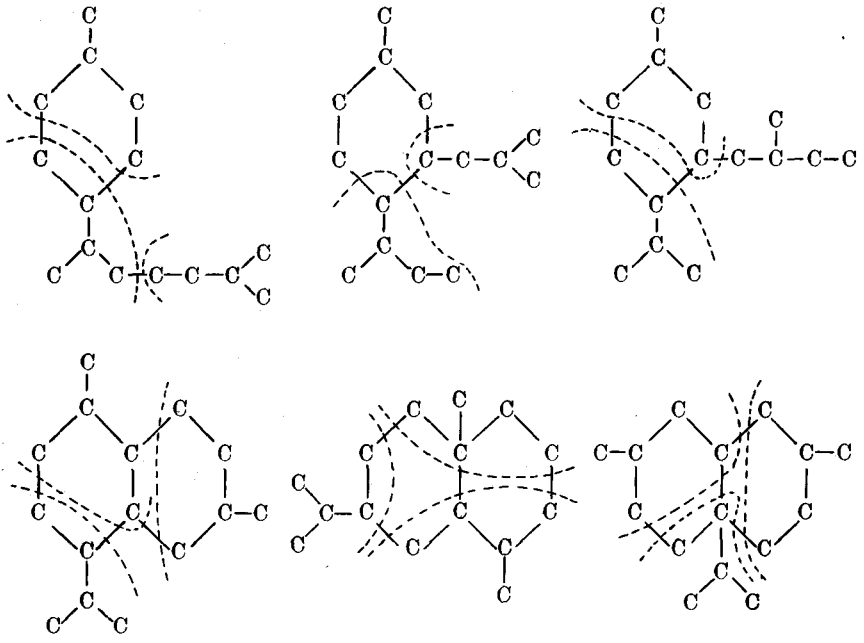
acyclic structure — *p*-cymene structure

All natural monocyclic monoterpenes are derivatives of *p*-cymene.

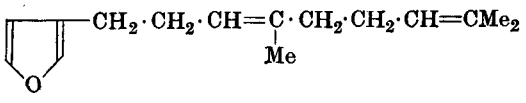
Bicyclic monoterpenes contain a six-membered ring and a three-, four- or five-membered ring. Ingold (1921) also pointed out that *cyclo*propane and *cyclo*butane rings require the introduction of a *gem*-dimethyl group to render them sufficiently stable to be capable of occurrence in nature. Thus closure of the C_{10} open chain gives three possible bicyclic structures; all three types are known.



If we use these ideas with the sesquiterpene acyclic structure, then we find that only three monocyclic and three bicyclic structures are possible (not all are known; see the sesquiterpenes).



Recently some furano-terpenes have been isolated, *e.g.*, dendrolasin, which is believed to have the following structure (Quilico *et al.*, 1957); it contains three isoprene units joined head to tail.



§2. Isolation of monoterpenes and sesquiterpenes. Plants containing essential oils usually have the greatest concentration at some particular time, *e.g.*, jasmine at sunset. In general, there are four methods of extraction of the terpenes: (i) expression; (ii) steam distillation; (iii) extraction by means of volatile solvents; (iv) adsorption in purified fats (*enfleurage*). Method (ii) is the one most widely used; the plant is macerated and then steam distilled. If the compound decomposes under these conditions, it may be extracted with light petrol at 50°, and the solvent then removed by distillation under reduced pressure. Alternatively, the method of adsorption in fats is used. The fat is warmed to about 50°, and then the flower petals are spread on the surface of the fat until the latter is saturated. The fat is now digested with ethanol, any fat that dissolves being removed by cooling to 20°. The essential oils so obtained usually contain a number of terpenes, and these are separated by fractional distillation. The terpene hydrocarbons distil first, and these are followed by the oxygenated derivatives. Distillation of the residue under reduced pressure gives the sesquiterpenes, and these are separated by fractional distillation.

§3. General methods of determining structure. The following brief account gives an indication of the various methods used in elucidating the structures of the terpenes (see the text for details).

(i) A pure specimen is obtained, and the molecular formula is ascertained by the usual methods. If the terpene is optically active, its specific rotation

is measured. Optical activity may be used as a means of distinguishing structures (see, *e.g.*, §12).

(ii) If oxygen is present in the molecule, its functional nature is ascertained, *i.e.*, whether it is present as hydroxyl, aldehyde, ketone, etc. (*cf.* alkaloids, §4. XIV).

(iii) The presence of olefinic bonds is ascertained by means of bromine, and the number of double bonds is determined by analysis of the bromide, or by quantitative hydrogenation, or by titration with monoperphthalic acid. These facts lead to the molecular formula of the parent hydrocarbon, from which the number of rings present in the structure may be deduced.

(iv) The preparation of nitrosochlorides and a study of their behaviour (see also the nitroso compounds, Vol. I).

(v) Dehydrogenation of terpenes with sulphur or selenium, and an examination of the products thereby obtained (see also §2 vii. X).

(vi) Measurement of the refractive index leads to a value for the molecular refractivity. From this may be deduced the nature of the carbon skeleton (see, in particular, sesquiterpenes). Also, optical exaltation indicates the presence of double bonds in conjugation (*cf.* §11. I).

(vii) Measurement of the ultraviolet, infra-red and Raman spectra. More recently X-ray analysis of crystals has also been used.

(viii) Degradative oxidation. The usual reagents used for this purpose are ozone, acid or alkaline permanganate, chromic acid and sodium hypobromite. In general, degradative oxidation is the most powerful tool for elucidating the structures of the terpenes.

(ix) After the analytical evidence has led to a tentative structure (or structures), the final proof of structure depends on synthesis. In terpene chemistry, many of the syntheses are ambiguous, and in such cases analytical evidence is used in conjunction with the synthesis. Many terpenes have not yet been synthesised.

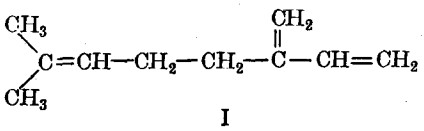
MONOTERPENES

The monoterpenes may be subdivided into three groups: acyclic, monocyclic and bicyclic. This classification affords a convenient means of study of the monoterpenes.

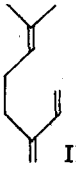
ACYCLIC MONOTERPENES

§4. Myrcene, $C_{10}H_{16}$, is an acyclic monoterpene hydrocarbon which occurs in verbena and bay oils. It is a liquid, b.p. 166–168°. Catalytic hydrogenation (platinum) converts myrcene into a decane, $C_{10}H_{22}$; thus myrcene contains three double bonds, and is an open-chain compound. Furthermore, since myrcene forms an adduct with maleic anhydride, two of the double bonds are conjugated (Diels *et al.*, 1929; see the Diels–Alder reaction, Vol. I). This conjugation is supported by evidence obtained from the ultraviolet spectrum of myrcene (Booker *et al.*, 1940). These facts, *i.e.*, that myrcene contains three double bonds, two of which are in conjugation, had been established by earlier investigators (*e.g.*, Semmler, 1901). Ozonolysis of myrcene produces acetone, formaldehyde and a ketodialdehyde, $C_5H_6O_3$, and the latter, on oxidation with chromic acid, gives succinic acid and carbon dioxide (Ruzicka *et al.*, 1924). These results can be explained by assigning structure I to myrcene. In terpene chemistry it has become customary to use conventional formulæ rather than those of the type I. In these conventional formulæ only lines are used; carbon atoms are at the junctions of pairs of lines or at the end of a line, and unsaturation is indicated by double bonds. Furthermore, the carbon skeleton is usually

drawn in a ring fashion (the *cyclo*hexane ring). Thus myrcene may be represented as II, and this type of structural formula will, in general, be

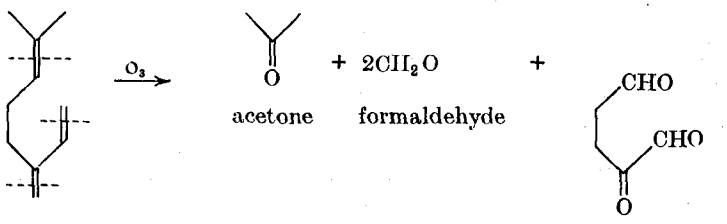


I

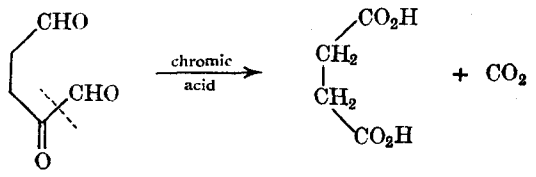


II

used in this book. Thus the process of ozonolysis and oxidation of the ketodialdehyde may be written:

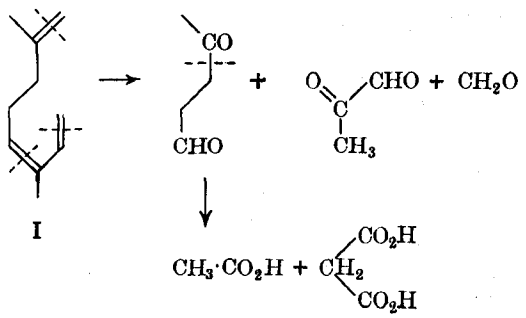


ketodialdehyde

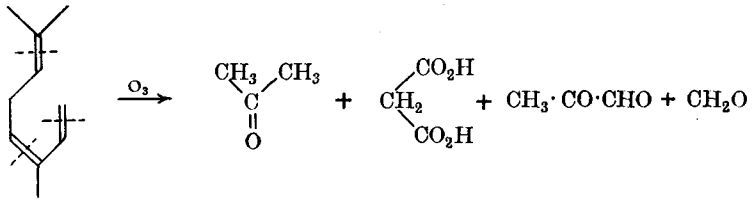


This structure for myrcene is supported by the fact that on hydration (under the influence of sulphuric acid), myrcene forms an alcohol which, on oxidation, gives citral. The structure of this compound is known (see §5), and its formation is in accord with the structure given to myrcene.

§4a. Ocimene, $C_{10}H_{16}$, b.p. 81°/30 mm. When catalytically hydrogenated, ocimene adds on three molecules of hydrogen to form a decane. Thus ocimene is an acyclic compound which contains three double bonds. Furthermore, since ocimene forms an adduct with maleic anhydride, two of the double bonds are conjugated. On ozonolysis, ocimene produces formaldehyde, methylglyoxal, lævulaldehyde, acetic and malonic acids, and some acetone. All of these products, except acetone, are accounted for by structure I for ocimene (this has an *iso*propenyl end-group). In order to account for the appearance of acetone in the oxidation products, ocimene



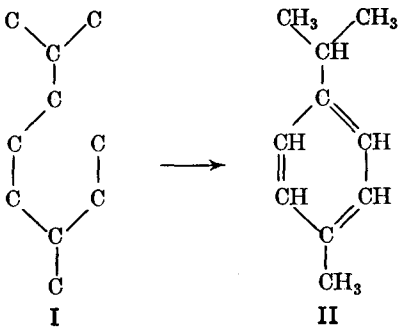
is also believed to exist in the *iso*propylidene form, II, *i.e.*, ocimene is a mixture of I and II, with I predominating (but see citral, §5).



II

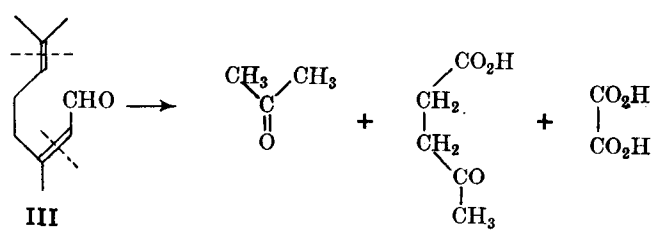
§5. Citral, $C_{10}H_{16}O$. This is the most important member of the acyclic monoterpenes, since the structures of most of the other compounds in this group are based on that of citral. Citral is widely distributed and occurs to an extent of 60–80 per cent. in lemon grass oil. Citral is a liquid which has the smell of lemons.

Citral was shown to contain an oxo group, *e.g.*, it forms an oxime, etc. On heating with potassium hydrogen sulphate, citral forms *p*-cymene, II (Semmler, 1891). This reaction was used by Semmler to determine the positions of the methyl and *iso*propyl groups in citral; Semmler realised that the citral molecule was acyclic, and gave it the skeleton structure, I (two



I II

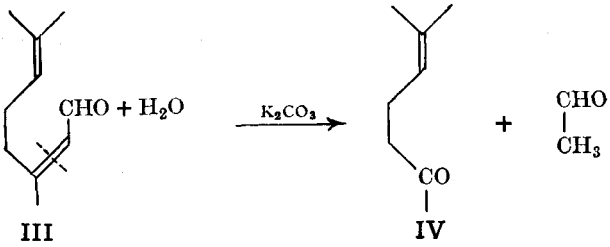
isoprene units joined head to tail). Citral can be reduced by sodium amalgam to an alcohol, geraniol, $C_{10}H_{18}O$, and is oxidised by silver oxide to geranic acid, $C_{10}H_{16}O_2$; since there is no loss of carbon on oxidation to the acid, the oxo group in citral is therefore an aldehyde group (Semmler, 1890). Oxidation of citral with alkaline permanganate, followed by chromic acid, gives acetone, oxalic and lævulic acids (Tiemann and Semmler, 1895). Thus, if citral has structure III, the formation of these oxidation products may be



III

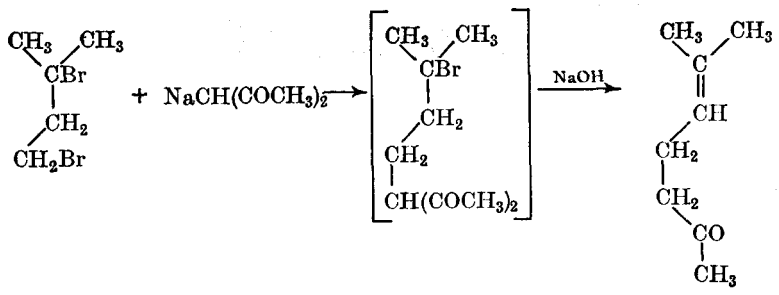
accounted for. This structure is supported by the work of Verley (1897), who found that aqueous potassium carbonate converted citral into 6-methylhept-5-en-2-one, IV, and acetaldehyde. The formation of these products

is readily explained by assuming III undergoes cleavage at the $\alpha:\beta$-double bond; this cleavage by alkaline reagents is a general reaction of $\alpha:\beta$-unsaturated oxo compounds (see Vol. I). Furthermore, methylheptenone itself is also oxidised to acetone and lævulic acid; this is again in accord with

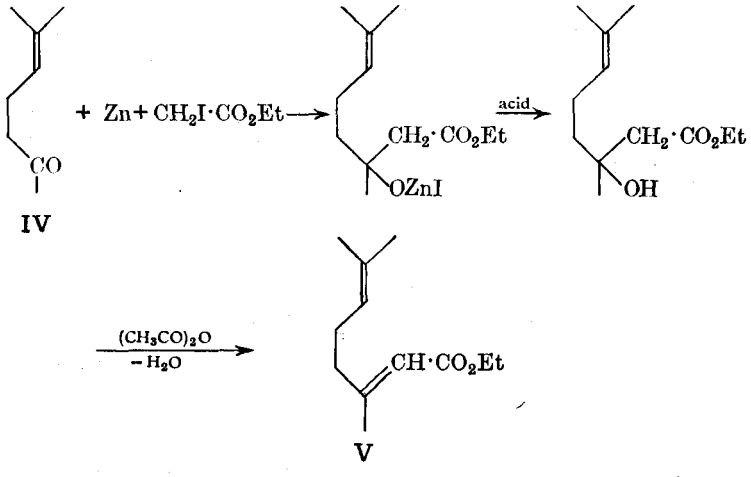


III IV

structure III. The structure of methylheptenone was already known from its synthesis by Barbier and Bouveault (1896). These workers condensed 2 : 4-dibromo-2-methylbutane with sodio-acetylacetone, and heated the resulting compound with concentrated sodium hydroxide solution. Barbier

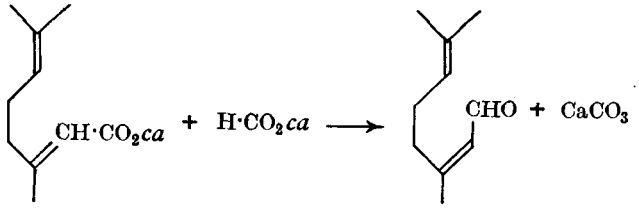


and Bouveault (1896) then converted methylheptenone into geranic ester, V, by means of the Reformatsky reaction, using zinc and ethyl iodoacetate. The synthesis of citral was completed by Tiemann (1898) by distilling a

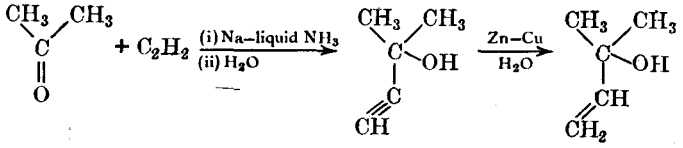


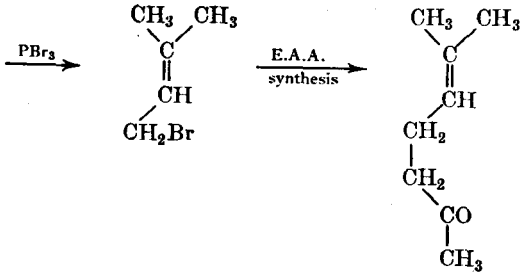
IV V

mixture of the calcium salts of geranic and formic acids (*ca* represents "half an atom of calcium"):

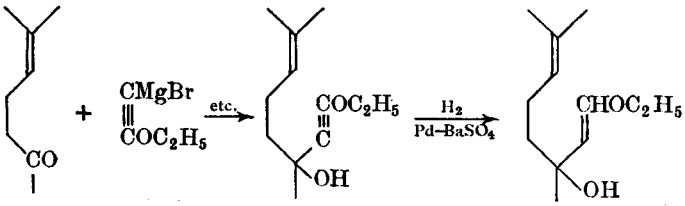


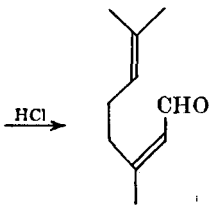
A more recent synthesis of citral is that of Arens and van Dorp (1948). Methylheptenone was first prepared as follows:





Then the methylheptenone was treated with ethoxyacetylene-magnesium bromide, the product reduced and then de-alkylated. It should be noted

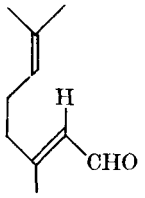




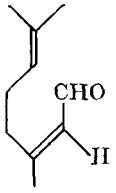
that an *allylic rearrangement* occurs in both parts of this synthesis (see also §8). Ethoxyacetylenemagnesium bromide may conveniently be prepared from chloroacetaldehyde diethyl acetal as follows (Jones *et al.*, 1954):

$$CH_2Cl{\cdot}CH(OC_2H_5)_2 \xrightarrow{NaNH_2} CH{\equiv}C{\cdot}OC_2H_5 \xrightarrow{RMgBr} BrMgC{\equiv}C{\cdot}OC_2H_5$$

Examination of the formula of citral shows that two geometrical isomers are possible:



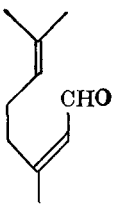
trans-form;
citral-*a*;
geranial



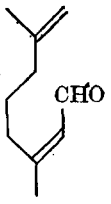
cis-form;
citral-*b*;
neral

Both isomers occur in natural citral, *e.g.*, *two* semicarbazones are formed by citral; both forms of citral itself have also been obtained: **citral-a** (also known as *geranial*) has a b.p. 118–119°/20 mm., and **citral-b** (also known as *neral*) has a b.p. 117–118°/20 mm. The configurations of these two forms have been determined from a consideration of the ring closures of the corresponding alcohols (see geraniol, §7).

The problem of the structure of citral is further complicated for the following reasons. Ozonolysis of citral gives acetone, lævulaldehyde and glyoxal (Harries, 1903, 1907); these products are to be expected from structure III. On the other hand, Grignard *et al.* (1924) also isolated a small amount of formaldehyde from the products of ozonolysis; this points towards structure VI, which has an *iso*propenyl end-group. Thus citral

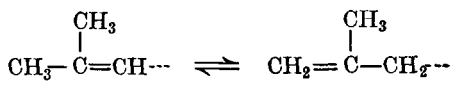


III



VI

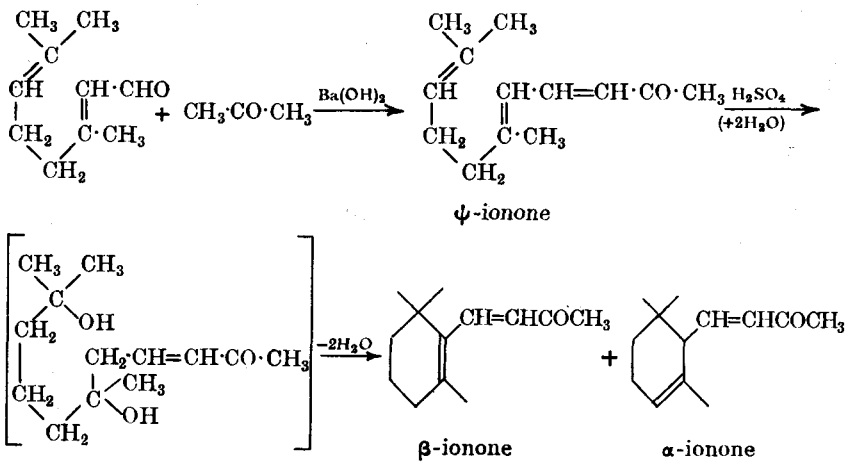
has been regarded as a mixture of *four* substances, two geranials and two nerals. Assuming, then, that both the *iso*propylidene and *iso*propenyl forms are present, it is possible that these two structures form a three-carbon tautomeric system:



Recent work, however, has cast doubt on the existence of these two forms in citral. According to infra-red spectroscopic studies, it appears that naturally occurring acyclic monoterpenes *as a class* possess only the *iso*propylidene end-group structure (Barnard, Bateman *et al.*, 1950). According to these authors, *during oxidative degradation*, partial rearrangement from the *iso*propylidene to the *iso*propenyl structure occurs, and so this method of determining fine structure is unreliable (see also geraniol, §7). Oliver (1961) has developed a chemical together with a chromatographic

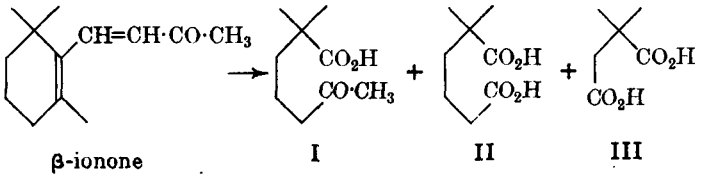
method for separating a mixture of *iso*propylidene and *iso*propenyl isomers. This should be of value in the studies of natural terpenes.

§6. Ionones. When citral is condensed with acetone in the presence of barium hydroxide, *ψ*-ionone is formed and this, on heating with dilute sulphuric acid in the presence of glycerol, forms a mixture of α- and β-ionones (Tiemann and Krüger, 1893). The proportion of α to β varies with the nature of the cyclising agent used, *e.g.*, with sulphuric acid, β-ionone is the main product; with phosphoric acid, α-ionone is the main product. Both ionones have been obtained from natural sources; the β-isomer is optically inactive, whereas the α-isomer can exist in optically active forms

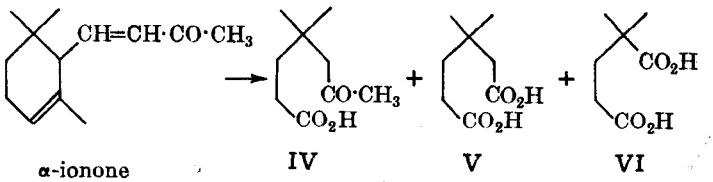


since it contains one asymmetric carbon atom. Actually, the (+)-, (−)- and (±)-forms of α-ionone occur naturally. Very dilute ethanolic solutions of β-ionone have the odour of violets.

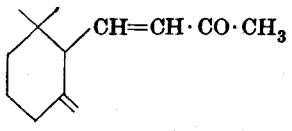
The structures of the ionones were established by a study of the oxidation products produced by potassium permanganate (Tiemann, 1898, 1900);



β-ionone gave geronic acid, I, α : α-dimethyladipic acid, II, and α : α-dimethylsuccinic acid, III. On the other hand, α-ionone gave a mixture of *iso*geronic acid, IV, β : β-dimethyladipic acid, V, and α : α-dimethylglutaric acid, VI.

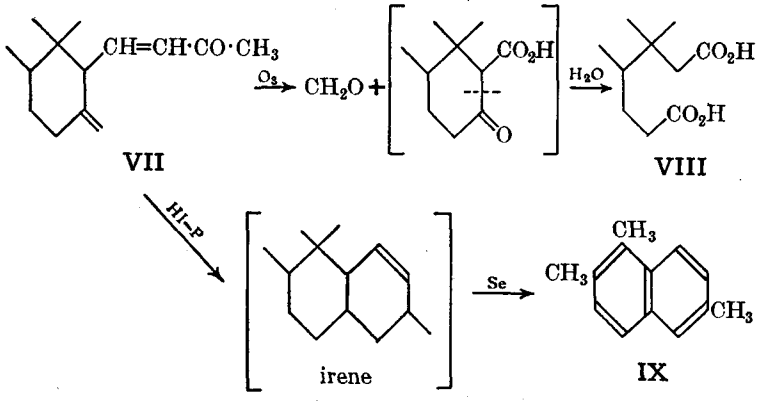


Theimer *et al.* (1962) have isolated γ-ionone (by vapour-phase chromato-

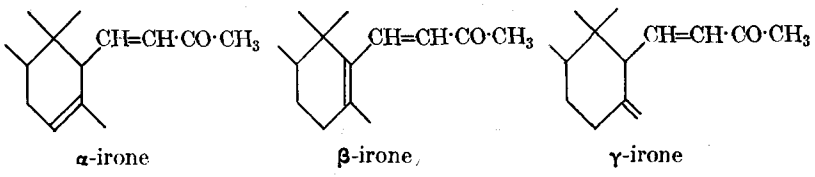


graphy) from the mixture of ionones obtained above (this ionone corresponds to the γ-irone; see below).

The ionones are related to **irone,** $C_{14}H_{22}O$; this occurs in the oil obtained from the orris root. The structure of irone was established by Ruzicka *et al.* (1947), who showed that on ozonolysis, irone gives formaldehyde and $\beta:\beta:\gamma$-trimethylpimelic acid, VIII; also, reduction of irone with hydriodic acid and red phosphorus, followed by dehydrogenation with selenium, gives 1 : 2 : 6-trimethylnaphthalene, IX. Ruzicka therefore proposed structure

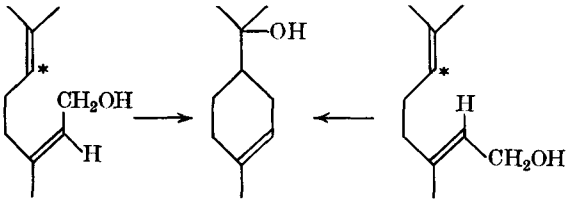


VII for irone. Ruzicka (1947) further showed that irone was a mixture of three isomers (VII is γ-irone):



§7. **Geraniol,** $C_{10}H_{18}O$, b.p. 229–230°/757 mm. This is found in many essential oils, particularly rose oil. Geraniol was shown to be a primary alcohol, *e.g.*, on oxidation it gives an aldehyde (citral-*a*); and since it forms a tetrabromide, geraniol therefore contains two double bonds. Reduction of citral produces geraniol, but at the same time some **nerol** is formed. The *structural* identity of geraniol and nerol is shown by the following facts. Both add on two molecules of hydrogen when hydrogenated catalytically; thus both contain two double bonds. Both give the same saturated alcohol, $C_{10}H_{22}O$. Also, on oxidation, geraniol and nerol give the same oxidation products which, at the same time, show the positions of the double bonds to be 2 and 7 (*cf.* citral, §5). Thus geraniol and nerol are geometrical isomers. Geraniol has been assigned the *trans* configuration and nerol the *cis* on the fact that cyclisation to α-terpineol (§11) by means of dilute sulphuric acid takes place about 9 times as fast with nerol as it does with geraniol;

this faster rate with nerol is due to the proximity of the alcoholic group to the carbon (*) which is involved in the ring formation. Thus:

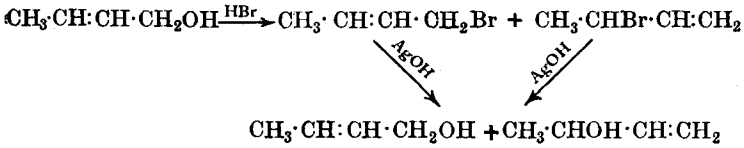


nerol (*cis*-) α-terpineol geranol (*trans*-)

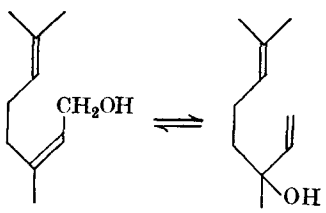
Nerol also occurs naturally in various essential oils, *e.g.*, oil of neroli, bergamot, etc.; its b.p. is 225–226°.

Knights *et al.* (1955) have found that, on ozonolysis, geranyl acetate gives less than 3 per cent. of formaldehyde, and have concluded that the acetate and geraniol itself have predominantly the *iso*propylidene structure (*cf.* citral, §5).

§8. Linalool, $C_{10}H_{18}O$, b.p. 198–199°. This is an optically active compound; the (−)-form occurs in rose oil and the (+)-form in orange oil. It was shown to be a tertiary alcohol, and since it adds on two molecules of hydrogen on catalytic hydrogenation, it must contain two double bonds. When heated with acetic anhydride, linalool is converted into geranyl acetate; and the latter is converted into the former by heating with steam at 200° under pressure. Also, heating linalool with hydrogen chloride in toluene solution at 100° produces geranyl chloride, and this, when treated with moist silver oxide in benzene solution, is reconverted into linalool. These reactions are parallel to those which occur when crotyl alcohol is treated with hydrogen bromide; a mixture of crotyl bromide and methylvinylcarbinyl bromide is obtained. When *either* of these products is treated with moist silver oxide, a mixture of crotyl alcohol and methylvinylcarbinol is obtained.

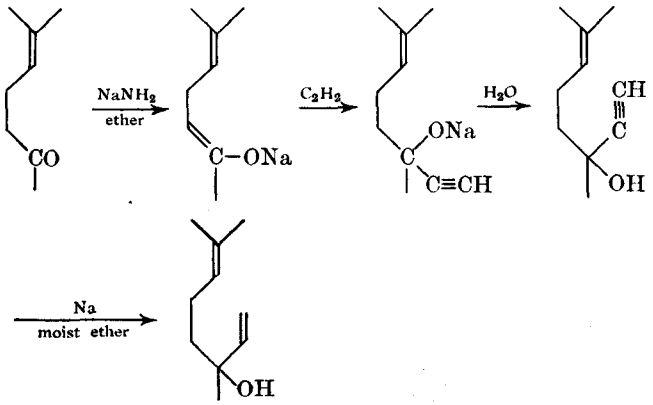


Thus the elucidation of the structure of linalool is complicated by the ease with which the *allylic rearrangement* occurs (see also Vol. I). Since the structure of geraniol is known, a possible structure for linalool is obtained on the basis of this allylic rearrangement.



geraniol linalool

This structure has been confirmed by synthesis of linalool (Ruzicka *et al.*, 1919); 6-methylhept-5-en-2-one was treated as follows:

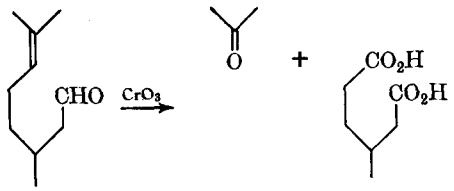


(±)-linalool

Normant (1955) has synthesised linalool in one step by the action of vinylmagnesium bromide on methylheptenone.

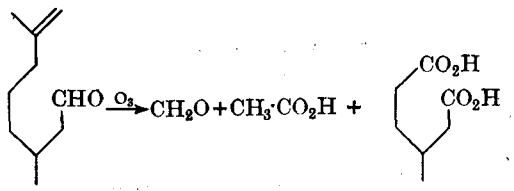
§9. **Citronellal,** $C_{10}H_{18}O$. This is an optically active compound which occurs in citronella oil. Citronellal is an aldehyde; reduction with sodium amalgam converts it into the alcohol citronellol, $C_{10}H_{20}O$, and oxidation gives citronellic acid, $C_{10}H_{18}O_2$. Now there is another aldehyde, **rhodinal,** which is isomeric with citronellal, and on reduction, rhodinal gives the alcohol, rhodinol, which is isomeric with citronellol. Furthermore, reduction of ethyl geranate with sodium and ethanol gives rhodinol (Bouveault *et al.*, 1900).

Oxidation of citronellal with chromic acid gives β-methyladipic acid and acetone (Tiemann *et al.*, 1896, 1897). Rhodinal also gives the *same* products



I

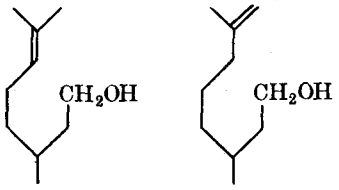
on oxidation. Thus structure I would fit the facts for both citronellal and rhodinal. On the other hand, ozonolysis of citronellal gives β-methyladipic acid, acetone and some formaldehyde (Harries *et al.*, 1908). These results point towards structure II for citronellal, as well as I. Thus citronellal appears to be a mixture of I (*iso*propylidene end-group) and II (*iso*propenyl end-group). Furthermore, a detailed study of rhodinal has shown that this



II

compound is identical with citronellal, but consists of a mixture of the two forms in different proportions (but *cf.* citral, §5).

§9a. Citronellol and Rhodinol, $C_{10}H_{20}O$. (−)-Citronellol occurs in rose and geranium oils, and is a mixture of the two forms:

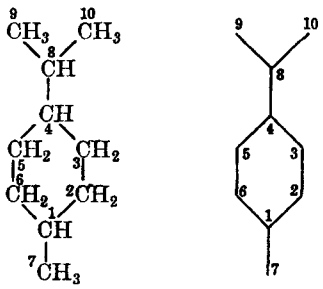


The (+)-form of citronellol is made commercially by reduction of citronellal with sodium or aluminium amalgam; it also occurs in Java citronella oil.

Rhodinol is identical with citronellol, but the proportions of the two forms are different from those which occur in citronellol; the identity of citronellol and rhodinol is shown by the products of ozonolysis.

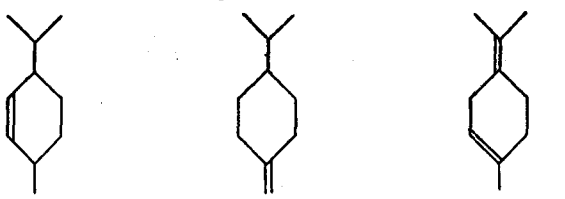
MONOCYCLIC MONOTERPENES

§10. Nomenclature. For the purposes of nomenclature of the monocyclic monoterpenes, the fully saturated compound *p*-methyl*iso*propyl*cyclo*hexane, hexahydro-*p*-cymene or *p*-menthane, $C_{10}H_{20}$, is used as the parent substance; it is a synthetic compound, b.p. 170°. *p*-Menthane is I, and II is a conventional method of drawing formula I. The positions of substituents and double bonds are indicated by numbers, the method of numbering being shown in I (and II). When a compound derived from *p*-menthane



I II

contains one or more double bonds, ambiguity may arise as to the position of a double bond when this is indicated in the usual way by a number which locates the *first* carbon atom joined by the double bond. To prevent ambiguity, the *second* carbon atom joined to the double bond is also shown,



Δ^2-*p*-menthene; 2-*p*-menthene; *p*-menth-2-ene; *p*-menthene-2.

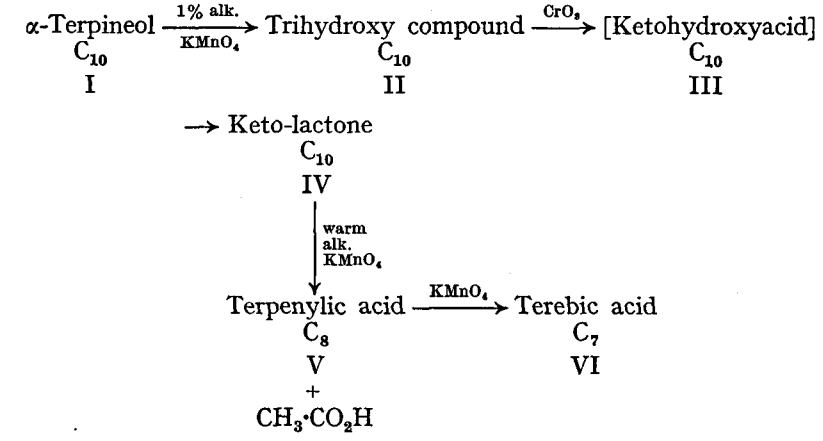
p-menth-1(7)-ene

p-mentha-1:4(8)-diene

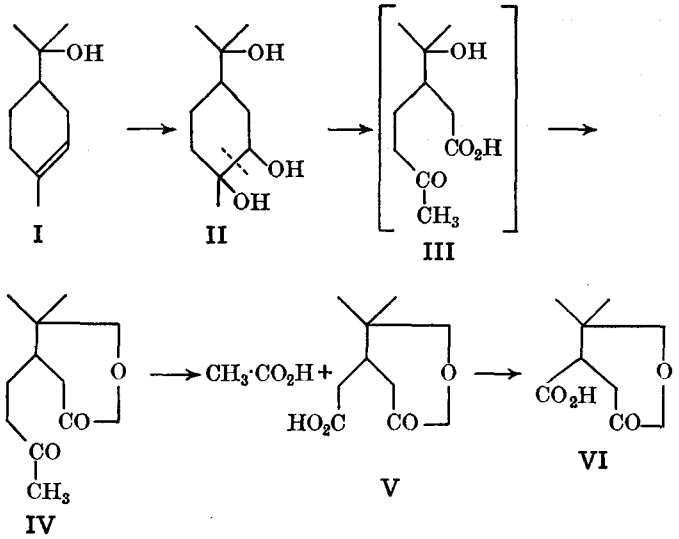
but is placed in parentheses. The previous examples illustrate the method of nomenclature; in the first example, all the types of methods of nomenclature have been given; in the second and third examples, only the nomenclature that will be used in this book is given.

§11a. α-Terpineol. This is an optically active monoterpene that occurs naturally in the (+)-, (−)- and (±)-forms; it is a solid, m.p. (of the racemic modification) 35°. The molecular formula of α-terpineol is $C_{10}H_{18}O$, and the oxygen atom is present as a tertiary alcoholic group (as shown by the reactions of α-terpineol). Since α-terpineol adds on two bromine atoms, it therefore contains one double bond. Thus the parent (saturated) hydrocarbon of α-terpineol has the molecular formula $C_{10}H_{20}$. This corresponds to C_nH_{2n}, the general formula of the (monocyclic) *cyclo*alkanes, and so it follows that α-terpineol is a monocyclic compound.

When heated with sulphuric acid, α-terpineol forms some *p*-cymene. Taking this in conjunction with the tentative proposal that α-terpineol is monocyclic, it is reasonable to infer that α-terpineol contains the *p*-cymene skeleton. Thus we may conclude that α-terpineol is probably *p*-menthane with one double bond and a tertiary alcoholic group. The positions of these functional groups were ascertained by Wallach (1893, 1895) by means of *graded* oxidation. The following chart gives the results of Wallach's work; only the carbon content is indicated to show the fate of these carbon atoms (the formulæ are given in the text).



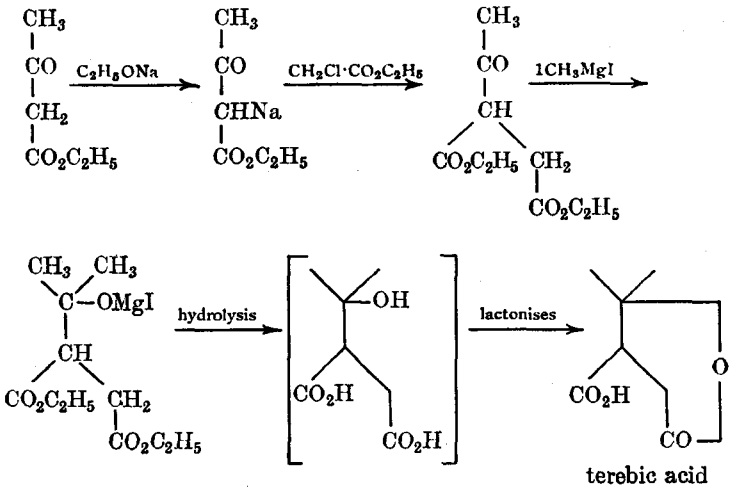
Oxidation of α-terpineol, I, with 1 per cent. alkaline potassium permanganate hydroxylates the double bond to produce the trihydroxy compound II, $C_{10}H_{20}O_3$. This, on oxidation with chromic acid (chromium trioxide in acetic acid), produces a compound with the molecular formula $C_{10}H_{16}O_3$ (IV). This compound was shown to contain a ketonic group, and that it was neutral, *e.g.*, it gave no reaction with sodium carbonate solution. When, however, IV was refluxed with excess of standard sodium hydroxide solution, and then back titrated, it was found that alkali had been consumed, the amount corresponding to the presence of one carboxyl group. Thus compound IV appears to be the *lactone* of a monocarboxylic acid. Furthermore, since it is the lactone that is isolated and not the hydroxy acid, this *spontaneous* lactonisation may be interpreted as being produced from a γ-hydroxyacid, *i.e.*, IV is a γ-lactone, and therefore III is a γ-hydroxyacid. It is possible, however, for δ-hydroxyacids to spontaneously lactonise, and so whether IV is a γ- or δ-lactone is uncertain at this stage of the evidence. Now, since IV is formed from II by scission of the glycol bond, and since



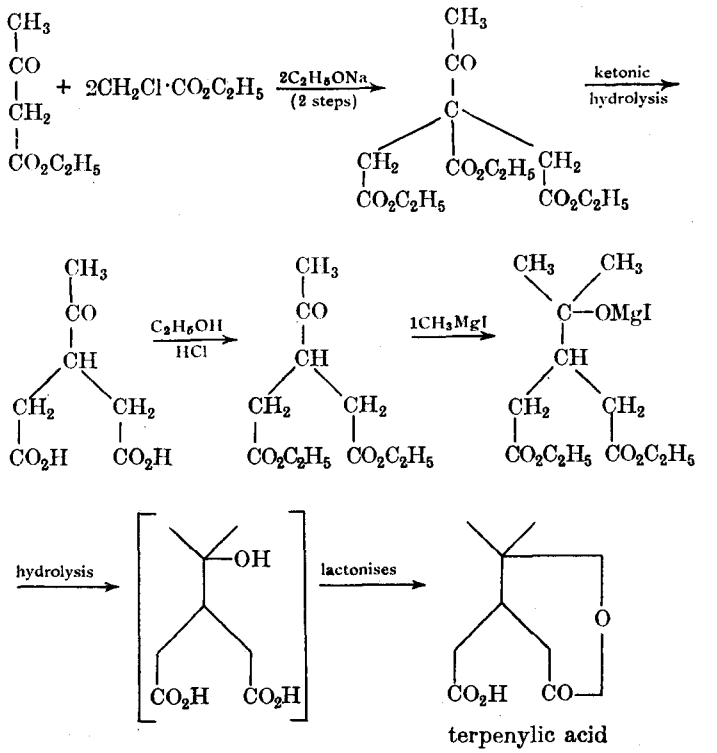
there is *no loss of carbon atoms* in the process, the double bond must therefore be in the ring in I. On warming with alkaline permanganate, IV gave acetic acid and a compound $C_8H_{12}O_4$ (V). The formation of acetic acid suggests that IV is a *methyl ketone, i.e.*, a CH_3·CO group is present. Thus IV is a methyl ketone and a lactone; it is known as homoterpenyl methyl ketone, and the structure assigned to it has been confirmed by synthesis (Simonsen *et al.*, 1932). A study of the properties of terpenylic acid, V, showed that it was the lactone of a monohydroxydicarboxylic acid. Further oxidation of terpenylic acid gives terebic acid $C_7H_{10}O_4$ (VI), which is also the lactone of a monohydroxydicarboxylic acid.

The above reactions can be formulated as shown, *assuming* I (*p*-menth-1-en-8-ol) as the structure of α-terpineol. These reactions were formulated by Wallach, who adopted formula I which had been proposed by Wagner (1894). The structures of terpenylic (V) and terebic (VI) acids were established by synthesis, *e.g.*, those of Simonsen (1907).

Terebic acid, m.p. 175°.

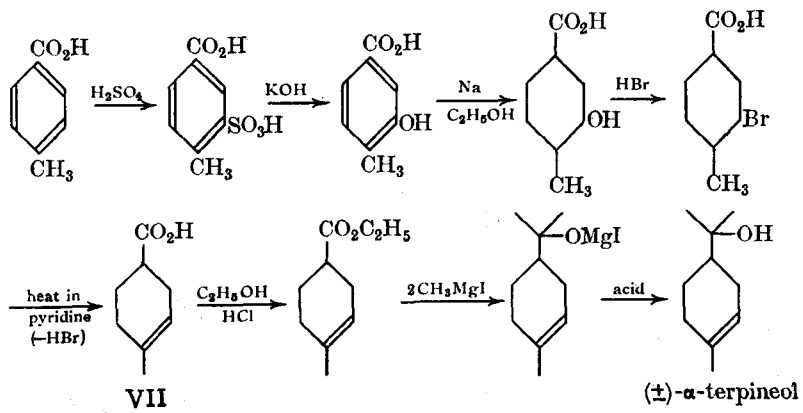


Terpenylic acid, m.p. 90°.

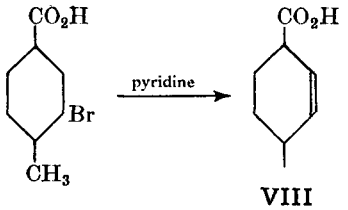


It is of interest to note here that Sandberg (1957) has prepared the β-acetotricarballylate in *one* step from acetoacetic ester and ethyl bromoacetate in the presence of sodium hydride (in benzene solution).

These syntheses strengthen the evidence for the structure assigned to α-terpineol, but final proof rests with a synthesis of α-terpineol itself. This has been carried out by Perkin, junior (1904), and by Perkin, junior, with Meldrum and Fisher (1908). Only the second synthesis is given here; this starts with *p*-toluic acid.



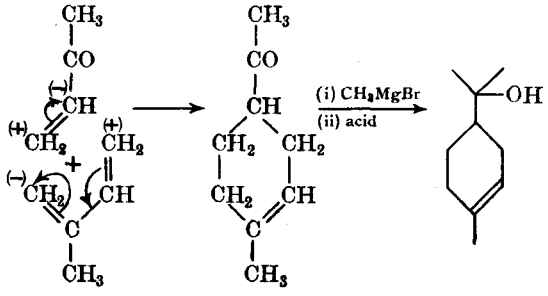
Compound VII was also resolved with strychnine, each enantiomorph treated as shown above (esterified, etc.), and thereby resulted in the formation of (+)- and (−)-terpineol. It should be noted that in the above synthesis



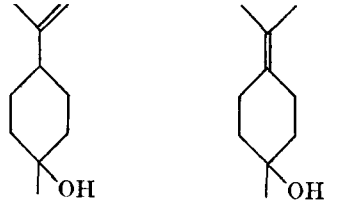
VIII

the removal of a molecule of hydrogen bromide from 3-bromo-4-methyl-*cyclo*hexane-1-carboxylic acid to give VII is an ambiguous step; instead of VII, compound VIII could have been formed. That VII and not VIII is formed rests on the analytical evidence for the position of this double bond; VIII cannot give the products of oxidation that are actually obtained from α-terpineol.

A much simpler synthesis of α-terpineol has been carried out by Alder and Vogt (1949); this makes use of the Diels–Alder reaction, using isoprene and methyl vinyl ketone as the starting materials (see also Vol. I).



Two other terpineols are also known, *viz.*, **β-terpineol** and **γ-terpineol**; both occur naturally.

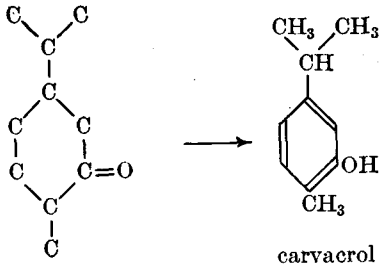


β-terpineol
m.p. 32-33°

γ-terpineol
m.p. 68-70°

§12. Carvone, $C_{10}H_{14}O$, b.p. 230°/755 mm. This occurs in various essential oils, *e.g.*, spearmint and caraway oils, in optically active forms and also as the racemic modification.

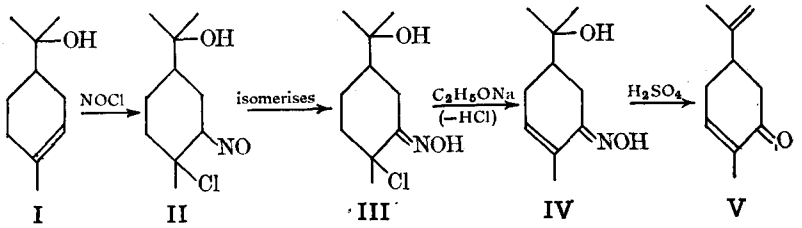
Carvone behaves as a ketone and, since it adds on four bromine atoms, it therefore contains two double bonds. Thus the parent hydrocarbon is $C_{10}H_{20}$, and since this corresponds to the general formula C_nH_{2n}, carvone is monocyclic. When heated with phosphoric acid, carvone forms carvacrol; this suggests that carvone probably contains the *p*-cymene structure, and that the keto group is in the ring in the *ortho*-position with respect to the methyl group.



carvone skeleton

carvacrol

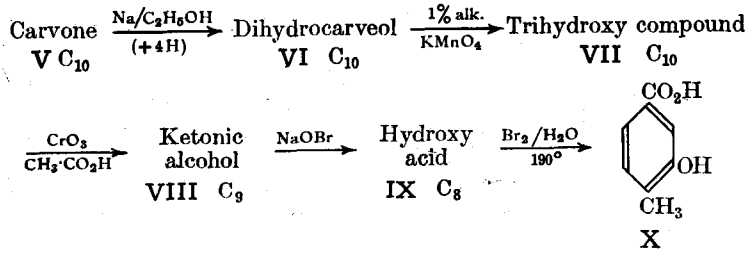
The structure of carvone is largely based on the fact that carvone may be prepared from α-terpineol as follows:



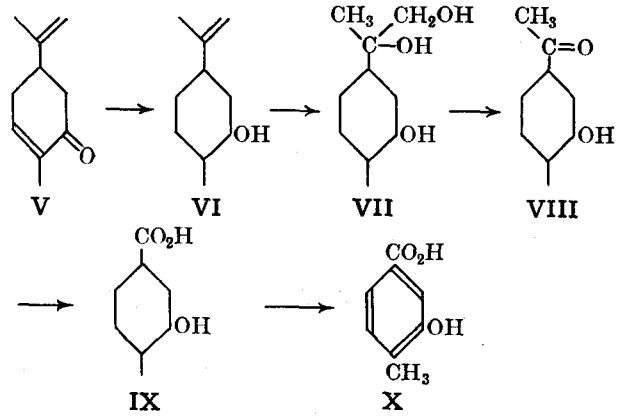
I II III IV V

The addition of nitrosyl chloride to α-terpineol, I, produces α-terpineol nitrosochloride, II, the addition occurring according to Markownikoff's rule (the chlorine is the negative part of the addendum; see Vol. I). This nitrosochloride rearranges spontaneously to the oximino compound, III (see nitroso-compounds, Vol. I; it might be noted that this rearrangement proves the orientation of the addition of the nitrosyl chloride to the double bond; addition the other way could not give an oxime, since there is no hydrogen atom at position 1 in α-terpineol). Removal of a molecule of hydrogen chloride from III by means of sodium ethoxide produces IV, and this, on warming with dilute sulphuric acid, loses a molecule of water with simultaneous hydrolysis of the oxime to form carvone, V. Thus, according to this interpretation of the reactions, carvone is *p*-menth-6 : 8-dien-2-one. Actually, these reactions show that carvone has the same carbon skeleton as α-terpineol, and also confirm the position of the keto group. They do not prove conclusively the positions of the two double bonds; instead of position 6 (in IV), the double bond could have been 1(7), and instead of position 8 (as in V), the double bond could have been 4(8). Thus the above reactions constitute an ambiguous synthesis of carvone (α-terpineol has already been synthesised). The exact positions of these two double bonds have been determined *analytically* as follows.

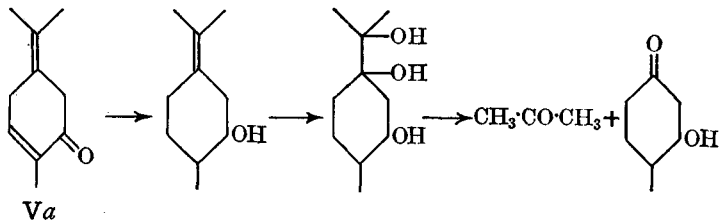
The double bond in the 8*-position.* The following reactions were carried out by Tiemann and Semmler (1895).



Reduction of carvone, V, with sodium and ethanol gives dihydrocarveol, $C_{10}H_{18}O$ (VI); this is a secondary alcohol and contains *one* double bond, *i.e.*, the keto group and *one* of the two double bonds in carvone have been reduced. Hydroxylation of the double bond in dihydrocarveol by means of 1 per cent. alkaline permanganate produces the trihydroxy compound $C_{10}H_{20}O_3$ (VII). Oxidation of VII with chromic acid causes scission of the glycol bond to produce a compound $C_9H_{16}O_2$ (VIII); this was shown to contain a keto group and a hydroxyl (alcoholic) group. The action of sodium hypobromite on VIII caused the loss of one carbon atom to produce the compound $C_8H_{14}O_3$ (IX); this was shown to be a hydroxymonocarboxylic acid, and since *one* carbon is lost in its formation, its precursor VIII must therefore be a methyl ketone. Finally, dehydrogenation of IX by heating with bromine-water at 190° under pressure produced *m*-hydroxy-*p*-toluic acid, X (a *known* compound). Tiemann and Semmler explained these reactions on the assumption that one double bond in carvone is in the 8-position. Thus:



Had the double bond been in the 4(8)-position (structure V*a*), then compound VIII, and consequently X, could not have been obtained, since *three* carbon atoms would have been lost during the oxidation.

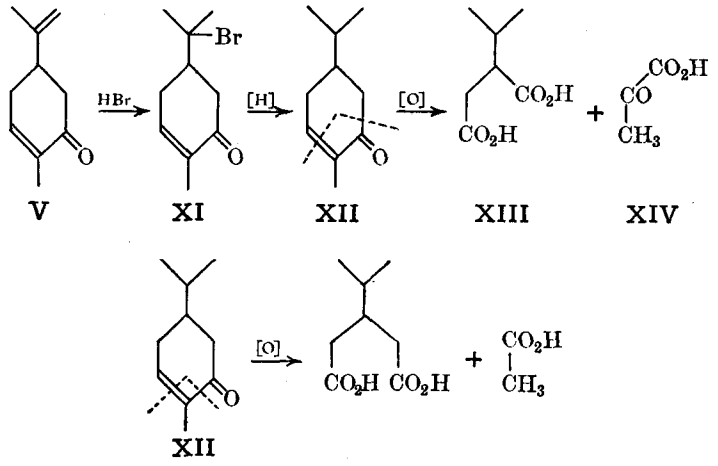


V*a*

It might be noted in passing that V contains an asymmetric carbon atom, whereas V*a* is a symmetrical molecule and so cannot exhibit optical activity. Since carvone is known in optically active forms, structure V*a* must be rejected on these grounds.

The double bond in the 6-position. Carvone adds on one molecule of hydrogen bromide to form carvone hydrobromide, $C_{10}H_{15}OBr$ (XI), and this, on treatment with zinc dust and methanol, is converted into carvotanacetone, $C_{10}H_{16}O$ (XII), by replacement of the bromine atom by hydrogen. Thus the final result of these reactions is to saturate *one* of the two double bonds in carvone. Carvotanacetone, on oxidation with permanganate, gives *iso*propylsuccinic acid, XIII, and pyruvic acid, XIV (Semmler,

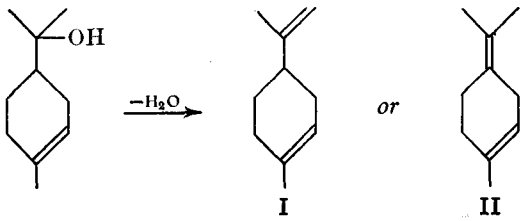
1900). These products are obtainable only if the ring contains the double bond in the 6-position. Had the double bond been in the 1(7)-position,



formic acid and not pyruvic acid would have been obtained. Further support for the 6-position is provided by the work of Simonsen *et al.* (1922), who obtained *β-iso*propylglutaric acid and acetic acid on oxidation of carvotanacetone with permanganate.

§13. **Limonene,** $C_{10}H_{16}$, b.p. 175·5–176·5°. This is optically active; the (+)-form occurs in lemon and orange oils, the (−)-form in peppermint oil, and the (±)-form in turpentine oil. The racemic modification is also produced by racemisation of the optically active forms at about 250°. The racemic modification is also known as **dipentene;** this name was given to the inactive form before its relation to the active form (limonene) was known.

Since limonene adds on four bromine atoms, it therefore contains two double bonds. (+)-Limonene may be prepared by dehydrating (+)-α-terpineol with potassium hydrogen sulphate, and limonene (or dipentene) may be converted into α-terpineol on shaking with dilute sulphuric acid.



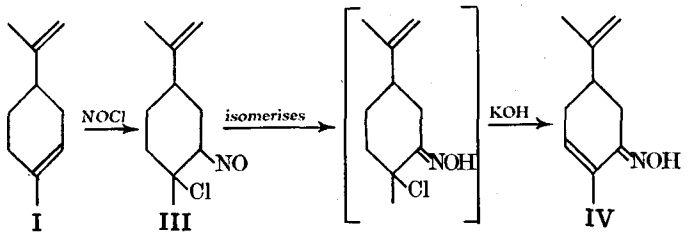
Thus the carbon skeleton and the position of one double bond in limonene are known. The position of the other double bond, however, remains uncertain from this preparation; I or II is possible.

Proof for position 8. Structure I contains an asymmetric carbon atom (C_4), and hence can exhibit optical activity. II is a symmetrical molecule and so cannot be optically active. Therefore I must be limonene.

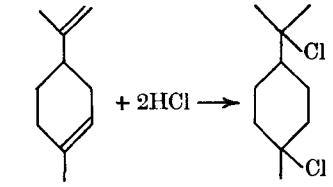
Chemical proof for position 8 is afforded by the following reactions:

$$\underset{\text{I}}{\text{Limonene}} \xrightarrow{\text{NOCl}} \underset{\text{III}}{\text{Limonene nitrosochloride}} \xrightarrow[C_2H_5OH]{\text{KOH}} \underset{\text{IV}}{\text{carvoxime}}$$

Since the structure of carvoxime is known, it therefore follows that I must have one double bond in position 8; thus the above reactions may be written:

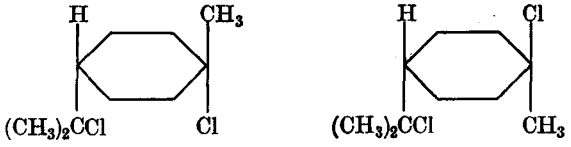


The connection between limonene and dipentene is shown by the fact that (+)- or (−)-limonene adds on two molecules of hydrogen chloride in the presence of moisture to form limonene dihydrochloride, and this is identical with dipentene dihydrochloride.



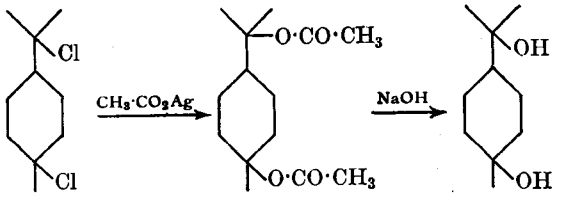
(+)- or (−)-limonene

Limonene dihydrochloride no longer contains an asymmetric carbon atom, and so is optically inactive. It can, however, exhibit geometrical isomerism; the *cis*-form is produced from limonene, and the *trans*-form from cineole (§14).



cis *trans*

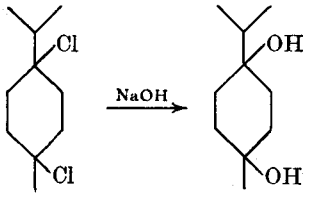
Dipentene can be regenerated by heating the dihydrochloride with sodium acetate in acetic acid, or boiling with aniline. On the other hand, when limonene dihydrochloride is heated with silver acetate in acetic acid, and then hydrolysing the ester with sodium hydroxide, **1 : 8-terpin** is formed; the direct action of sodium hydroxide on the dihydrochloride regenerates dipentene.



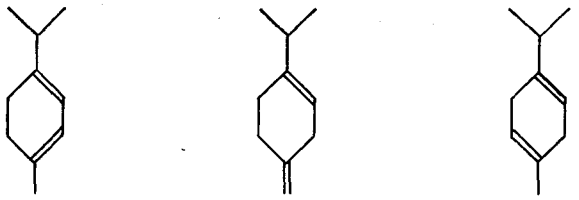
1:8-terpin

1 : 8-Terpin exists in two geometrical isomeric forms, corresponding to the *cis* and *trans* dipentene dihydrochlorides. *cis*-1 : 8-Terpin is the common form, m.p. 105°, and readily combines with one molecule of water to form terpin hydrate. The *trans*-form, m.p. 158–159°, does not form a hydrate (see also §14).

There is also a **1 : 4-terpin;** this was originally prepared by the action of dilute alkali on terpinene dihydrochloride.



Terpinenes, $C_{10}H_{16}$. There are three isomeric terpinenes, and all give the same terpinene dihydrochloride with hydrogen chloride.



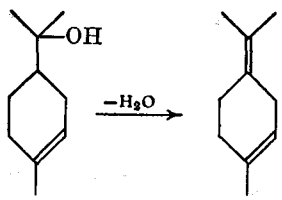
α-terpinene
b.p. 180-182°

β-terpinene
b.p. 173-174°

γ-terpinene
b.p. 69-73°/20mm.

All three occur naturally.

Terpinolene, $C_{10}H_{16}$, b.p. 67–68°/10 mm. This occurs naturally. It is not optically active, and since it may be prepared by dehydrating α-terpineol with oxalic acid, its structure is known (it is II, the alternative formula offered for limonene). Terpinolene adds on two molecules of hydrogen chloride to form dipentene dihydrochloride.



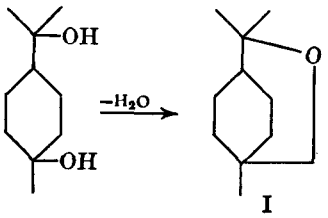
II

Phellandrenes, $C_{10}H_{16}$. There are two phellandrenes, both of which are optically active, and all the enantiomorphs occur naturally.

α-phellandrene
b.p. 58-59°/16 mm.

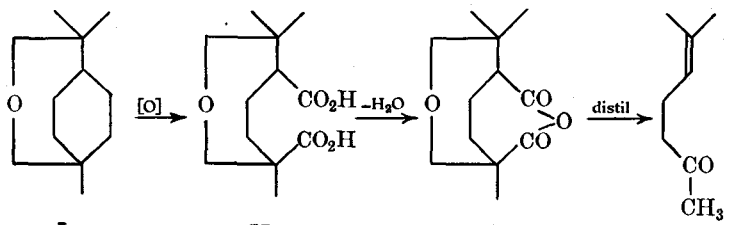
β-phellandrene
b.p. 171-172°

§14. 1 : 8-Cineole, $C_{10}H_{18}O$, b.p. 174·4°. This occurs in eucalyptus oils. It is isomeric with α-terpineol, but contains neither a hydroxyl group nor a double bond. The oxygen atom in cineole is inert, *e.g.*, it is not attacked by sodium or by the usual reducing agents. This inertness suggests that the oxygen atom is of the ether type. Support for this is obtained from the fact that dehydration of *cis*-1 : 8-terpin gives 1 : 8-cineole; at the same time, this reaction suggests that the structure of cineole is I.



I

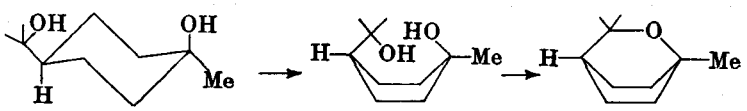
Further support for this structure is afforded by a study of the products obtained by oxidation (Wallach *et al.*, 1888, 1890, 1892). When oxidised with potassium permanganate, cineole forms cineolic acid, II, and this, on distillation with acetic anhydride, forms cineolic anhydride, III. When distilled at atmospheric pressure, cineolic anhydride forms 6-methylhept-5-en-2-one, IV, a *known* compound (§5). These reactions were interpreted by Wallach as follows:



I II III IV

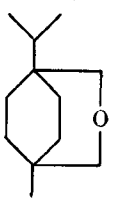
Further work on the structure of cineolic acid has confirmed the above sequence of reactions (Rupe, 1901, —).

It seems most probable that the 1 : 8-terpins have chair conformations, but when they form 1 : 8-cineole, the latter possesses the boat conformation; thus:



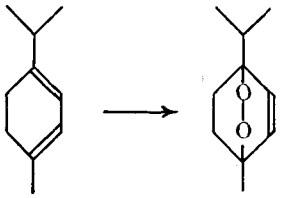
cis-terpin 1 : 8-cineole

There is also a **1 : 4-cineole;** this occurs naturally.



1:4-cineole
b.p. 172°

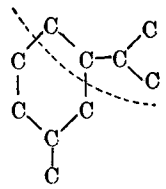
Ascaridole, $C_{10}H_{16}O_2$, b.p. 96–97°/8 mm. The cineoles are oxides; ascaridole, however, is a peroxide, the only known terpene peroxide, and it occurs naturally in, *e.g.*, chenopodium oil. When heated to 130–150°, ascaridole decomposes with explosive violence. When reduced catalytically, ascaridole forms 1 : 4-terpin (Wallach, 1912), and this led to the suggestion that



V

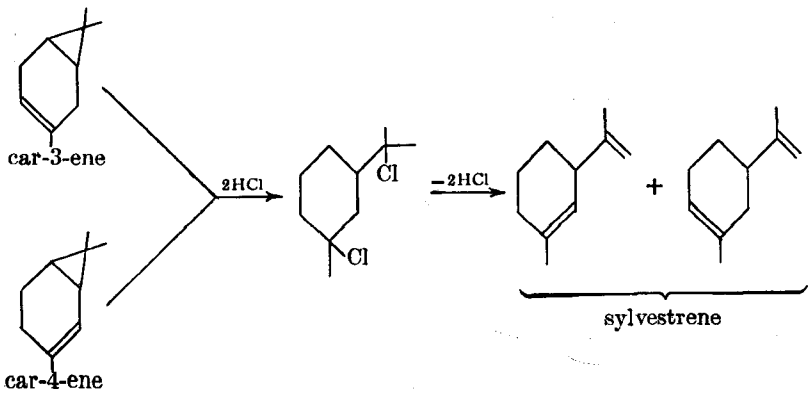
ascaridole is V. This structure has been confirmed by further analytical work. Ascaridole has been synthesised by Ziegler *et al.* (1944) by the irradiation of α-terpinene in dilute solution in the presence of chlorophyll.

§15. Sylvestrene, $C_{10}H_{16}$, b.p. 175–178°. This compound exists in (+)-, (−)- and (±)-forms; the racemic modification is also known as **carvestrene** (*cf.* limonene and dipentene, §13). The (+)-form of sylvestrene was first obtained from Swedish pine needle oil (Attenberg, 1877), and was shown to contain the *m*-cymene carbon skeleton (Baeyer *et al.*, 1898). Thus sylvestrene appeared to be the only monocyclic monoterpene which did not have the *p*-cymene structure and was obtainable from natural sources. Although the *m*-cymene structure can be divided into two isoprene units (Wallach's isoprene rule), these two units are not joined head to tail.



m-cymene skeleton

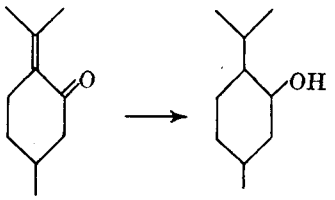
Subsequent work, however, showed that sylvestrene does not occur in pine oil. In the extraction of sylvestrene, the pine oil is heated with hydrogen chloride to give sylvestrene dihydrochloride. This compound was shown



by Simonsen *et al.* (1923, 1925) to be produced by the action of hydrogen chloride on car-3-ene, *i.e.*, these workers showed conclusively that the terpene originally present in Swedish pine oil is car-3-ene. Sylvestrene may be obtained from its dihydrochloride by heating the latter with aniline; removal of hydrogen chloride from the ring can give rise to two possible positions for the ring double bond. Analytical work has shown that the side-chain is *iso*propenyl (and not *iso*propylidene), and that sylvestrene is a mixture of the two forms, *m*-mentha-1 : 8-diene and *m*-mentha-6 : 8-diene. Furthermore, it has been shown that car-4-ene is also present in pine oil; both of these carenes are readily converted into sylvestrene, and so it appears that the precursor of sylvestrene (itself a mixture) is a mixture of the two carenes (see §21).

The enantiomorphs of sylvestrene have been synthesised (Perkin, junior, *et al.*, 1913), and it has also been shown that an equimolecular mixture of the dihydrochlorides of (+)- and (−)-sylvestrene is identical with carvestrene dihydrochloride.

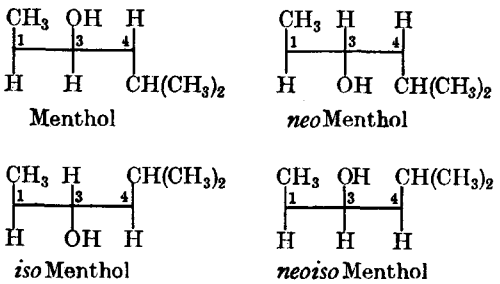
§16. Menthol and menthone. Menthol, $C_{10}H_{20}O$, is an optically active compound, but only the (−)-form occurs naturally, *e.g.*, in peppermint oils. (−)-Menthol, m.p. 34°, is a saturated compound, and the functional nature of the oxygen atom is alcoholic, as shown by its reactions, *e.g.*, menthol forms esters. Furthermore, since oxidation converts menthol into menthone, a *ketone*, the alcoholic group in menthol is therefore secondary. Also, since reduction with hydrogen iodide gives *p*-menthane, menthol most probably contains this carbon skeleton. Finally, since (+)-pulegone gives menthol on reduction, and since the structure of pulegone is known to be I (see §17), it therefore follows that menthol must be II. This structure,



I II

p-menth-3-ol, for menthol has been confirmed by consideration of the oxidation products of menthone (see below), and also by the synthesis of menthol.

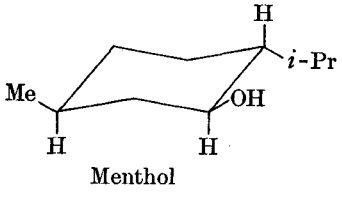
Examination of the menthol structure shows that three dissimilar asymmetric carbon atoms (1, 3 and 4) are present; thus eight optically active forms (four racemic modifications) are possible theoretically. All eight enantiomorphs are known and their configurations are as follows (the horizontal lines represent the plane of the *cyclo*hexane ring):



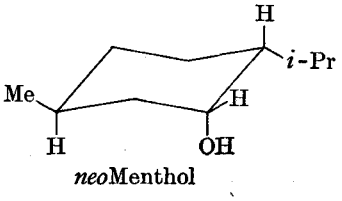
These configurations have been assigned from a study of chemical and optical relationships and the Auwers–Skita rule. More recently the application of conformational analysis has confirmed these results. Eliel (1953) applied the principle that the esterification of an axial hydroxyl group occurs less readily than with an equatorial one. Furthermore, Eliel postulated that the reaction proceeds *via* the conformation of the molecule in which the reactive hydroxyl group is equatorial, and that the rate differences should be attributed to that energy necessary to place the other substituents, if necessary, into the axial conformation (see also §12. IV). On this basis, the rates of esterification of the isomeric menthols will be:

menthol > *iso-* > *neoiso-* > *neo-*.

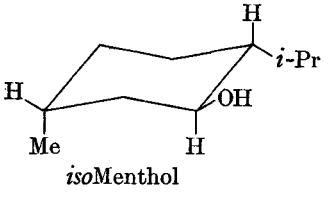
These are the orders of rates actually obtained by Read *et al.* (1934). The following conformations have been assigned by Eliel from chemical studies, and are supported by Cole *et al.* (1956) from their infra-red spectra and conformation studies.



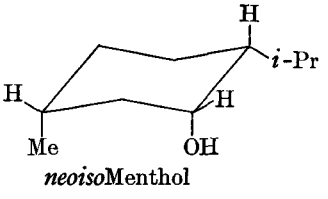
Menthol



*neo*Menthol



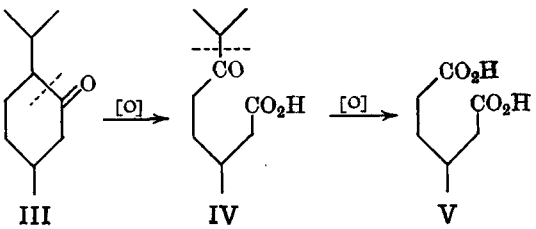
*iso*Menthol



*neoiso*Menthol

In menthol, *all* of the substituents are equatorial, and in the rest one is axial. It should also be noted that the larger of the two alkyl groups (*iso*-propyl) is always equatorial (*cf.* §11. IV).

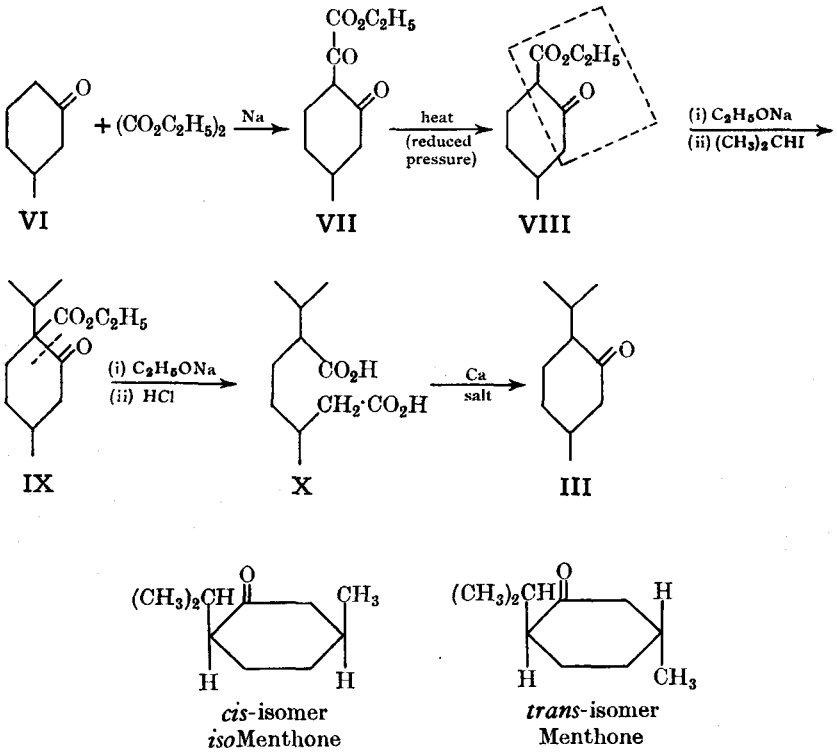
Menthone, $C_{10}H_{18}O$, b.p. 204°/750 mm. (—)-Menthone occurs in peppermint oil, and it may readily be prepared by the oxidation of (—)-menthol with chromic acid. Menthone is a saturated compound which has the characteristic properties of a ketone. When heated with hydriodic acid and red phosphorus, menthone is reduced to *p*-menthane; thus this skeleton is present in menthone. Oxidation of menthone with potassium permanganate produces a compound $C_{10}H_{18}O_3$; this compound was shown to contain a keto-group and one carboxyl group, and is known as ketomenthylic acid (IV). Ketomenthylic acid itself is very readily oxidised by permanganate to β-methyladipic acid (V) and some other acids (Arth, 1886; Manasse *et al.*, 1894). The foregoing oxidative reactions may be formulated as follows, on the *assumption* that III is the structure of menthone. This



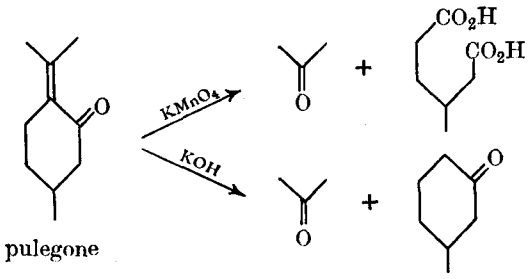
III IV V

structure for menthone has been confirmed by synthesis, *e.g.*, Kötz and Schwarz (1907) obtained menthone by the distillation of the calcium salt of β′-methyl-α-*iso*propylpimelic acid, which was prepared as follows. 3-Methyl-*cyclo*hexanone, VI, was condensed with ethyl oxalate in the presence of sodium, and the product VII then heated under reduced pressure; this gave the ethyl ester of 4-methyl*cyclo*hexan-2-one-1-carboxylic acid, VIII. VIII, on treatment with sodium ethoxide followed by *iso*propyl iodide, gave IX, and this when boiled with ethanolic sodium ethoxide and the product then acidified, gave β′-methyl-α-*iso*propylpimelic acid, X (note the acetoacetic ester fragment in VIII).

Structure III contains two dissimilar asymmetric carbon atoms (1 and 4), and so four optically active forms (and two racemic modifications) are possible. All are known, and correspond to the menthones and *iso*menthones; these are geometrical isomers, each one existing as a pair of enantiomorphs. The configurations have been assigned on physical evidence; the *cis*-isomer has the higher refractive index and density (Auwers–Skita rule; see §5 x. IV).

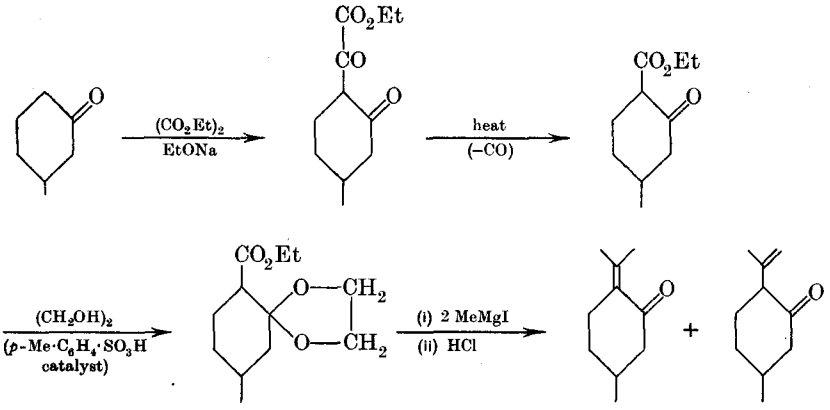


§17. **(±)-Pulegone,** $C_{10}H_{16}O$, b.p. 221–222°. This occurs in pennyroyal oils. Pulegone contains one double bond, and behaves as a ketone. On reduction, pulegone first gives menthone and this, on further reduction, gives menthol. When oxidised with permanganate, pulegone forms acetone and β-methyladipic acid (Semmler, 1892); when boiled with aqueous ethanolic potassium hydroxide, acetone and 3-methyl*cyclo*hexanone are obtained (Wallach, 1896). These reactions show that pulegone is *p*-menth-4(8)-en-3-one.



pulegone

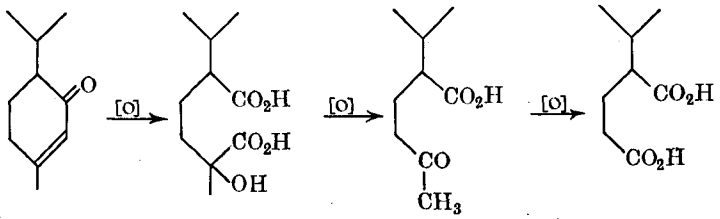
This structure has been confirmed by synthesis, starting from 3-methyl-*cyclo*hexanone (Black *et al.*, 1956: *cf.* menthone, §16).



cyclic ketal pulegone *iso*pulegone

*iso*Pulegone can be isomerised to pulegone by alkaline reagents (Kon *et al.*, 1927), and Black *et al.* found that, on treating their mixture with sodium ethoxide, the resulting compound was pure pulegone.

§18. (–)-Piperitone, $C_{10}H_{16}O$, b.p. 232–233°/768 mm. This occurs in eucalyptus oils, and is a valuable source of menthone and thymol. Piperitone contains one double bond, and behaves as a ketone. Piperitone, on catalytic hydrogenation (nickel), gives menthone in almost quantitative yield; on oxidation with ferric chloride, thymol is obtained (Smith *et al.*, 1920). These reactions show that piperitone is *p*-menthene-3-one, but do



I II III IV

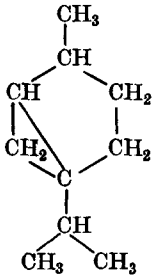
not show the position of the double bond. This had been shown by Schimmel (1910), who found that on oxidation with alkaline permanganate, piperitone gave α-hydroxy-α-methyl-α′-*iso*propyladipic acid, II, γ-acetyl-α-*iso*propylbutyric acid, III, and α-*iso*propylglutaric acid, IV. These results can be explained only if piperitone is *p*-menth-1-en-3-one, I. This structure for piperitone has been confirmed by various syntheses (*e.g.*, Henecka,

1948; Birch *et al.*, 1949). Bergmann *et al.* (1959) have shown that piperitone is formed directly by the condensation of mesityl oxide with methyl vinyl ketone.

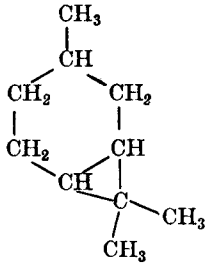
BICYCLIC MONOTERPENES

§19. Introduction. The bicyclic monoterpenes may be divided into three classes according to the size of the *second* ring, the first being a six-membered ring in each class.

Class I (6- + 3-membered ring).

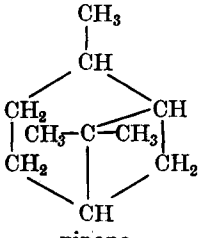


thujane



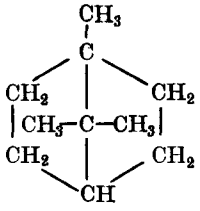
carane

Class II (6- + 4-membered ring).

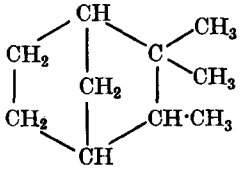


pinane

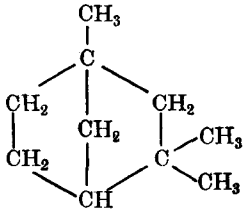
Class III (6- + 5-membered ring).



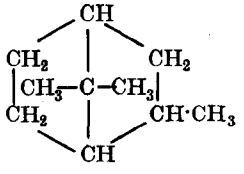
camphane



*iso*camphane



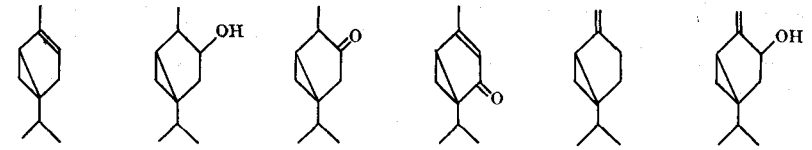
fenchane



*iso*bornylane

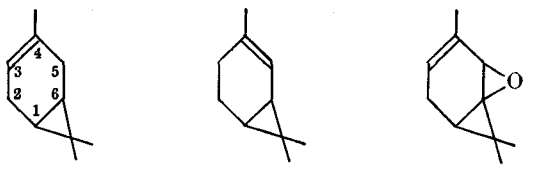
It is important to note that the two rings do not lie in one plane, but are almost perpendicular to each other (see, *e.g.*, §23b).

§20. Thujone and its derivatives. The members of this group which occur naturally are the following:



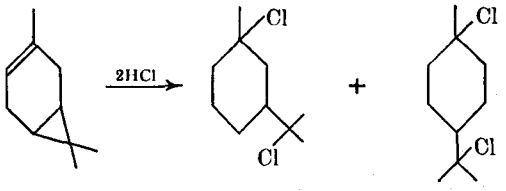
α-thujene thujyl alcohol thujone umbellulone sabinene sabinol

§21. Carane and its derivatives. It appears that only three carane derivatives occur naturally:



car-3-ene car-4-ene car-3-ene-5:6-epoxide

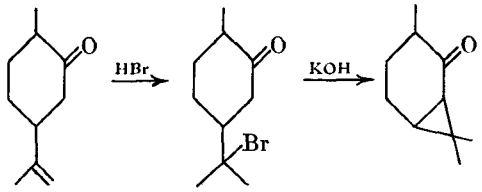
Car-3-ene occurs in Swedish pine needle oil. It is a liquid, b.p. 170°; when treated with hydrogen chloride it forms a mixture of sylvestrene dihydrochloride (see §15) and dipentene dihydrochloride (§13).



(+)-Car-4-ene, b.p. 165·5–167°/707 mm., occurs in various essential oils. It forms sylvestrene dihydrochloride on treatment with hydrogen chloride (§15).

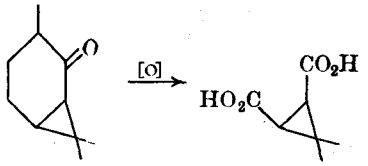
Car-3-ene-5 : 6-epoxide, b.p. 83–85°/14 mm., occurs in certain essential oils.

Carone, b.p. 99–100°/15 mm., is a synthetic compound, and is of some importance because of its relationship to carane. It was first prepared by

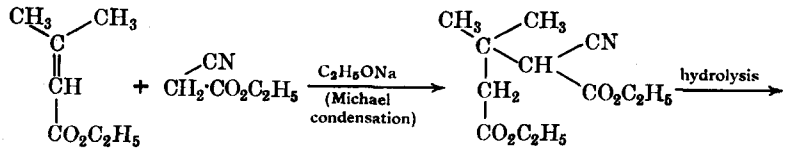


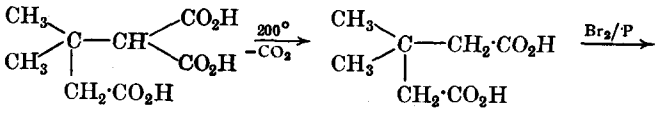
dihydrocarvone carone

Baeyer *et al.* (1894) by the action of hydrogen bromide on dihydrocarvone, which was then treated with ethanolic potassium hydroxide, whereupon carone was obtained.

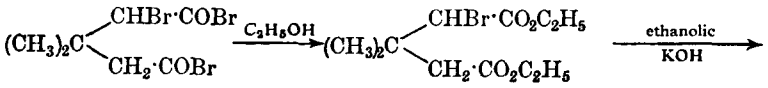


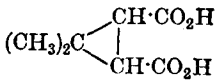
The structure of carone was established by Baeyer *et al.* (1896), who obtained caronic acid on oxidation of carone with permanganate. Baeyer suggested that caronic acid was a *cyclo*propane derivative, and this was confirmed by synthesis (Perkin, junior, and Thorpe, 1899), starting with ethyl β : β-dimethylacrylate and ethyl cyanoacetate.





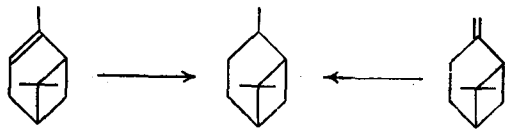
β:β-dimethylglutaric acid





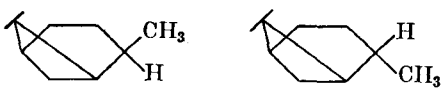
An interesting point about carone is that its ultraviolet absorption spectrum shows similarities to that of α : β-unsaturated ketones (Klotz, 1941).

§22. Pinane and its derivatives. Pinane, the parent compound of this group, is a synthetic substance which may be prepared by the catalytic hydrogenation (nickel or platinum) of either α- or β-pinene. Pinane exists



α-pinene pinane β-pinene

in two geometrical isomeric forms, ***cis*** and ***trans***, and each of these exists as a pair of enantiomorphs.

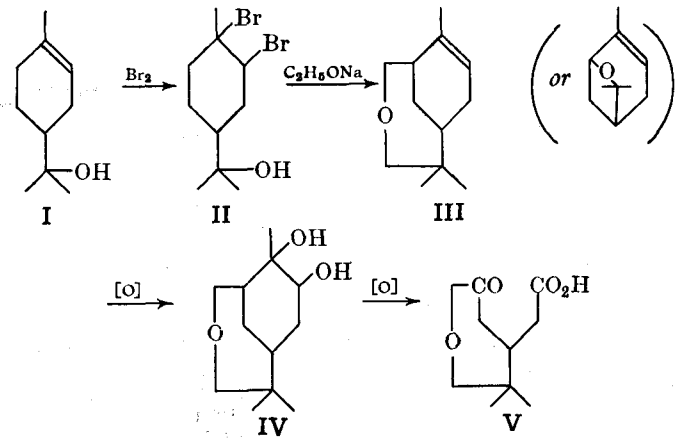


cis *trans*

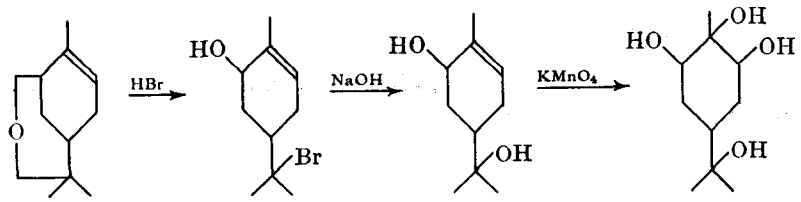
§22a. α-Pinene. This is the most important member of the pinane class. It occurs in both the (+)- and (−)-forms in all turpentine oils; it is a liquid, b.p. 156°.

The analytical evidence for the structure of α-pinene may conveniently be divided into two sections, each section leading independently to the structure, and the two taken together giving very powerful evidence for the structure assigned.

Method 1. The molecular formula of α-pinene is $C_{10}H_{16}$, and since α-pinene adds on two bromine atoms, one double bond is present in the molecule. Thus the parent hydrocarbon is $C_{10}H_{18}$, and since this corresponds to the general formula C_nH_{2n-2} the general formula of compounds containing *two* rings, it therefore follows that α-pinene is bicyclic (Wallach, 1887–1891). In the preparation of α-pinene nitrosochloride (by the action of nitrosyl chloride on α-pinene) the by-products which were formed were steam distilled, and the compound *pinol*, $C_{10}H_{16}O$, was thereby obtained. Pinol adds on one molecule of bromine to form pinol dibromide, and so pinol contains one double bond. Furthermore, the action of lead hydroxide on pinol dibromide converts the latter into pinol glycol, $C_{10}H_{16}O(OH)_2$, and this, on oxidation, gives terpenylic acid (Wallach *et al.*, 1889). Pinol (III) is also obtained by the action of sodium ethoxide on α-terpineol dibromide, II (Wallach, 1893). Wagner (1894) showed that the oxidation of pinol with permanganate gives pinol glycol (IV), which is further oxidised to terpenylic acid (V). All these facts can be explained as follows, based on I being the structure of α-terpineol (see also §11).



Support for the structure given for pinol (III) is obtained from the fact that oxidation of *sobrerol* (pinol hydrate) produces a tetrahydric alcohol, *sobrerythritol*. Sobrerol itself is readily prepared by the action of hydrogen bromide on pinol, followed by sodium hydroxide. These reactions may thus be formulated:

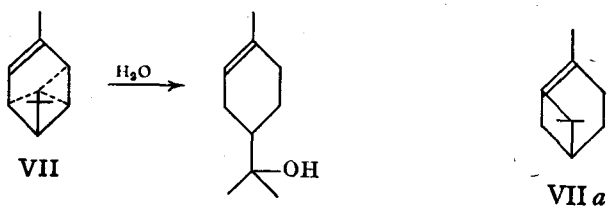


Thus, if the formula for α-pinene is VI, then the formation of the above substances can be explained. This structure also accounts for other reactions of α-pinene, *e.g.*, its ready hydration to α-terpineol (see later).

VI

Although the Wagner formula (VI) for α-pinene readily explains all the facts, there is no *direct* evidence for the existence of the *cyclo*butane ring. Such evidence was supplied by Baeyer (1896). This is described in method 2.

Method 2. As in *method* 1, α-pinene was shown to be bicyclic. When treated with ethanolic sulphuric acid, α-pinene is converted into α-terpineol (Flavitzky, 1879). Therefore α-pinene contains a six-membered ring and another ring (since it is bicyclic), the carbon skeleton of pinene being such as to give α-terpineol when this second ring opens. Since, in the formation of α-terpineol, one molecule of water is taken up and the hydroxyl group becomes attached to C_8, this suggests that the C_8 of α-terpineol is involved in forming the second ring in α-pinene. There are three possible points of union for this C_8, resulting in two three-membered and one four-membered ring (see VII); at the same time the position of the double bond in α-pinene is also shown by the conversion into α-terpineol (I).



VII I VII*a*

A point of interest here is that there are actually *four* possible points of union for C_8, the three shown in VII and the fourth being at the double bond to form a four-membered ring (VII*a*). This one, however, was rejected on the grounds of *Bredt's rule* (1924) which states that a double bond cannot be formed by a carbon atom occupying the bridge-head (of a bicyclic system). The explanation for this rule is that structures such as VII*a* have a large amount of strain.

This second ring was shown to be four-membered by Baeyer (1896), who carried out the following series of reactions.

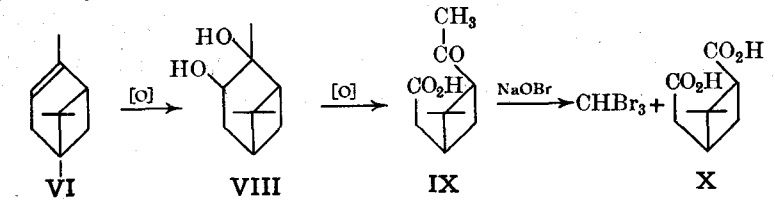
$$\underset{C_{10}}{\underset{VI}{\alpha\text{-Pinene}}} \xrightarrow[KMnO_4]{1\%\ alk.} \underset{C_{10}}{\underset{VIII}{\text{Pinene glycol}}} \xrightarrow[KMnO_4]{warm\ alk.} \underset{C_{10}}{\underset{IX}{\text{Pinonic acid}}}$$

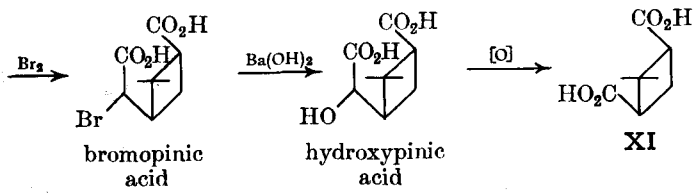
$$\xrightarrow{NaOBr} \underset{C_9}{\underset{X}{\text{Pinic acid}}} + CHBr_3 \xrightarrow[(iii)\ chromic\ acid]{(i)\ Br_2\ (ii)\ Ba(OH)_2} \underset{C_8}{\underset{XI}{\textit{cis}\text{-Norpinic acid}}}$$

Pinene glycol, $C_{10}H_{16}(OH)_2$, is produced by hydroxylation of the double bond in α-pinene, and pinonic acid, $C_{10}H_{16}O_3$, is produced by scission of the glycol bond. Pinonic acid was shown to be a saturated keto-monocarboxylic

acid. The formation of pinic acid, $C_9H_{14}O_4$, and bromoform, indicates the presence of an acetyl group in pinonic acid. Pinic acid, which was shown to be a saturated dicarboxylic acid, on treatment with bromine, then barium hydroxide, and finally the product oxidised with chromic acid, gives *cis*-norpinic acid, $C_8H_{12}O_4$. This was shown to be a saturated dicarboxylic acid, and so its formula may be written $C_6H_{10}(CO_2H)_2$. Furthermore, since α-pinene contains two methyl groups attached to a carbon atom in the second ring (see VII), and it is the *other* ring (the six-membered one containing the double bond) that has been opened by the above oxidation, then norpinic acid (with this second ring intact) contains these two methyl groups. Thus the formula for norpinic acid may be written $(CH_3)_2C_4H_4(CO_2H)_2$. Hence, regarding the methyl and carboxyl groups as substituents, the parent (saturated) hydrocarbon (from which norpinic acid is derived) is C_4H_8. This corresponds to *cyclo*butane, and so norpinic acid is (probably) a dimethyl-*cyclo*butanedicarboxylic acid. On this basis, pinic acid could therefore be a *cyclo*butane derivative with one side-chain of $—CH_2·CO_2H$.

Baeyer therefore *assumed* that pinic and norpinic acids contained a *cyclo*butane ring, and so suggested the following structures to account for the above reactions, accepting structure VI for α-pinene, the structure already proposed by Wagner (1894).



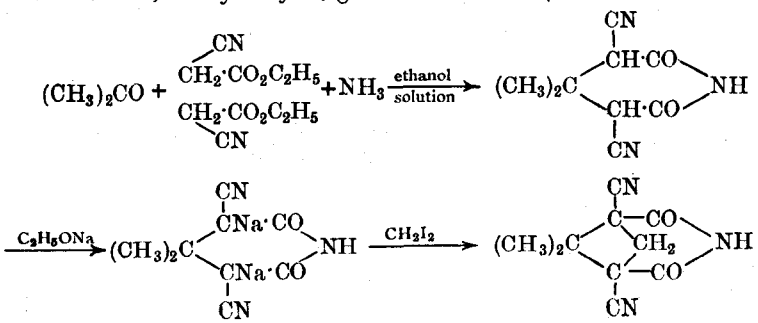


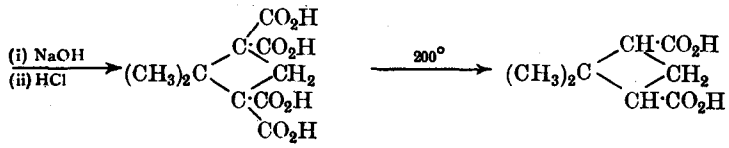
bromopinic acid

hydroxypinic acid

The synthesis of norpinic acid (to confirm the above reactions) proved to be a very difficult problem, and it was not carried out until 1929, when Kerr succeeded with the following ingenious method (apparently the presence of the *gem* dimethyl group prevents closure to form the *cyclo*butane ring).

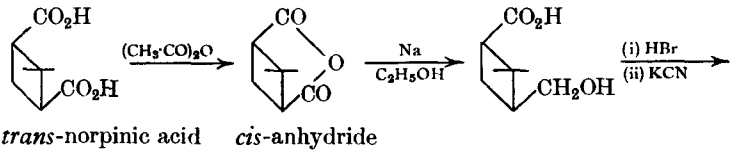
The norpinic acid obtained was the *trans*-isomer; this is readily converted into the *cis*-isomer (the isomer obtained from the oxidation of α-pinene) by heating the *trans* acid with acetic anhydride, whereupon the *cis* anhydride is formed and this, on hydrolysis, gives the *cis* acid (Simonsen *et al.*, 1929).



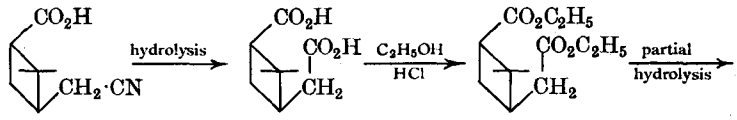


The total synthesis of α-pinene has now been carried out in the following way. Guha *et al.* (1937) synthesised pinic acid from norpinic acid, and Rao (1943) synthesised pinonic acid from synthetic pinic acid.

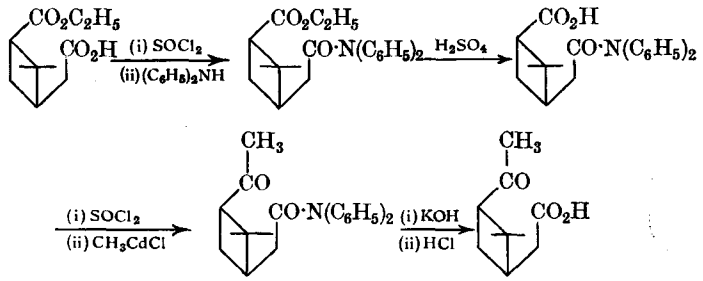
Ruzicka *et al.* (1920–1924) had already synthesised α-pinene starting from pinonic acid (obtained by the oxidation of α-pinene). Thus we now have



trans-norpinic acid *cis*-anhydride

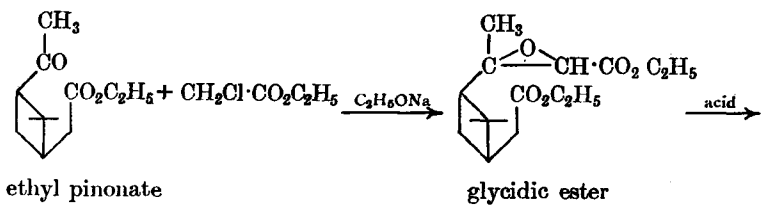


pinic acid

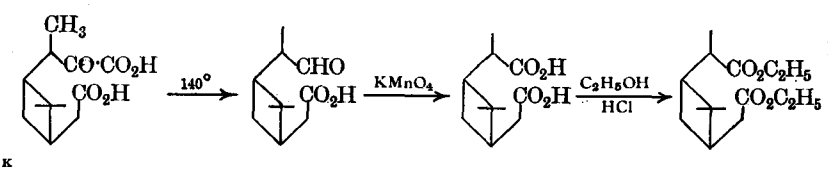


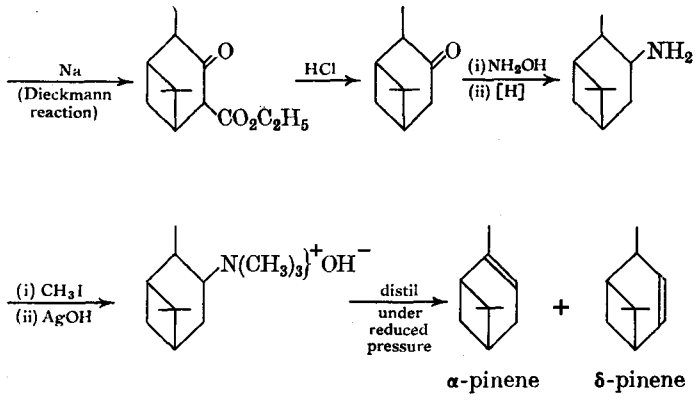
trans-pinonic acid

a total synthesis of α-pinene. Ruzicka's synthesis makes use of the Darzens glycidic ester synthesis (see Vol. I); the steps are:

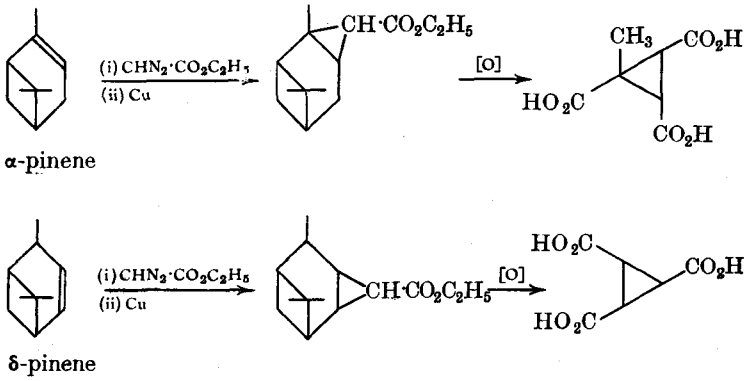


ethyl pinonate glycidic ester





The final step gives a mixture of two compounds, α- and δ-pinene. The former was identified by the preparation of the nitrosochloride; this proves that one of the products is α-pinene, but does not prove which is α and which is δ. These are differentiated by consideration of the analytical evidence; the following evidence also supports the structure given for α-pinene. This evidence is based on the fact that diazoacetic ester combines with compounds containing a double bond to form pyrazoline derivatives, and these, on heating alone or with copper powder, decompose to produce *cyclo*propane derivatives (see also §2a. XII). When the two pinenes were subjected to this



treatment, and the resulting compounds oxidised, α-pinene gave 1-methyl-*cyclo*propane-1 : 2 : 3-tricarboxylic acid, and δ-pinene *cyclo*propane-1 : 2 : 3-tricarboxylic acid. These products are in accord with the structures assigned to α- and δ-pinene.

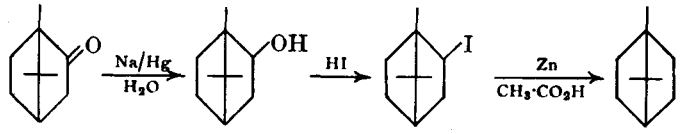
Examination of the α-pinene structure shows that two dissimilar asymmetric carbon atoms are present; thus two pairs of enantiomorphs are possible. In practice, however, only one pair is known. This is due to the fact that the four-membered ring can only be fused to the six-membered one in the *cis*-position; *trans* fusion is impossible. Thus only the enantiomorphs of the *cis*-isomer are known.

Isomeric with α-pinene are β- and δ-pinene; the former occurs naturally, the latter is synthetic (see Ruzicka's synthesis). Crowley (1962) has obtained a small amount of β-pinene by irradiating a one per cent. ethereal solution of myrcene (§4) with ultraviolet light. This is of some interest in connection with the biosynthesis of terpenes (see §32a).

β-pinene δ-pinene

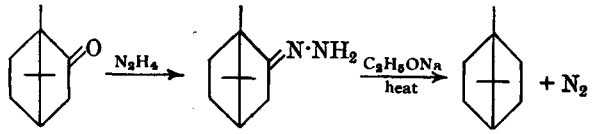
§23. Camphane and its derivatives. Camphane, $C_{10}H_{18}$, is a synthetic compound, and may be prepared from camphor, *e.g.*,

(i) By reduction of camphor to a mixture of borneols (§23b), these then converted to the bornyl iodides which are finally reduced to camphane (Aschan, 1900).



camphor camphane

(ii) Camphor may also be converted into camphane by means of the Wolff–Kishner reduction (see also Vol. I).



Camphane is a solid, m.p. 156°; it is optically inactive.

§23a. Camphor. This occurs in nature in the camphor tree of Formosa and Japan. It is a solid, m.p. 179°, and is optically active; the (+)- and (−)-forms occur naturally, and so does racemic camphor, which is the usual form of synthetic camphor (from α-pinene; see later).

A tremendous amount of work was done before the structure of camphor was successfully elucidated; in the following account only a small part of the work is described, but it is sufficient to justify the structure assigned to camphor.

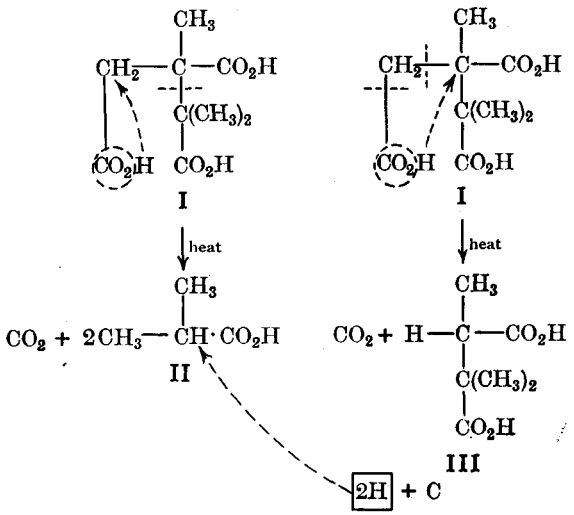
The molecular formula of camphor is $C_{10}H_{16}O$, and the general reactions and molecular refractivity of camphor show that it is saturated. The functional nature of the oxygen atom was shown to be oxo by the fact that camphor formed an oxime, etc., and that it was a keto group was deduced from the fact that oxidation of camphor gives a *dicarboxylic* acid containing 10 carbon atoms; a *monocarboxylic* acid containing 10 carbon atoms cannot be obtained (this type of acid would be expected if camphor contained an *aldehyde* group). From the foregoing facts it can be seen that the parent hydrocarbon of camphor has the molecular formula $C_{10}H_{18}$; this corresponds to C_nH_{2n-2}, and so camphor is therefore bicyclic. Camphor contains a $—CH_2·CO—$ group, since it forms an oxime with nitrous acid (*iso*amyl nitrite and hydrogen chloride). Finally, distillation of camphor with zinc chloride or phosphorus pentoxide produces *p*-cymene.

Bredt (1893) was the first to assign the correct formula to camphor (over 30 have been proposed). Bredt based his formula on the above facts and also on the facts that (*a*) oxidation of camphor with nitric acid gives **camphoric acid,** $C_{10}H_{16}O_4$ (Malaguti, 1837); (*b*) oxidation of camphoric acid

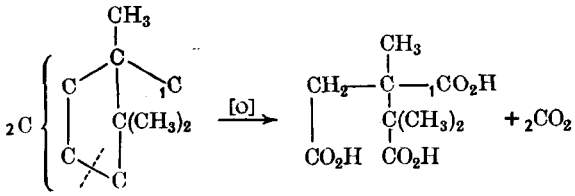
(or camphor) with nitric acid gives **camphoronic acid,** $C_9H_{14}O_6$ (Bredt, 1893).

Since camphoric acid contains the same number of carbon atoms as camphor, the keto group must be in one of the rings in camphor. Camphoric acid is a dicarboxylic acid, and its molecular refractivity showed that it is saturated. Thus, in the formation of camphoric acid from camphor, the ring containing the keto group is opened, and consequently camphoric acid must be a monocyclic compound.

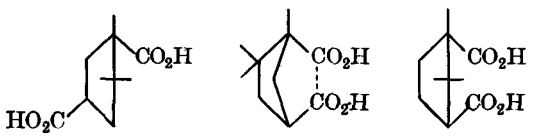
Camphoronic acid was shown to be a saturated tricarboxylic acid, and on distillation at atmospheric pressure, it gave *iso*butyric acid, II, trimethylsuccinic acid, III, carbon dioxide and carbon (and a small amount of some other products). Bredt (1893) therefore *suggested* that camphoronic acid is $\alpha : \alpha : \beta$-trimethyltricarballylic acid, I, since this structure would give the required decomposition products. In the following equations, the left-hand-side molecule is imagined to break up as shown; one molecule of carbon dioxide and two molecules of *iso*butyric acid are produced (but there is a shortage of two hydrogen atoms). The right-hand-side molecule breaks up to form one molecule of trimethylsuccinic acid, one molecule of carbon dioxide, one atom of carbon and *two atoms of hydrogen* which now make up the shortage of the left-hand-side molecule. Thus:



Hence, if camphoronic acid has structure I, then camphoric acid (and camphor) must contain *three methyl groups*. On this basis, the formula of camphoric acid, $C_{10}H_{16}O_4$, can be written as $(CH_3)_3C_5H_5(CO_2H)_2$. The parent (saturated) hydrocarbon of this is C_5H_{10}, which corresponds to C_nH_{2n}, *i.e.*, camphoric acid is a *cyclo*pentane derivative (this agrees with the previous evidence that camphoric acid is monocyclic). Thus the oxidation of camphoric acid to camphoronic acid may be written:

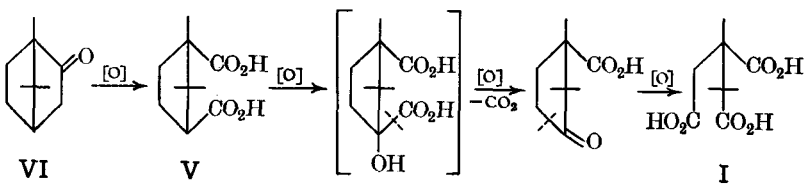


This skeleton, plus one carbon atom, arranged with two carboxyl groups, will therefore be the structure of camphoric acid. Now camphoric anhydride forms only one monobromo derivative (bromine and phosphorus); therefore there is only *one* α-hydrogen atom in camphoric acid. Thus the carbon atom of one carboxyl group must be $_1$C (this is the only carbon atom joined to a tertiary carbon atom). Furthermore, $_1$C must be the carbon of the keto or methylene group in camphor, since it is these two groups which produce the two carboxyl groups in camphoric acid. The problem is now to find the position of the other carboxyl group in camphoric acid. Its position must be such that when the *cyclo*pentane ring is opened to give camphoronic acid, one carbon atom is readily lost. Using this as a working hypothesis, then there are only two reasonable structures for camphoric



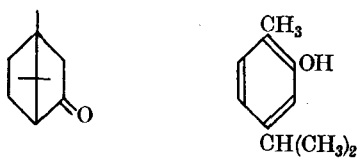
IV IV*a* V

acid, IV and V. IV may be rewritten as IV*a*, and since the two carboxyl groups are produced from the —CH_2·CO— group in camphor, the precursor of IV*a* (*i.e.*, camphor) will contain a six-membered ring with a *gem*-dimethyl group. This structure cannot account for the conversion of camphor into *p*-cymene. On the other hand, V accounts for all the facts given in the foregoing discussion. Bredt therefore assumed that V was the structure of camphoric acid, and that VI was the structure of camphor, and proposed the following reactions to show the relationships between camphor, camphoric acid and camphoronic acid.



VI V I

Bredt, however, realised that if camphor had structure VII, then all the foregoing facts would be equally satisfied, but he rejected VII in favour of VI for a number of reasons. One simple fact that may be used here for

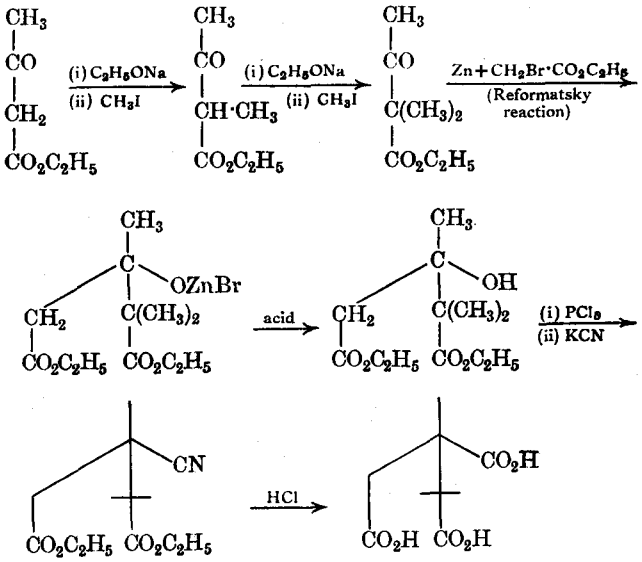


VII VIII

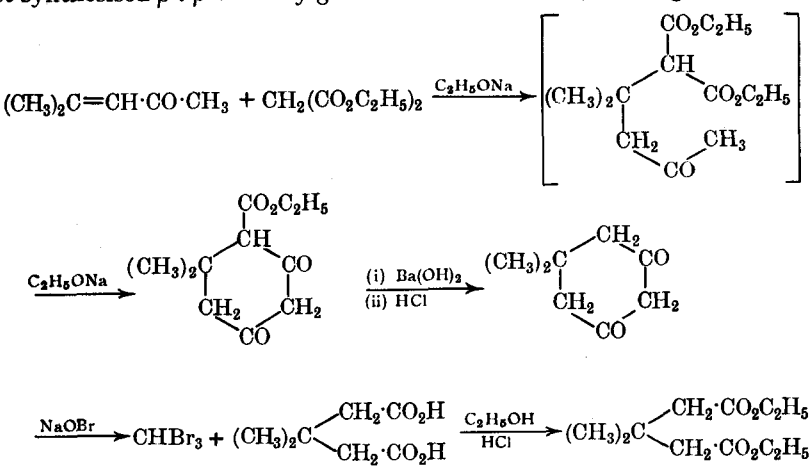
rejection of VII is that camphor gives carvacrol, VIII, when distilled with iodine. The formation of this compound can be expected from VI but not from VII.

Formula VI for camphor was accepted with reserve at the time when Bredt proposed it (in 1893), but by 1903 all the deductions of Bredt were confirmed by the syntheses of camphoronic acid, camphoric acid and camphor.

Synthesis of (±)-camphoronic acid (Perkin, junior, and Thorpe, 1897).

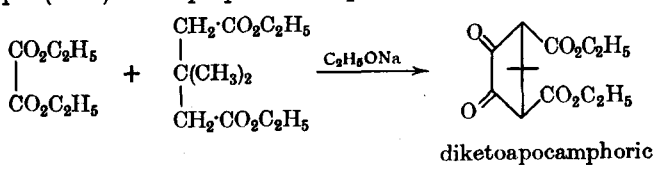


Synthesis of (±)-camphoric acid (Komppa, 1903). Komppa (1899) first synthesised β : β-dimethylglutaric ester as follows, starting with mesityl

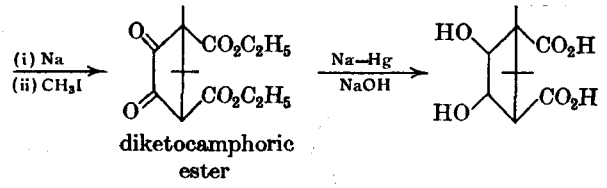


oxide and ethyl malonate. The product obtained was 6 : 6-dimethyl*cyclo*-hexane-2 : 4-dione-1-carboxylic ester (this is produced first by a Michael condensation, followed by a Dieckmann reaction). On hydrolysis, followed by oxidation with sodium hypobromite, β : β-dimethylglutaric acid was obtained (*cf.* carone, §21).

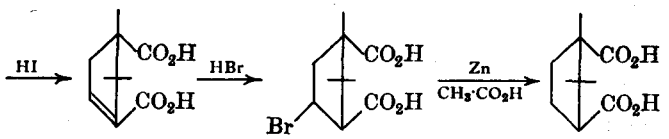
Komppa (1903) then prepared camphoric acid as follows:



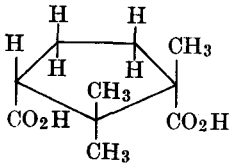
diketoapocamphoric ester



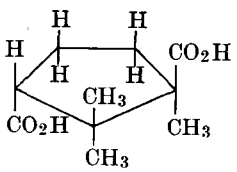
diketocamphoric ester



The structure given for camphoric acid can exist in two geometrical isomeric forms, *cis* and *trans*, neither of which has any elements of symmetry. Thus four optically active forms are possible; all are known, and correspond to the (+)- and (−)-forms of camphoric acid and *iso*camphoric acid. Since camphoric acid forms an anhydride, and *iso*camphoric acid does not, the former is the *cis*-isomer, and the latter the *trans*- (§5 i. IV).

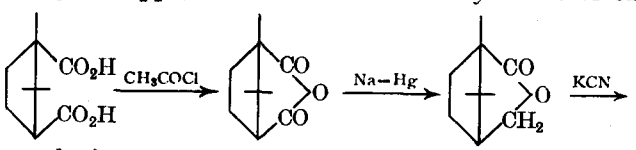


camphoric acid, m.p. 187°

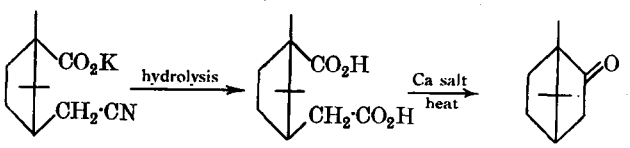


iso-camphoric acid, m.p. 171-172°

Synthesis of camphor (Haller, 1896). Haller started with camphoric acid prepared by the oxidation of camphor, but since the acid was synthesised later by Komppa, we now have a total synthesis of camphor.

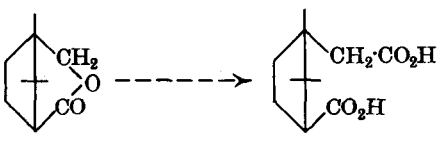


camphoric acid — camphoric anhydride — α-campholide



homocamphoric acid

This is *not* an unambiguous synthesis, since the campholide obtained might have had the structure IX (this is actually β-campholide).

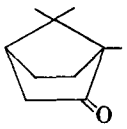


β-campholide IX — X

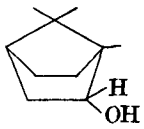
In this case, homocamphoric acid would have had structure X, and this would have given camphor with structure VII which, as we have seen, was rejected. Sauers (1959) has now oxidised camphor directly to α-campholide by means of peracetic acid. It is also of interest to note that Otvös *et al.* (1960) have shown, using labelled —CH_2·C*O_2H (^{14}C), that in the pyrolysis of the calcium salt of homocamphoric acid to camphor, it is the labelled carboxyl group that is lost.

Stereochemistry of camphor. Camphor has two dissimilar asymmetric carbon atoms (the same two as in camphoric acid), but only one pair of enantiomorphs is known. This is due to the fact that only the *cis*-form is possible; *trans* fusion of the *gem*-dimethylmethylene bridge to the *cyclo*hexane ring is impossible. Thus only the enantiomorphs of the *cis*-isomer are known (*cf.* α-pinene, §22a).

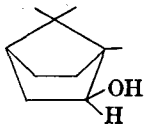
Camphor and its derivatives exist in the boat conformation. Since the *gem*-dimethyl bridge must be *cis*, the *cyclo*hexane ring must have the boat form (see also §23b for the usual way of drawing these conformations; the viewing point is different):



camphor

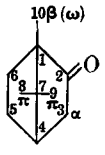


borneol



*iso*borneol

Some derivatives of camphor. The positions of substituent groups in camphor are indicated by numbers or by the Greek letters α (= 3), β or ω (= 10) and π (= 8 or 9). When (+)-camphor is heated with bromine at 100°, α-bromo-(+)-camphor is produced. This, on warming with sulphuric acid, is converted into α-bromo-(+)-camphor-π-sulphonic acid which,



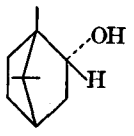
on reduction, forms (+)-camphor-π-sulphonic acid. (±)-Camphor-π-sulphonic acid is obtained by the sulphonation of (+)-camphor with fuming sulphuric acid; under these conditions, (+)-camphor is racemised. On the other hand, sulphonation of (+)-camphor with sulphuric acid in acetic anhydride solution produces (+)-camphor-β-sulphonic acid. These various (+)-camphorsulphonic acids are very valuable reagents for resolving racemic bases (§10 iv. II).

Commercial preparation of camphor. Synthetic camphor is usually obtained as the racemic modification. The starting material is α-pinene, and the formation of camphor involves the Wagner–Meerwein rearrangements (see §23d). Scheme (i) is the earlier method, and (ii) is the one that is mainly used now.

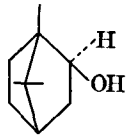
(i) α-Pinene $\xrightarrow[10°]{\text{HCl gas}}$ Bornyl chloride $\xrightarrow[-\text{HCl}]{\text{CH}_3\cdot\text{CO}_2\text{Na}}$ Camphene $\xrightarrow[\text{H}_2\text{SO}_4]{\text{CH}_3\cdot\text{CO}_2\text{H}}$ *iso*Bornyl acetate $\xrightarrow{\text{NaOH}}$ *iso*Borneol $\xrightarrow{\text{C}_6\text{H}_5\cdot\text{NO}_2}$ Camphor

(ii) α-Pinene $\xrightarrow[10°]{\text{HCl gas}}$ Bornyl chloride $\xrightarrow[-\text{HCl}]{\text{CH}_3\cdot\text{CO}_2\text{Na}}$ Camphene $\xrightarrow{\text{H}\cdot\text{CO}_2\text{H}}$ *iso*Bornyl formate $\xrightarrow{\text{NaOH}}$ *iso*Borneol $\xrightarrow[\text{Ni; } 200°]{\text{O}_2}$ Camphor

§23b. Borneols, $C_{10}H_{18}O$. There are two stereoisomeric compounds of the formula $C_{10}H_{18}O$; these correspond to **borneol** and ***iso*borneol,** and both are known in the (+)- and (−)-forms. The borneols occur widely distributed in essential oils, but it appears that the *iso*borneols have been isolated from only one essential oil. Borneol and *iso*borneol are secondary alcohols, and the evidence now appears to be conclusive that borneol has the *endo*-configuration in which the *gem*-dimethyl bridge is above the plane



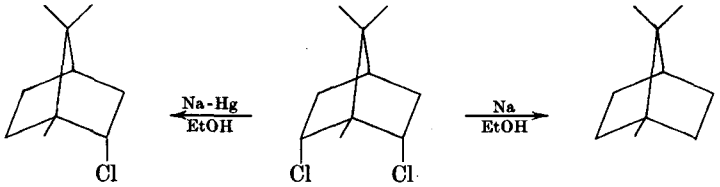
borneol
m.p. 208·5°



*iso*borneol
m.p. 217°

of the *cyclo*hexane ring and the hydroxyl group is below the plane. *iso*-Borneol has the *exo*-configuration in which the bridge and the hydroxyl group are both above the plane of the *cyclo*hexane ring (see also §23a).

Kwart *et al.* (1956) have now obtained direct evidence on the configuration of bornyl chloride. Bornyl dichloride (I), the structure of which has been established by Kwart (1953), is converted into bornyl chloride (II) by sodium amalgam and ethanol, and into camphane (III) by sodium and ethanol.

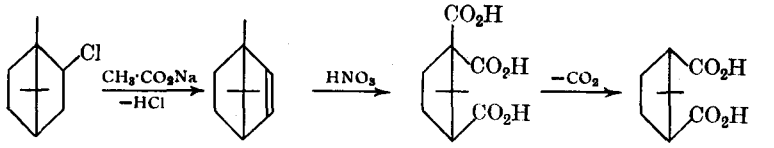


II I III

Both borneol and *iso*borneol are produced when camphor is reduced, but the relative amounts of each are influenced by the nature of the reducing agent used, *e.g.*, electrolytic reduction gives mainly borneol, whereas catalytic hydrogenation (platinum) gives mainly *iso*borneol; *iso*borneol is also the main product when aluminium *iso*propoxide is used as the reducing agent (the Meerwein–Ponndorf–Verley reduction; see Vol. I). Borneol is converted into a mixture of bornyl and *iso*bornyl chlorides by the action of phosphorus pentachloride. Borneol and *iso*borneol are both dehydrated to camphene (§23c), but the dehydration occurs more readily with *iso*borneols than with borneol. Both alcohols are oxidised to camphor, but whereas borneol can be dehydrogenated to camphor by means of a copper catalyst, *iso*borneol cannot.

§23c. Camphene and Bornylene. Camphene, $C_{10}H_{16}$, m.p. 51–52°, occurs naturally in the (+)-, (−)- and (±)-forms. It may be prepared by the removal of a molecule of hydrogen chloride from bornyl and *iso*bornyl chlorides by means of sodium acetate, or by the dehydration of the borneols with potassium hydrogen sulphate. These methods of preparation suggest that camphene contains a double bond, and this is supported by the fact that camphene adds on one molecule of bromine or one molecule of hydrogen chloride. Oxidation of camphene with dilute nitric acid produces carboxyapocamphoric acid, $C_{10}H_{14}O_6$, and apocamphoric acid, $C_9H_{14}O_4$ (Marsh *et al.*, 1891). The formation of the former acid, which contains the same number of carbon atoms as camphene, implies that the double bond in camphene is in a ring; and the fact that carboxyapocamphoric acid is converted into

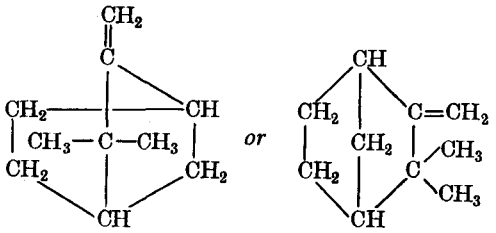
apocamphoric acid when heated above its melting point implies that the former contains two carboxyl groups attached to the same carbon atom



bornyl chloride — *camphene* I — carboxyapocamphoric acid — apocamphoric acid

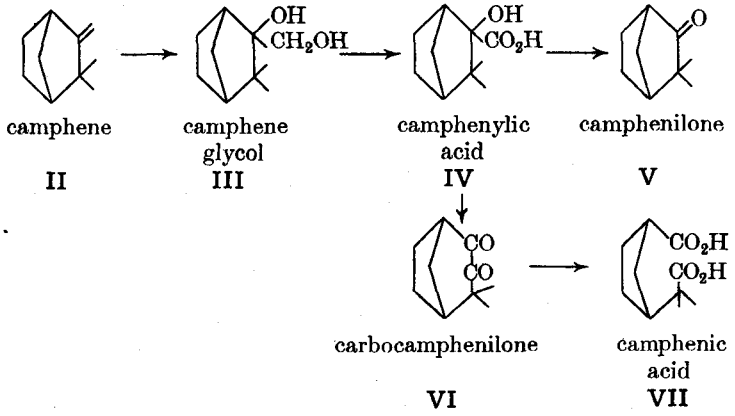
(*cf.* malonic ester syntheses). These facts were explained by giving camphene the formula shown (I). The structure of apocamphoric acid was later proved by synthesis (Komppa, 1901; *cf.* camphoric acid, §23a).

This structure for camphene, however, was opposed by Wagner. The oxidation of camphene with dilute permanganate gives camphene glycol, $C_{10}H_{16}(OH)_2$ [Wagner, 1890]. This glycol is saturated, and so camphene is a bicyclic compound (so, of course, is structure I). On further oxidation of camphene glycol, Wagner (1896, 1897) obtained camphenic acid, $C_{10}H_{16}O_4$ (a dibasic acid), and camphenylic acid, $C_{10}H_{16}O_3$ (a hydroxy-monobasic acid), which, on oxidation with lead dioxide, gave camphenilone, $C_9H_{14}O$ (a ketone). According to Wagner, it was difficult to explain the formation of these compounds if camphene had structure I. Wagner (1899) therefore suggested that camphene is formed by a molecular rearrangement when the borneols or bornyl chlorides are converted into camphene, and proposed structure II for camphene (see also §23d).



II

With this formula, the formation of camphene glycol, camphenylic acid and camphenilone could be explained as follows:



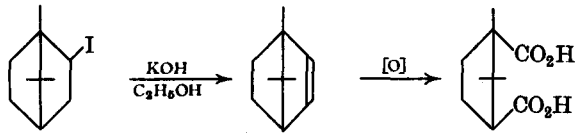
camphene II — camphene glycol III — camphenylic acid IV — camphenilone V

carbocamphenilone VI — camphenic acid VII

Although it was easy to explain the formation of III, IV and V, it was difficult to explain the formation of VII. The formation of VII was ex-

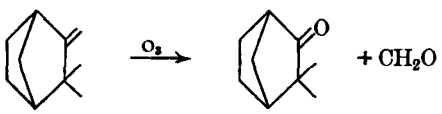
plained by later workers, who suggested it was produced *via* carbocamphenilone, VI. Another difficulty of the camphene formula, II, is that it does not explain the formation of apocamphoric acid when camphene is oxidised with nitric acid (see above). The course of its formation has been suggested by Komppa (1908, 1911), who proposed a mechanism involving a Wagner rearrangement.

Structure II for camphene is supported by the fact that treatment of bornyl iodide with ethanolic potassium hydroxide at 170° gives **bornylene,** $C_{10}H_{16}$ (m.p. 98°), as well as camphene (Wagner *et al.*, 1899). Bornylene is readily oxidised by permanganate to camphoric acid; it therefore follows that bornylene has the structure I, the structure originally assigned to camphene; no rearrangement occurs in the formation of bornylene.



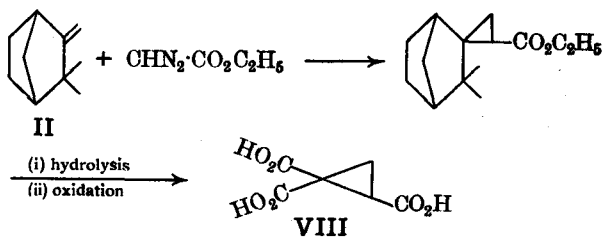
bornyl iodide — bornylene — camphoric acid

Ozonolysis of camphene gives camphenilone and formaldehyde (Harries *et al.*, 1910); these products are in keeping with the Wagner formula for camphene.



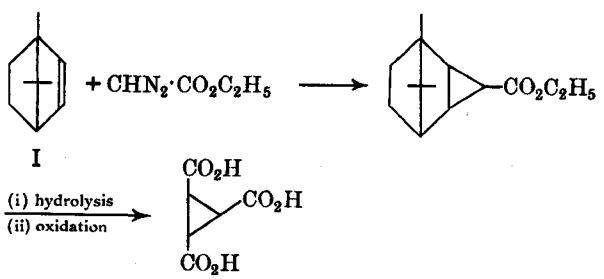
II V

Further support for this structure for camphene is afforded by the work of Buchner *et al.* (1913). These workers showed that camphene reacts with diazoacetic ester, and when the product is hydrolysed and then oxidised,



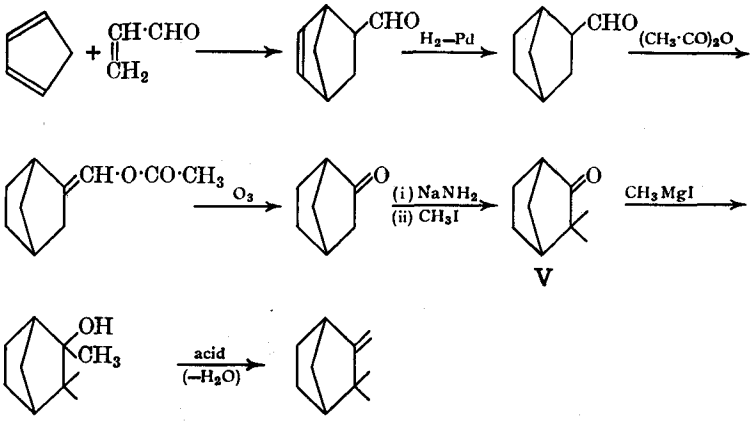
VIII

*cyclo*propane-1 : 1 : 2-tricarboxylic acid, VIII, is produced. VIII is to be expected from structure II, but not from I; I (bornylene) would give *cyclo*propane-1 : 2 : 3-tricarboxylic acid, IX.



IX

Lipp (1914) has synthesised camphenic acid (VII), and showed that it has the structure assigned to it by Wagner. Finally, camphene has been synthesised as follows (Diels and Alder, 1928–1931).



V

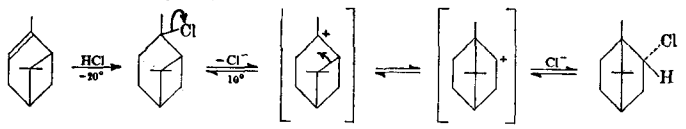
II

§23d. Wagner–Meerwein rearrangements. Wagner, as we have seen, proposed a molecular rearrangement to explain the formation of camphene from the borneols and bornyl chlorides. Wagner also recognised that a molecular rearrangement occurred when α-pinene was converted into bornyl chloride. Many other investigations concerning rearrangements in the terpene field were carried out by Meerwein and his co-workers, *e.g.*, when α-pinene is treated in ethereal solution at −20° with hydrogen chloride, the product is pinene hydrochloride. This is unstable, and if the temperature is allowed to rise to about 10°, the pinene hydrochloride rearranges to bornyl chloride (Meerwein *et al.*, 1922). Rearrangements such as these which occur with bicyclic monoterpenes are known as *Wagner–Meerwein rearrangements*. Furthermore, Meerwein extended the range of these rearrangements to compounds outside bicyclic terpenes; these compounds were monocyclic. Finally, the range was extended to acyclic compounds, the classical example being that of *neo*pentyl into *t*-pentyl compounds (Whitmore *et al.*, 1932–).

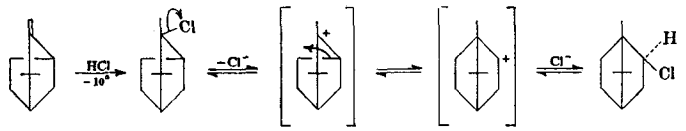
All of these rearrangements conform to a common pattern, ionisation to a carbonium ion followed by rearrangement. Most rearrangements in the terpene field involve a change in ring structure, and in a few cases the migration of a methyl group. All of these rearrangements are examples of the 1,2-shifts (Vol. I, Ch. V).

The following are examples, and the details of the mechanisms are discussed later; (but see Vol. I for a discussion of example v).

(i) *The conversion of α-pinene hydrochloride into bornyl chloride.*

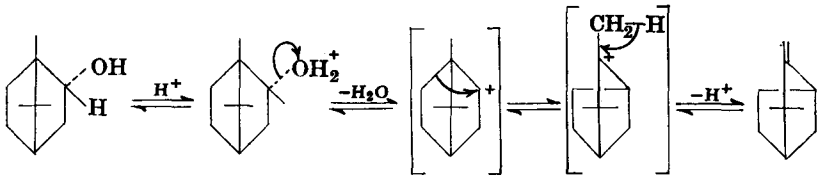


(ii) *The conversion of camphene hydrochloride into* isobornyl chloride.

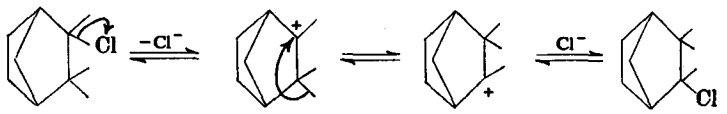


(i) and (ii) are of particular interest since both appear to proceed through the same carbonium ion. Why the epimers should be obtained is not certain (but see later).

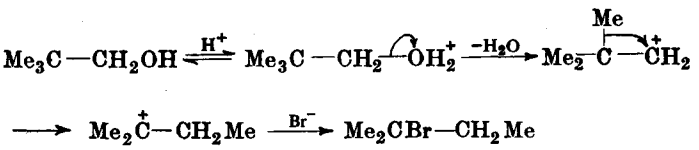
(iii) *The dehydration of borneol to camphene (with acids).*



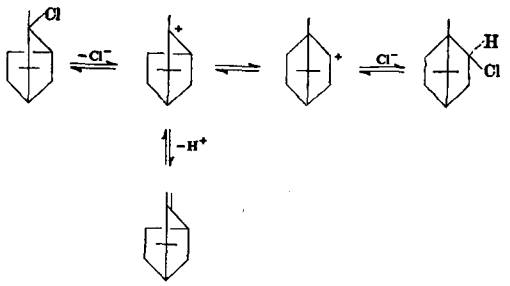
(iv) *The racemisation of camphene hydrochloride.*



(v) *Rearrangements in the* neo*pentyl system; e.g.*, the action of hydrobromic acid on *neo*pentyl alcohol to give *t*-pentyl bromide.



Evidence for the intermediate formation of a carbonium ion in the Wagner–Meerwein rearrangement. Meerwein *et al.* (1922), in their detailed investigation of the reversible conversion of camphene hydrochloride into *iso*bornyl chloride (example ii), concluded that the first step was ionisation, and this was then followed by rearrangement of the carbonium ion:



Their evidence for this mechanism was that the rate of the rearrangement was first order, and that the rate depended on the nature of the solvent, the rate being faster the greater the ionising power of the solvent. The order observed for some solvents was:

$$SO_2 > MeNO_2 > MeCN > PhOMe > PhBr > PhH > Et_2O$$

This dependence of rate on solvent was more clearly shown by also studying the solvolysis rates of triphenylmethyl chloride in the same solvents. It was found that the rate of the rearrangement of camphene hydrochloride was faster in those solvents in which triphenylmethyl chloride undergoes solvolysis more readily. Meerwein also found that the rearrangement was strongly catalysed by Lewis acids such as stannic chloride, ferric chloride, etc. All of these form complexes with triphenylmethyl chloride. Furthermore, halides such as phosphorus trichloride and silicon tetrachloride, which do not form complexes with triphenylmethyl chloride, did not catalyse the rearrangement. Further evidence

by Meerwein *et al.* (1927) and by Ingold (1928) also supports the mechanism given above.

Meerwein, however, recognised a difficulty in his proposed mechanism. The carbonium ion formed in the rearrangement of camphene hydrochloride would presumably be the same as that formed in the rearrangement of pinene hydrochloride to bornyl chloride (example i). The reason why the epimers are obtained is not certain; one possibility is that the ions are *not* the same, and as we shall see later, the ions are not identical if we assume there is neighbouring group participation producing a non-classical carbonium ion.

Bartlett *et al.* (1937, 1938) showed that the rearrangement of camphene hydrochloride in non-hydroxylic solvents is strongly catalysed by hydrogen chloride, and pointed out that the formation of *iso*bornyl chloride requires a Walden inversion at the new asymmetric carbon atom. According to these authors, the function of the hydrochloric acid is to help the ionisation of the chloride ion (from the camphene hydrochloride). Evidence for this is that phenols have a catalytic effect on the rearrangement rate of camphene hydrochloride, and that the order of this catalytic activity of substituted phenols is the same as the order of the increase in acid strength of hydrogen chloride which phenols promote in dioxan as solvent. These catalytic effects were explained by Bartlett *et al.* (1941) as being due to hydrogen bonding between the phenolic hydroxyl group and the receding chloride ion.

Nevell *et al.* (1939) suggested that the type of resonance hybrid Z is involved in the rearrangement. Thus the hydrogen chloride-catalysed reaction in the inert solvents used would produce an ion-pair $[Z^+][HCl_2^-]$ (§2e. III). Z^+ can

Z

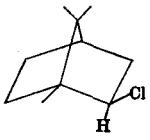
now react with HCl_2^- at position 1 to regenerate camphene hydrochloride, or at position 2 to give *iso*bornyl chloride. This interpretation is supported by experimental work.

(i) Nevell *et al.* found that the rate of radioactive chlorine (^{36}Cl) exchange between HCl* and camphene hydrochloride is 15 times faster than the rate of rearrangement to *iso*bornyl chloride. It therefore follows that the rate-determining step of the rearrangement is *not* the ionisation step, but is the reaction of the bridged-ion with HCl_2^- at position 2. It also follows, from the principle of microscopic reversibility (Vol. I), that the rate-determining step of the rearrangement of *iso*bornyl chloride back to camphene hydrochloride is the reaction with hydrogen chloride to produce the ion-pair directly.

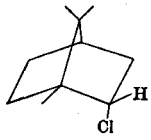
(ii) On the basis of the bridged-ion being an intermediate in the rearrangement in inert solvents and also for solvolytic reactions of both camphene hydrochloride and *iso*bornyl chloride, then both isomers should give the *same* products Meerwein *et al.* (1922) found that methanolysis, in the cold, of camphene hydrochloride gave at first the *t*-methyl ether (attack at position 1) and this, on long standing, gave *iso*bornyl methyl ether. *iso*Bornyl chloride also gave *iso*bornyl methyl ether, but in this case the reaction was slower. These results can be explained by the presence of the liberated hydrogen chloride which would make the methanolysis reversible.

(iii) Neighbouring group participation in solvolytic reactions of camphene hydrochloride would be expected to accelerate these reactions (anchimeric assistance) as compared with the formation of a classical carbonium ion intermediate. This will be so because the formation of the bridge will assist the expulsion of the chloride ion. Hughes, Ingold *et al.* (1951) have found that the ethanolysis of camphene hydrochloride is 6000 times faster (at 0°) than the corresponding reaction with *t*-butyl chloride. Also, from the reaction rates of the solvolysis of 1-chloro-1-methyl*cyclo*pentane, it followed that camphene hydrochloride is 370 times more reactive than this *cyclo*pentyl derivative. Purely on the basis of ring strain, the camphene compound should have been *less* reactive. Thus the high reactivity of the camphene compound is very strong evidence for neighbouring group participation.

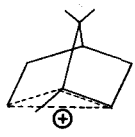
The relative rates of solvolysis of *cyclo*pentyl chloride, bornyl chloride, and *iso*bornyl chloride (in 80 per cent. ethanol at 85°) are respectively 9·4, 1·0 and 36,000 (Roberts *et al.*, 1949; Winstein *et al.*, 1952). This very large difference between the behaviour of bornyl and *iso*bornyl chlorides is readily explained by neighbouring group participation. In *iso*bornyl chloride the methylene group that forms the bridged ion is *trans* to the chloride ion ejected and so can readily



*iso*bornyl chloride

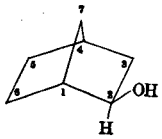


bornyl chloride

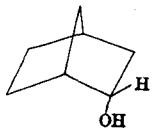

I

attack the C+ (of the C—Cl) at the rear, thereby assisting ionisation; this neighbouring group participation cannot occur with bornyl chloride. Various representations of this bridged-ion are possible; I has been proposed by Winstein *et al.* (1952).

Very strong evidence for the participation of a neighbouring saturated hydrocarbon radical has been obtained by Winstein *et al.* (1952) in their detailed examination of some reactions of the parent norbornyl systems.

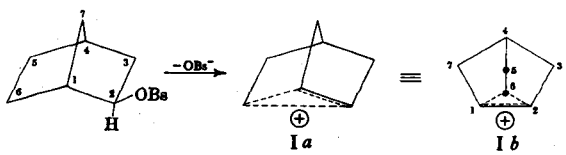


exo-norbornyl alcohol



endo-norbornyl alcohol

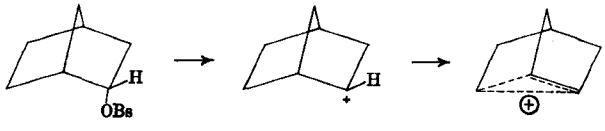
These authors showed that the relative rates of acetolysis of the brosylates (*p*-bromobenzenesulphonates) of *exo/endo* norbornyl alcohols in acetic acid at 25° are 350/1. The explanation offered for the large relative rate of the *exo*-isomer acetolysis was neighbouring group participation to form the non-classical carbonium ion (I*a*). As the OBs^- ion is leaving from the front, the neighbouring group (group C_6) can attack from the rear to form the bridged-ion. This



I*a* I*b*

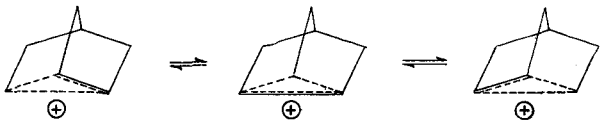
sequence is not possible as such for the *endo*-compound, and so the latter reacts far more slowly. Further support for the formation of (I*a*) is as follows. This ion has a plane of symmetry (see I*b*) and hence is optically inactive. It has been shown that solvolysis of *exo*-norbornyl brosylate in aqueous acetone, ethanol or acetic acid gives only *exo*-products, but in these products the carbon atoms have become " shuffled " (see below). Winstein *et al.* (1952) also showed that acetolysis of optically active *exo*-norbornyl brosylate gave racemic *exo*-norbornyl acetate. Attack must be from the back of the CH_2 bridge and so this results in the *exo*-product; also, since positions 1 and 2 are equivalent, equal amounts of the enantiomorphs (*i.e.*, racemate) will be produced.

When *endo*-norbornyl brosylate undergoes acetolysis, ionisation of the OBs^- group leaves the *endo*-norbornyl carbonium ion. This is probably originally the

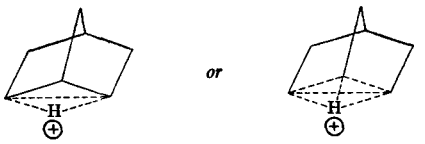


classical carbonium ion, but it then rearranges to the more stable *exo*-bridged-ion. The formation of the latter is shown by the fact that acetolysis of the optically active *endo*-brosylate produces racemic *exo*-acetate.

The structure of the bridged carbonium ion, however, appears to be more complicated than that shown by formula (I*a*). Examination of (I*b*) shows the equivalence of positions 1 and 2, and of positions 3 and 7. Thus labelling the brosylate with ^{14}C at positions 2 and 3 should give products equally labelled at positions 1, 2, 3 and 7. Roberts *et al.* (1954) carried out the acetolysis of this labelled *exo*-brosylate, and the tracer atom was found at 1, 2, 3 and 7, but positions 5 and 6 also contained labelled carbon (15 per cent. of the total radioactivity). These results can be explained on the basis that there is also a 1,3-hydride shift from position 2 to position 6. Thus positions 1, 2 and 6 become shuffled to a certain extent, and there is also the same amount of interchange among positions 3, 5 and 7. This raises the question as to whether some ions



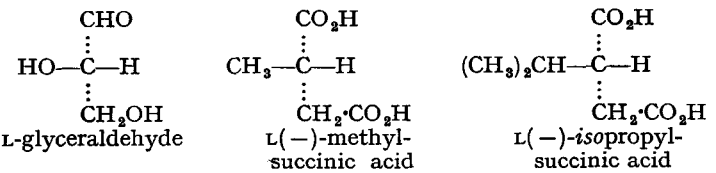
C-bridging



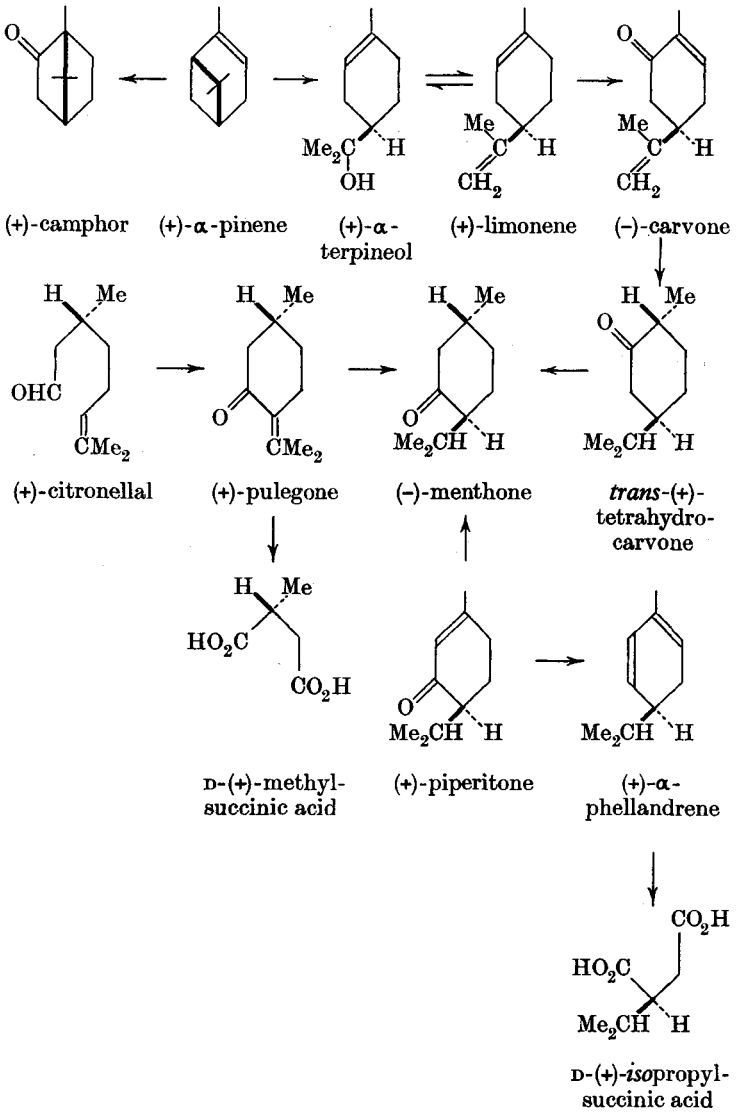
H-bridging

have both carbon and hydrogen bridging. Winstein (1955) has pointed out that the " extra " carbon shuffling (to positions 5 and 6) depends on the nucleophilic activity of the solvent, and is zero for very reactive solvents in which the life of the carbonium ion is short. This suggests that the hydrogen shift competes with the solvent attack and so occurs *after* the formation of the purely carbon bridged-ion.

§23e. Correlation of configurations of terpenes. This has been made possible by the work of Fredga on quasi-racemic compounds (see §9a. II). This author has established the following configurations:

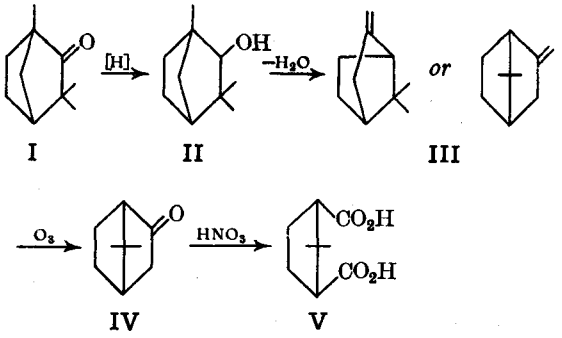


By means of these configurations, combined with various interrelations obtained by oxidative degradations and by molecular rearrangements, it has been possible to correlate the configurations of many mono- and bicyclic terpenes with L-glyceraldehyde, *e.g.*,

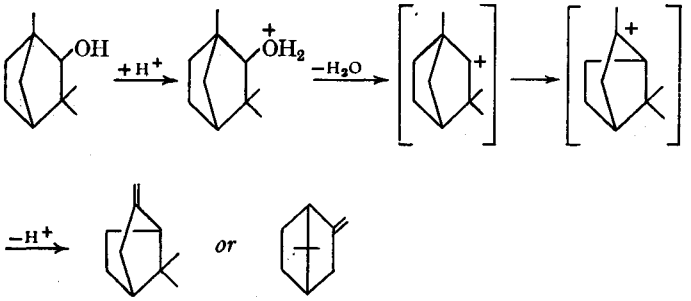


§24. Fenchane and its derivatives. The most important natural terpene of this group is **fenchone;** this occurs in oil of fennel. It is a liquid, b.p. 192–193°, and is optically active, both enantiomorphs occurring naturally.

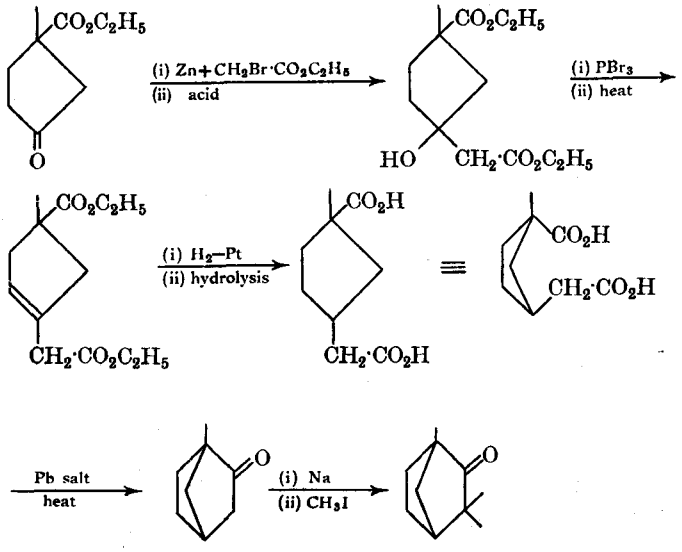
The molecular formula of fenchone is $C_{10}H_{16}O$, and the compound behaves as a ketone. When fenchone (I) is reduced with sodium and ethanol, fenchyl alcohol, $C_{10}H_{18}O$ (II), is produced, and this, on dehydration under the influence of acids, gives α-fenchene, $C_{10}H_{16}$ (III). On ozonolysis, α-fenchene is converted into α-fenchocamphorone, $C_9H_{14}O$ (IV), which, on oxidation with nitric acid, forms apocamphoric acid, V, a compound of known structure. This work was carried out by Wallach *et al.* (1890–1898), but it was Semmler (1905) who was the first to assign the correct structure to fenchone; the foregoing reactions may be formulated:



It should be noted that the dehydration of fenchyl alcohol, II, to α-fenchene, III, occurs *via* a Wagner–Meerwein rearrangement; the mechanism for this reaction may thus be written (*cf.* §23d):



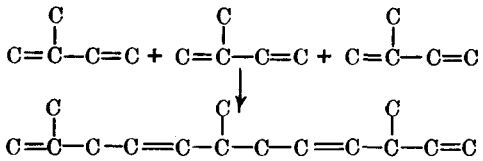
The structure of fenchone has been confirmed by synthesis (Ruzicka, 1917).



SESQUITERPENES

§25. Introduction. The sesquiterpenes, in general, form the higher boiling fraction of the essential oils; this provides their chief source. Wallach (1887) was the first to suggest that the sesquiterpene structure is built up of three isoprene units; this has been shown to be the case for the majority of the known sesquiterpenes, but there are some exceptions.

The sesquiterpenes are classified into four groups according to the number of rings present in the structure. If we use the *isoprene rule*, then when three isoprene units are linked (head to tail) to form an acyclic sesquiterpene hydrocarbon, the latter will contain *four* double bonds. Each isoprene unit contains *two* double bonds, but one disappears for each pair that is connected:



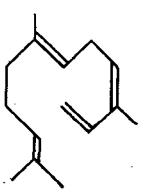
When this open-chain compound is converted into a monocyclic structure, another double bond is utilised in the process, and so monocyclic sesquiterpene hydrocarbons contain three double bonds. In a similar manner, it will be found that a bicyclic structure contains two double bonds, and a tricyclic one. Thus the nature of the sesquiterpene skeleton is also characterised by the number of double bonds present in the molecule. The sesquiterpene hydrocarbon structures may also be distinguished by the calculation of the molecular refractivities for the various types of structures, and then using these values to help elucidate the structures of new sesquiterpenes; *e.g.*, zingiberene (§27a).

Class of sesquiterpene	Number of double bonds	Molecular refractivity
Acyclic	4	69·5
Monocyclic . . .	3	67·8
Bicyclic	2	66·1
Tricyclic	1	64·4

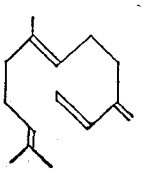
This type of information can also be used with the monoterpenes, but in this case it has not been so useful as in the sesquiterpenes. It might be noted here that the non-acyclic members of the sesquiterpenoid group may have rings of various sizes: 4, 5, 6, 7, 9, 10 and 11; and in many of these the rings are fused.

ACYCLIC SESQUITERPENES

§26. Farnesene, $C_{15}H_{24}$, b.p. 128–130°/12 mm., is obtained by the dehydration of farnesol with potassium hydrogen sulphate (Harries *et al.*,



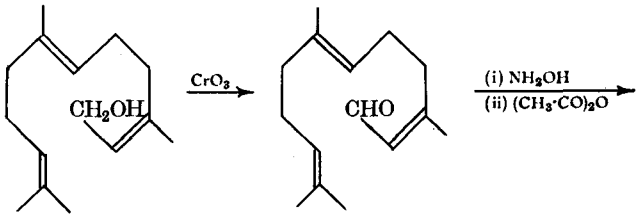
α-farnesene



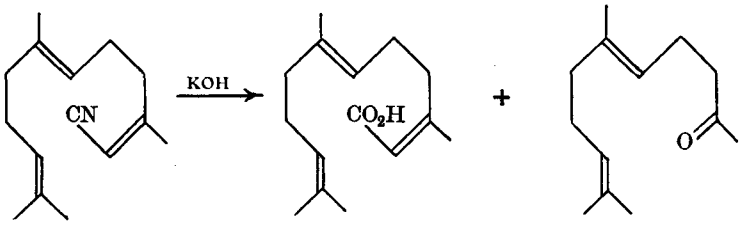
β-farnesene

1913). This compound is the α-isomer, and it has now been shown that the β-isomer occurs naturally (in oil of hops), and Sorm *et al.* (1949, 1950) have assigned it the structure shown. β-Farnesene is also obtained by the dehydration of nerolidol.

§26a. Farnesol, $C_{15}H_{26}O$, b.p. 120°/0·3 mm., occurs in the oil of ambrette seeds, etc. Its structure was elucidated by Kerschbaum (1913) as follows. When oxidised with chromic acid, farnesol is converted into farnesal, $C_{15}H_{24}O$, a compound which behaves as an aldehyde. Thus farnesol is a primary alcohol. Conversion of farnesal into its oxime, followed by dehydration with acetic anhydride, produces a cyanide which, on hydrolysis with alkali, forms farnesenic acid, $C_{15}H_{24}O_2$, and a ketone, $C_{13}H_{22}O$. This ketone was then found to be dihydro-*pseudo*-ionone (geranylacetone). In the formation of this ketone, two carbon atoms are removed from its precursor. This reaction is characteristic of α : β-unsaturated carbonyl compounds, and so it is inferred that the precursor, farnesenic acid (or its nitrile), is an α : β-unsaturated compound. Thus the foregoing facts may be formulated as follows, on the basis of the known structure of geranylacetone.

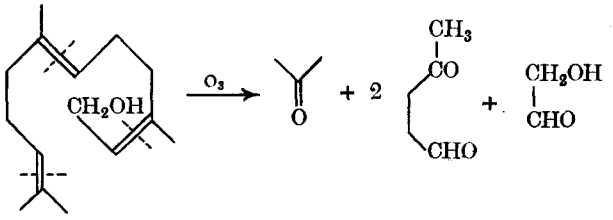


farnesol farnesal

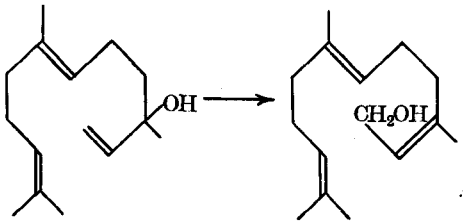


farnesenonitrile farnesenic acid geranylacetone

Kerschbaum's formula has been confirmed by Harries *et al.* (1913), who obtained acetone, lævulaldehyde and glycolaldehyde on the ozonolysis of farnesol.

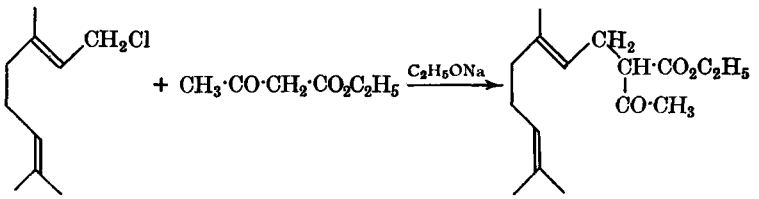


Ozonolysis, however, also gave some formaldehyde, thus indicating the presence of the *iso*propenyl end-group as well as the *iso*propylidene end-group (but *cf.* citral, §5). Ruzicka (1923) synthesised farnesol (with the *iso*propylidene end-group) by the action of acetic anhydride on synthetic nerolidol (*cf.* linalool, §8).

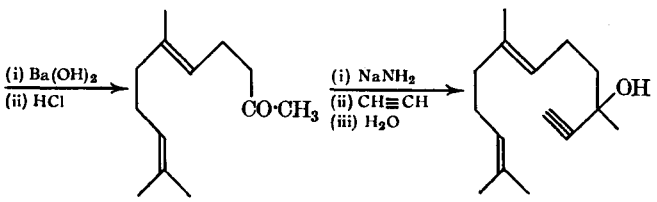


nerolidol farnesol

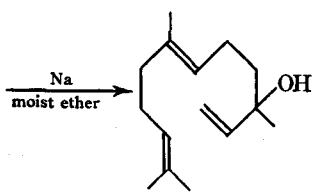
§26b. Nerolidol, $C_{15}H_{26}O$, b.p. 125–127°/4·5 mm., occurs in the oil of neroli, etc., in the (+)-form. Nerolidol is isomeric with farnesol, and Ruzicka (1923) showed that the relationship between the two is the same as that between linalool and geraniol (see §8). Ruzicka (1923) confirmed the structure of nerolidol by synthesis.



geranyl chloride



geranylacetone

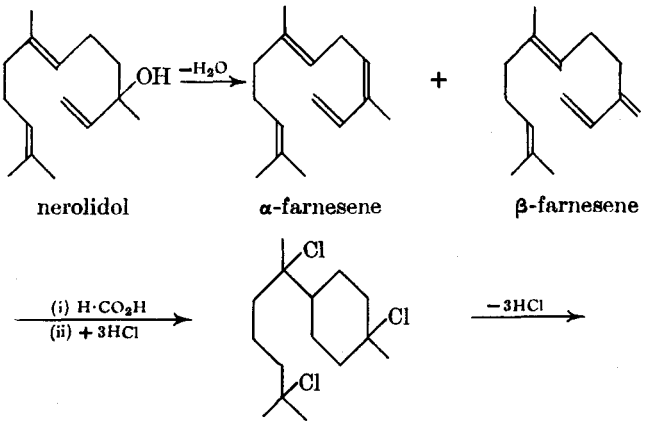


(±)-nerolidol

MONOCYCLIC SESQUITERPENES

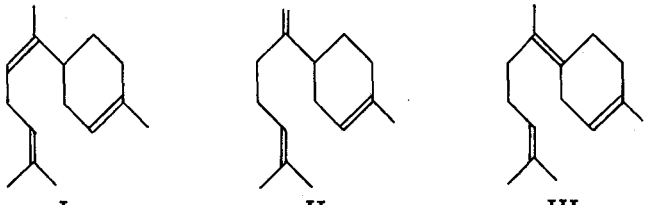
§27. Bisabolene, $C_{15}H_{24}$, b.p. 133–134°/12 mm., occurs in the oil of myrrh and in other essential oils. The structure of bisabolene was determined by Ruzicka *et al.* (1925). Bisabolene adds on three molecules of hydrogen chloride to form bisabolene trihydrochloride, and this regenerates bisabolene when heated with sodium acetate in acetic acid solution. Thus bisabolene contains three double bonds and is therefore monocyclic (see §25). Nerolidol may be dehydrated to a mixture of α- and β-farnesenes (*cf.* §26). This mixture, on treatment with formic acid, forms a monocyclic sesquiterpene (or possibly a mixture) which combines with hydrogen chloride to form bisabolene trihydrochloride. Removal of these three molecules of

hydrogen chloride (by means of sodium acetate in acetic acid) produces bisabolene; thus bisabolene could be I, II or III, since all three would give the *same* bisabolene trihydrochloride.



nerolidol α-farnesene β-farnesene

bisabolene trihydrochloride



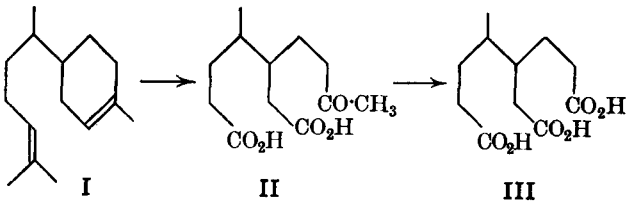
I α-bisabolene II β-bisabolene III γ-bisabolene

Ruzicka *et al.* (1929) showed that synthetic and natural bisabolene consisted mainly of the γ-isomer (III), since on ozonolysis of bisabolene, the products were acetone, lævulic acid and a small amount of succinic acid. These products are readily accounted for by III; and this structure has been confirmed by synthesis (Ruzicka *et al.*, 1932).

§27a. **Zingiberene,** $C_{15}H_{24}$, b.p. 134°/14 mm., occurs in the (−)-form in ginger oil. It forms a dihydrochloride with hydrogen chloride, and thus apparently contains two double bonds. The molecular refractivity, however, indicates the presence of three double bonds and, if this be the case, zingiberene is monocyclic (see §25). The presence of these three double bonds is conclusively shown by the fact that catalytic hydrogenation (platinum) converts zingiberene into hexahydrozingiberene, $C_{15}H_{30}$. Zingiberene can be reduced by means of sodium and ethanol to dihydrozingiberene, $C_{15}H_{26}$; this indicates that two of the double bonds are probably conjugated (Semmler *et al.*, 1913). Further evidence for this conjugation is afforded by the fact that zingiberene shows optical exaltation, whereas dihydrozingiberene does not. The absorption spectrum of zingiberene also shows the presence of conjugated double bonds (Gillam *et al.*, 1940).

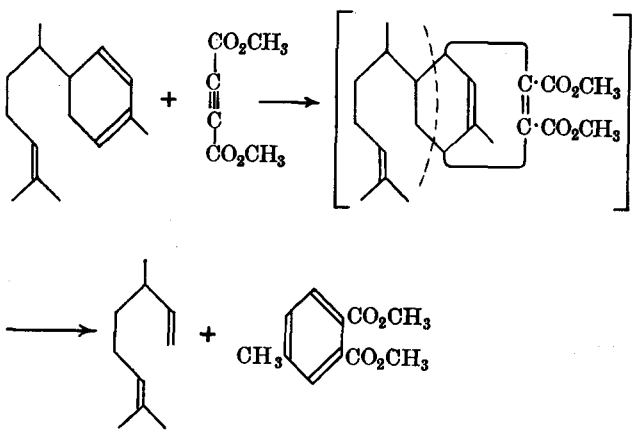
Ozonolysis of zingiberene gives acetone, lævulic acid and succinic acid (Ruzicka *et al.*, 1929). Since these products are also obtained from bisabolene (§27), it appears probable that zingiberene and bisabolene have the same carbon skeleton. Oxidation of dihydrozingiberene, I, with permanganate gives a keto-dicarboxylic acid, $C_{12}H_{20}O_5$ (II), which, on oxidation

with sodium hypobromite, forms a tricarboxylic acid, $C_{11}H_{18}O_6$ (III). Thus II must contain a methyl ketone group ($CH_3 \cdot CO—$), and so, if I be assumed as the structure of dihydrozingiberene, the foregoing oxidation reactions may be formulated:

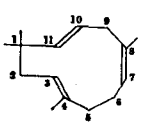


I II III

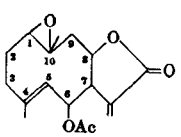
Thus I, with another double bond in conjugation with one already present, will be (probably) the structure of zingiberene. The position of this third double bond was shown as follows (Eschenmoser *et al.*, 1950). Zingiberene forms an adduct with methyl acetylenedicarboxylate, and this adduct (which was not isolated), on pyrolysis, gives 2 : 6-dimethylocta-2 : 7-diene and methyl 4-methylphthalate. These reactions can be explained on the assumption that zingiberene has the structure shown below.



§27b. Humulene (α-caryophyllene), $C_{15}H_{24}$, b.p. 264°, is an eleven-membered ring compound which contains three double bonds. Its structure is very closely related to that of caryophyllene (§28c).



humulene



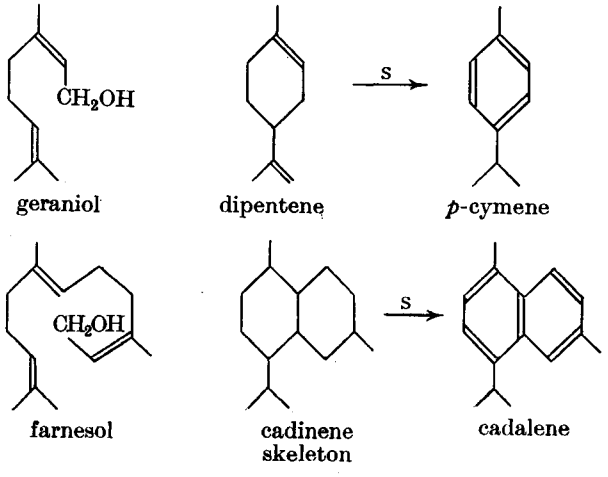
pyrethrosin

Pyrethrosin is also a monocyclic sesquiterpene; it is a γ-lactone which contains a ten-membered ring.

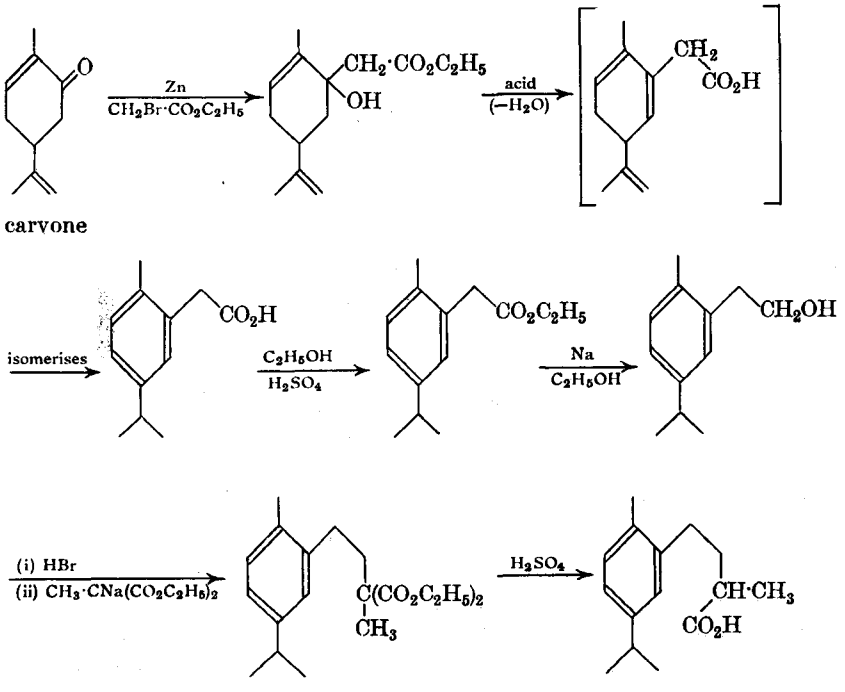
BICYCLIC SESQUITERPENES

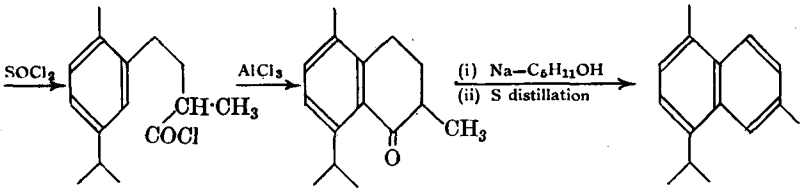
§28. Cadinene, $C_{15}H_{24}$, b.p. 134–136°/11 mm., occurs in the (—)-form in oil of cubebs, etc. Catalytic hydrogenation converts cadinene into tetrahydrocadinene, $C_{15}H_{28}$. Thus cadinene contains two double bonds and is

bicyclic. On dehydrogenation with sulphur, cadinene forms cadalene, $C_{15}H_{18}$ (Ruzicka *et al.*, 1921). Cadalene does not add on bromine, and forms a picrate. This led to the belief that cadalene was an aromatic compound, and its structure was deduced as follows. Ruzicka assumed that the relationship of farnesol (§26a) to cadinene was analogous to that of geraniol (§7) to dipentene (§13). Furthermore, since dipentene gives *p*-cymene when dehydrogenated with sulphur, then cadalene should be, if the analogy is correct, 1 : 6-dimethyl-4-*iso*propylnaphthalene; thus:

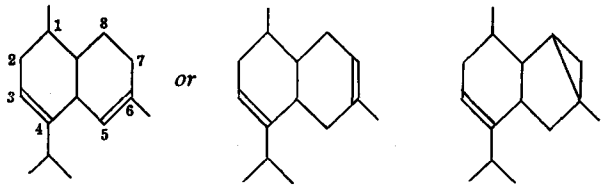


1 : 6-Dimethyl-4-*iso*propylnaphthalene was synthesised by Ruzicka *et al.* (1922), and was found to be identical with cadalene.

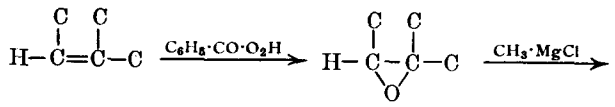


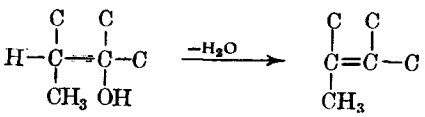


Thus cadinene has the carbon skeleton assumed. The only remaining problem is to ascertain the positions of the two double bonds in cadinene. Since the molecular refractivity shows no optical exaltation, the two double bonds are not conjugated (§11. I); this is supported by the fact that cadinene is not reduced by sodium and amyl alcohol. Ozonolysis of cadinene produces a compound containing the *same* number of carbon atoms as cadinene. The two double bonds are therefore in ring systems, but they cannot be in the *same* ring, since in this case carbon would have been lost on ozonolysis. Ruzicka *et al.* (1924) were thus led to suggest I (α *or* β) for the structure of cadinene, basing it on the relationship of cadinene to copaene, which had been given structure II by Semmler (1914). I was proposed mainly on the



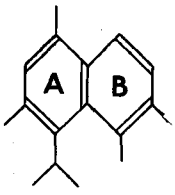
fact that copaene adds two molecules of hydrogen chloride to form copaene dihydrochloride, which is *identical* with cadinene dihydrochloride (both the α and β structures of I would give the *same* dihydrochloride as II). Structure I (α or β) was accepted for cadinene until 1942, when Campbell and Soffer re-investigated the problem. These authors converted cadinene into its monoxide and dioxide by means of perbenzoic acid, treated these oxides with excess of methylmagnesium chloride, and then dehydrogenated the product with selenium. By this means, Campbell and Soffer obtained a monomethylcadalene from cadinene monoxide, and a dimethylcadalene from cadinene dioxide. Now the introduction of a methyl group *via* the oxide takes place according to the following scheme:



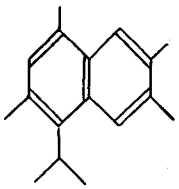


Thus the positions of the additional methyl groups show the positions of the double bonds in cadinene. The Ruzicka formula for cadinene would give dimethylcadalene III (from the α isomer) or IV (from the β), and the monomethylcadalenes would be V (from α or β), VI (from α) and VII (from β). Campbell and Soffer oxidised their dimethylcadalene, first with chromic acid and then with nitric acid, and thereby obtained pyromellitic acid

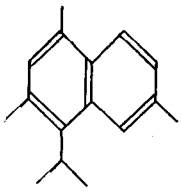
(benzene-1 : 2 : 4 : 5-tetracarboxylic acid), VIII. The formation of VIII therefore rules out III as the structure of dimethylcadalene, but IV, with



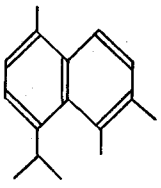
III



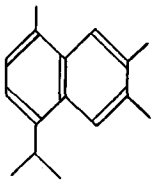
IV



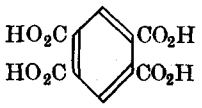
V



VI

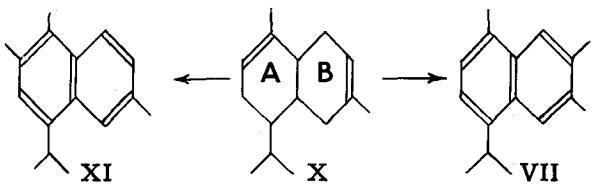


VII

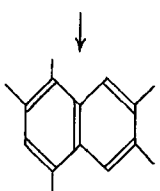


VIII

the two methyl groups at positions 6 and 7 in ring B, could give VIII. Therefore the double bond in cadinene in ring B is 6 : 7. From this it follows that VI is also eliminated. If the double bond in ring A is as in structure I, then dimethylcadalene is IV, and monomethylcadalene is V or VII. Campbell and Soffer synthesised IV and VII, and found that each was different from the methylcadalenes they had obtained from cadinene. Thus IV and VII are incorrect; consequently the double bond in ring A cannot be 3 : 4. The only other dimethylcadalene which could give VIII on oxidation is IX. This was synthesised, and was found to be identical with the dimethylcadalene from cadinene. Cadinene must therefore be X, and the introduction of one or two methyl groups may thus be formulated as follows:



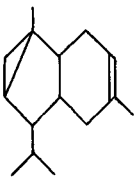
XI X VII



IX

X could give two monoxides (oxidation of ring A or B), and one of these (ring B oxidised) would give VII. This, as pointed out above, was different from the monomethylcadalene actually obtained. Therefore, if X is the structure of cadinene, the monomethylcadalene obtained from cadinene must be XI. XI was synthesised, and was found to be identical with the compound obtained from cadinene. Thus X is the structure of cadinene.

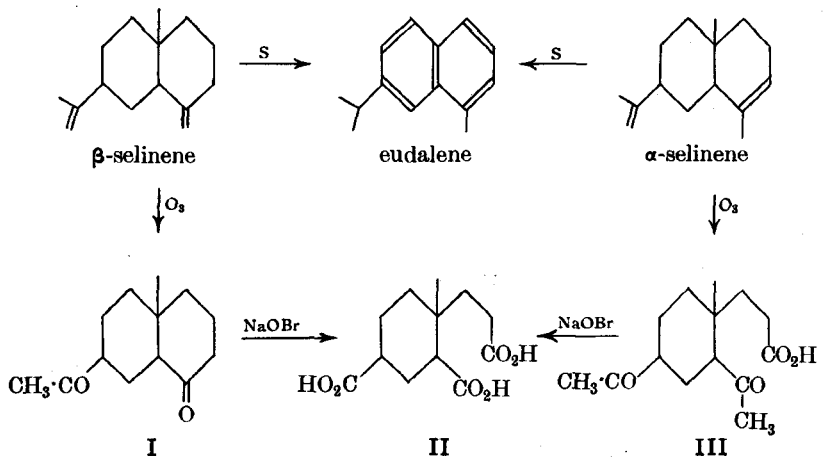
It should be noted, in passing, that this new structure for cadinene has necessitated revision of the structure of copaene. Briggs and Taylor (1947), using a technique similar to that of Campbell and Soffer, have assigned the following structure to copaene.



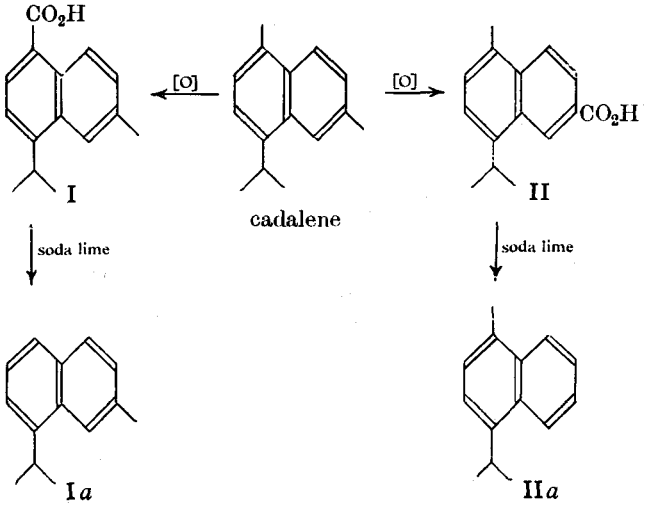
copaene

The absolute configurations of the cadinenes (and cadinols) have now been established (Motl *et al.*, 1958; Soffer *et al.*, 1958).

§28a. Selinenes, $C_{15}H_{24}$. Selinene occurs in celery oil; when treated with hydrogen chloride, it forms a dihydrochloride which, when warmed with aniline, is converted into the compound $C_{15}H_{24}$. This is isomeric with selinene, and the natural compound was called β-selinene, and the synthetic isomer α-selinene (Semmler *et al.*, 1912). Semmler showed that the catalytic hydrogenation of the two selinenes gives the same tetrahydroselinene, $C_{15}H_{28}$. Thus they each contain two double bonds, and are bicyclic. Ozonolysis of β-selinene produces a diketone (I) with the loss of two carbon atoms, and oxidation of I with sodium hypobromite gives a tricarboxylic acid (II), with the loss of one carbon atom. From this it follows that I contains a $CH_3{\cdot}CO$— group. Ozonolysis of α-selinene gives a diketo-monocarboxylic acid (III) with loss of one carbon atom, and III, on oxidation with sodium hypobromite, loses two carbon atoms to form II. Thus III contains *two* $CH_3{\cdot}CO$— groups (Semmler *et al.*, 1912). Ruzicka *et al.* (1922) distilled β-selinene with sulphur, and thereby obtained eudalene (see §28b for the evidence for the structure of this compound). If we use the isoprene rule, all the foregoing facts are explained by giving the selinenes the following structures (Ruzicka *et al.*, 1922). The relationship of the selinenes to eudesmol (§28b) confirms the nature of the carbon skeleton given to the selinenes.

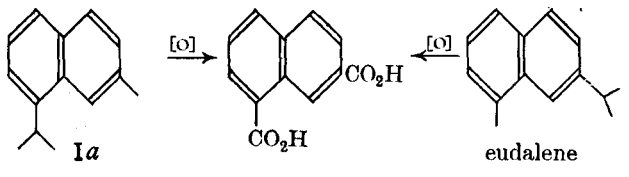


§28b. Eudesmol, $C_{15}H_{26}O$, occurs in eucalyptus oil. Catalytic hydrogenation converts eudesmol into dihydroeudesmol, $C_{15}H_{28}O$. Thus one double bond is present in the molecule, and since eudesmol behaves as a tertiary alcohol, the parent hydrocarbon is $C_{15}H_{28}$≡C_nH_{2n-2}; eudesmol is therefore bicyclic. When dehydrogenated with sulphur, eudesmol forms eudalene, $C_{14}H_{16}$, and methanethiol (Ruzicka *et al.*, 1922). Eudalene behaved as an aromatic compound (*cf.* cadalene, §28), and its structure was deduced as follows. Since eudalene was a naphthalene derivative, and since it contained one carbon atom less than cadalene, it was thought to be an apocadalene, *i.e.*, cadalene minus one methyl group. Thus eudalene is either 1-methyl-4-*iso*propylnaphthalene (II*a*) or 7-methyl-1-*iso*propylnaphthalene (I*a*). To test this hypothesis, Ruzicka oxidised cadalene with chromic acid, and thereby obtained a naphthoic acid, $C_{15}H_{16}O_2$, which must

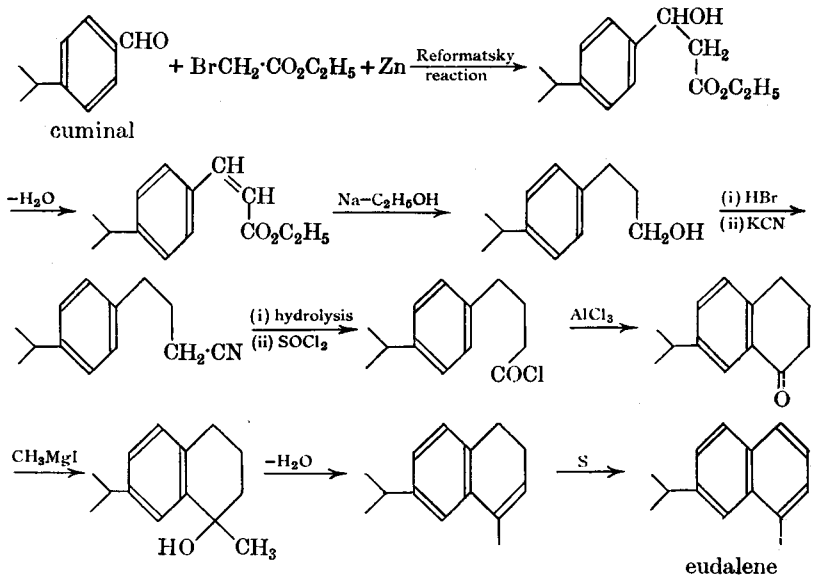


be I or II. Distillation of this acid with soda-lime gives a methyl*iso*propylnaphthalene which must be I*a* or II*a*. II*a* was synthesised from carvone (the synthesis is the same as for cadalene except that ethyl malonate is used instead of ethyl methylmalonate; see §28). The synthetic compound (II*a*) was found to be different from the hydrocarbon obtained by the distillation of the naphthoic acid from cadalene. Thus the apocadalene obtained must be I*a*, *i.e.*, 7-methyl-1-*iso*propylnaphthalene.

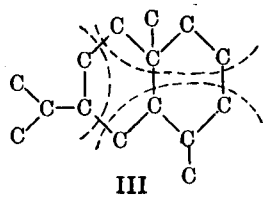
Ruzicka now found that eudalene was not identical with either I*a* or II*a*. On oxidation, however, eudalene gives the same naphthalenedicarboxylic acid as that which is obtained by the oxidation of I*a*. This is only possible if in eudalene the two side-chains in I*a* are interchanged, *i.e.*, eudalene is 1-methyl-7-*iso*propylnaphthalene; thus:



This structure for eudalene was proved by synthesis (Ruzicka *et al.*, 1922).

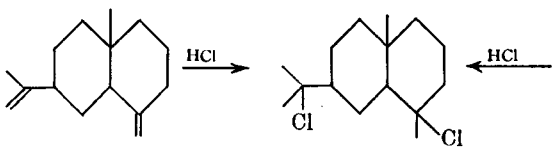


To develop the sesquiterpene carbon skeleton from that of eudalene, it is necessary to introduce one carbon atom in such a position that it is eliminated as methanethiol during the sulphur dehydrogenation (see above). If we use the *isoprene rule* with the units joined head to tail, then there is only one possible structure that fits the requirements, *viz.*, III (*cf.* §1).



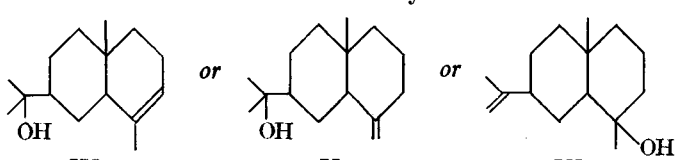
III

Now β-selinene combines with hydrogen chloride to form selinene dihydrochloride, which is also obtained by the action of hydrogen chloride on eudesmol (Ruzicka *et al.*, **1927, 1931**). Since eudesmol contains one double bond and a tertiary alcoholic group, it follows that the double bond must be in the side-chain, and the hydroxyl group in the ring, or *vice versa*, *i.e.*, IV, V or VI is the structure of eudesmol.



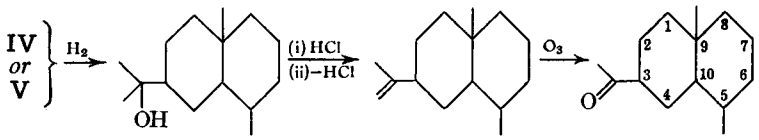
β-selinene

selinene dihydrochloride



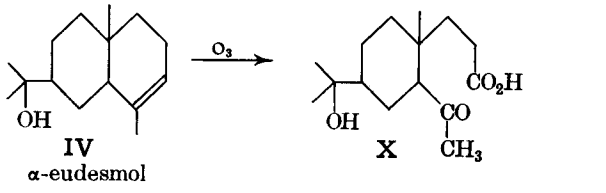
IV V VI

Hydrogenation of eudesmol forms dihydroeudesmol, VII, and this, on treatment with hydrogen chloride followed by boiling with aniline (to remove a molecule of hydrogen chloride), gives dihydroeudesmene, VIII. VIII, on ozonolysis, forms 3-acetyl-5 : 9-dimethyldecalin, IX, *with the elimination of one carbon atom.* These results are explained if IV or V is the structure of eudesmol, but not by VI. Thus the hydroxyl group is in the *iso*propyl side-chain.



VII VIII IX

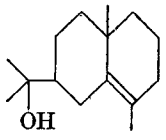
The final problem is to ascertain the position of the double bond in eudesmol, *i.e.*, Is the structure IV or V? Ozonolysis of eudesmol showed that eudesmol is a mixture of IV (α-eudesmol) and V (β-eudesmol), since *two* products are obtained: a hydroxyketo-acid X, with no loss of carbon, and a hydroxyketone XI, with the loss of one carbon atom (but *cf.* citral, §5).



IV α-eudesmol X

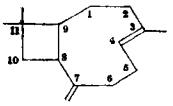
V β-eudesmol XI

The proportions of these two isomers vary with the source, and McQuillin *et al.* (1956) have succeeded in separating them (*via* their 3 : 5-dinitrobenzoates), and at the same time have characterised a third, synthetic γ-isomer.

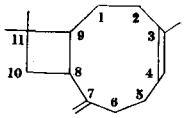


γ-eudesmol

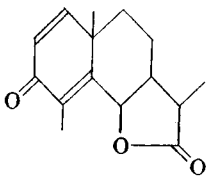
§28c. Caryophyllene, $C_{15}H_{24}$, b.p. 123–125°/10 mm., is a bicyclic sesquiterpene containing a fused system of a four- and a nine-membered ring. The main source of this compound is the sesquiterpene fraction of oil of cloves, and three isomeric hydrocarbons have been isolated. These were originally called



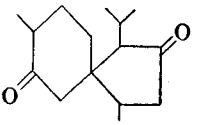
caryophyllene



*iso*caryophyllene



santonin



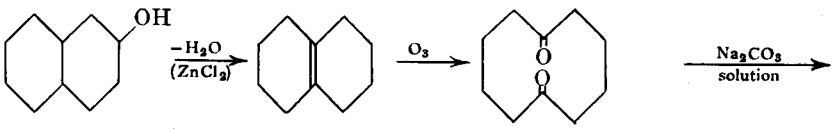
acorone

α-, β-, and γ-caryophyllene, but it has now been shown that the α-isomer is identical with humulene (§27b); the β-isomer (the main hydrocarbon) is called caryophyllene; and the γ-isomer (which is believed to be produced by thermal isomerisation) is known as iso*caryophyllene*.

Santonin is a lactone sesquiterpene of the decalin type (*cf.* pyrethrosin, §27b).

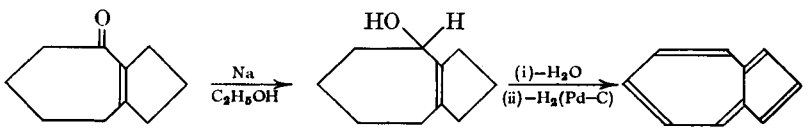
Acorone is a most interesting bicyclic sesquiterpene in that it is a carbocyclic spiran, the first example of such a compound to be found in nature.

§**29. Azulenes.** Many essential oils contain blue or violet compounds, or may form such compounds after distillation at atmospheric pressure or dehydrogenation with sulphur, selenium or palladium-charcoal (Ruzicka *et al.*, 1923). These coloured compounds may be extracted by shaking an ethereal solution of the essential oil with phosphoric acid (Sherndal, 1915). These coloured substances are known as **azulenes.** Their molecular formula is $C_{15}H_{18}$, and they are sesquiterpenes, the parent substance being azulene, $C_{10}H_8$, which contains a seven-membered ring fused to a five-membered one. Azulene has been synthesised as follows (Plattner *et al.*, 1936).



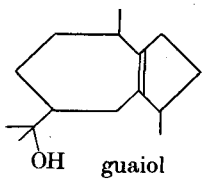
2-decalol

*cyclo*decane-1:6-dione

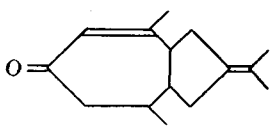


azulene

Azulene is a deep blue solid, m.p. 99°; its systematic name is bicyclo[5 : 3 : 0]-decane. Two sesquiterpenes containing this bicyclodecane skeleton are

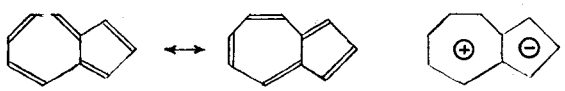


guaiol



vetivone

Azulene is a *non-benzenoid* aromatic compound in which $n = 2$ (aromatics



dipolar structure

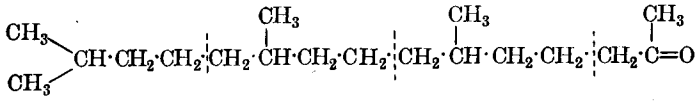
contain $(4n + 2)$ π-electrons in a " circular " system; see Vol. I, Ch. XX). It undergoes many typical aromatic substitution reactions.

DITERPENES

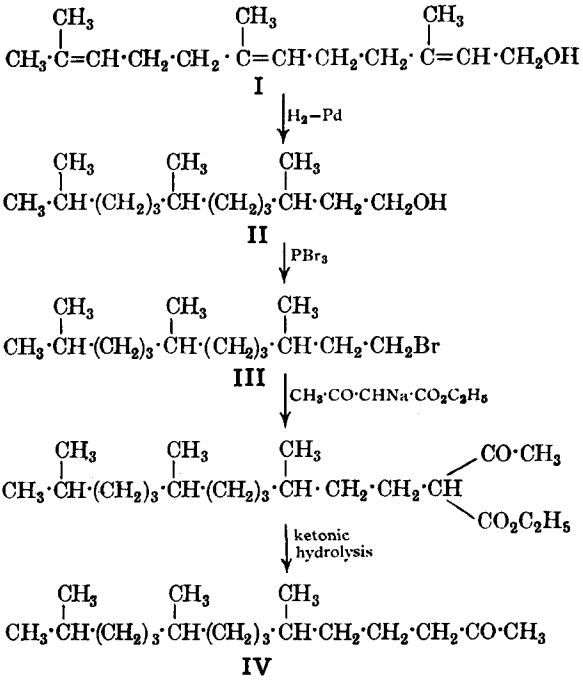
§30. Phytol, $C_{20}H_{40}O$, b.p. 145°/0·03 mm., is an acyclic diterpene; it is produced from the hydrolysis of chlorophyll (§6. XIX), and it also forms part of the molecules of vitamins E and K (see Ch. XVII). The reactions of phytol showed that it is a primary alcohol (Willstätter *et al.*, 1907), and since on catalytic reduction phytol forms dihydrophytol, $C_{20}H_{42}O$, it therefore follows that phytol contains one double bond. Thus the parent hydrocarbon is $C_{20}H_{42}$ ($\equiv C_nH_{2n+2}$), and so phytol is acyclic. Ozonolysis of phytol gives glycolaldehyde and a saturated ketone, $C_{18}H_{36}O$ (F. Fischer *et al.*, 1928). Thus this reaction may be written:

$$C_{18}H_{36}{=}CH{\cdot}CH_2OH \xrightarrow{O_3} C_{18}H_{36}O + CHO{\cdot}CH_2OH$$

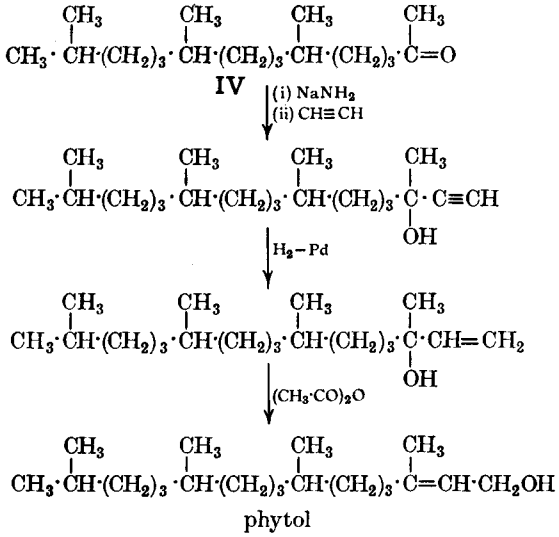
The formula of phytol led to the suggestion that it was composed of four reduced isoprene units. If this were so, and assuming that the units are joined head to tail, the structure of the saturated ketone would be:



This structure was proved to be correct by the synthesis of the ketone from farnesol (F. Fischer *et al.*, 1928). The catalytic hydrogenation of farnesol, I, produces hexahydrofarnesol, II, which, on treatment with phos-

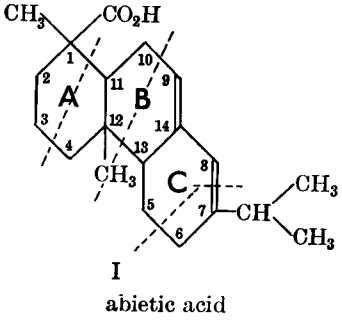


phorus tribromide, gives hexahydrofarnesyl bromide, III. III, on treatment with sodio-acetoacetic ester, followed by ketonic hydrolysis, forms the saturated ketone, IV. This ketone (IV) was then converted into phytol as follows (F. Fischer *et al.*, 1929); it should be noted that the last step involves an allylic rearrangement (*cf.* linalool, §8).

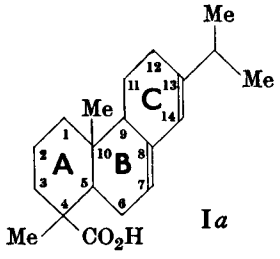


It appears that natural phytol has a very small optical rotation; Karrer *et al.* (1943) have isolated a (+)-form from nettles.

§31. Abietic acid, $C_{20}H_{30}O_2$, m.p. 170–174°, is a tricyclic diterpene. The non-steam volatile residue from turpentine is known as rosin (or colophony), and consists of a mixture of resin acids which are derived from the diterpenes. Abietic acid is one of the most useful of these acids.



I
abietic acid

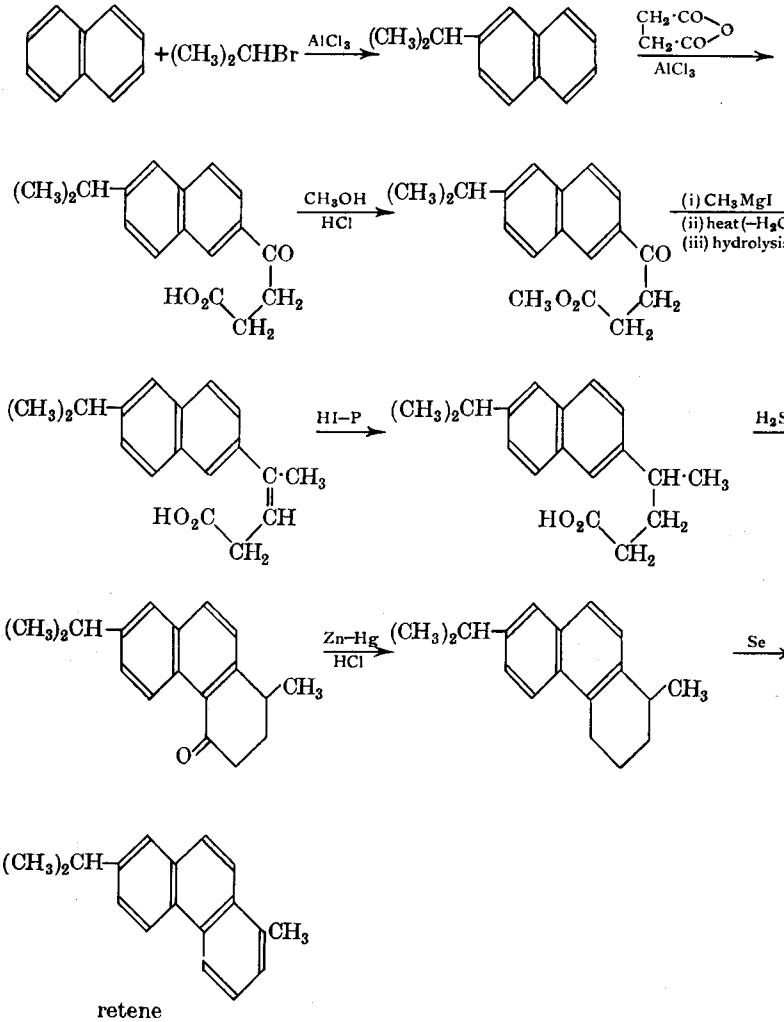


Ia

A great amount of work was done before the structure of abietic acid was elucidated. For our purpose it is useful to have the structure of abietic acid as a reference, and then describe the evidence that led to this structure. I is the structure of abietic acid; the system of numbering is shown, and also the four isoprene units comprising it. This way of numbering abietic acid follows the phenanthrene numbering. There has been recently, however, a tendency to bring the numbering of all diterpenes in line with the steroids (§3. XI); this is shown in I*a*. In the following discussion I has been used (the reader should work out the change-over for himself).

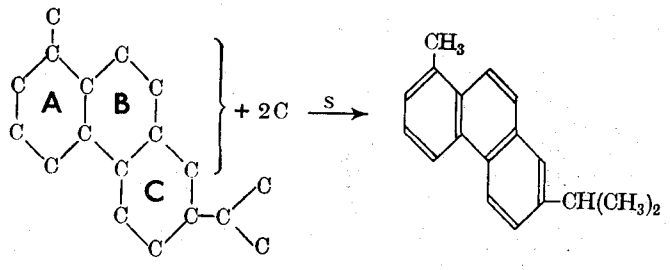
The general reactions of abietic acid showed that it was a monocarboxylic acid. On dehydrogenation with sulphur, abietic acid gives retene (Vesterberg, 1903); better yields of retene are obtained by dehydrogenating with selenium (Diels *et al.*, 1927), or with palladised charcoal (Ruzicka *et al.*, 1933). Retene, $C_{18}H_{18}$, m.p. 99°, was shown by oxidative degradation to

be 1-methyl-7-*iso*propylphenanthrene (Bucher, 1910), and this structure was later confirmed by synthesis, *e.g.*, that of Haworth *et al.* (1932).



retene

Hence we may assume that this carbon skeleton is present in abietic acid. Thus:



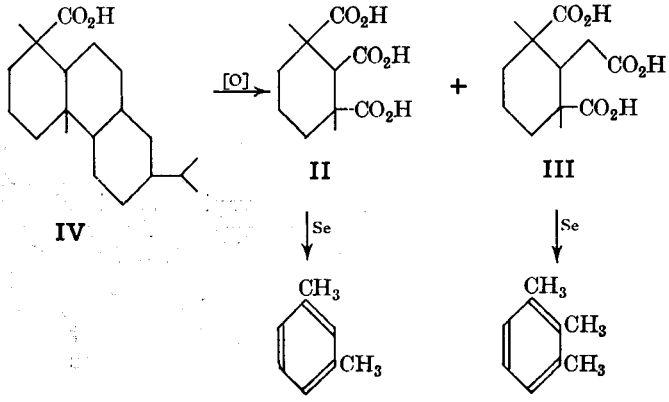
Now it is known that in sulphur dehydrogenations, carboxyl groups and

angular methyl groups can be eliminated (see §2 vii. X). It is therefore possible that the two carbon atoms lost may have been originally the carboxyl group (in abietic acid) and an angular methyl group.

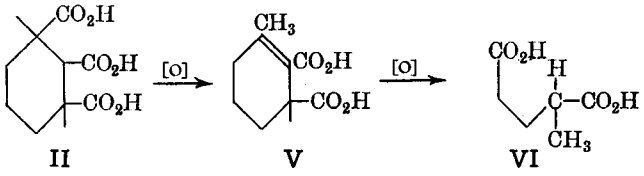
Abietic acid is very difficult to esterify, and since this is characteristic of a carboxyl group attached to a tertiary carbon atom, it suggests that abietic acid contains a carboxyl group in this state. This is supported by the fact that abietic acid evolves carbon monoxide when warmed with concentrated sulphuric acid; this reaction is also characteristic of a carboxyl group attached to a tertiary carbon atom.

Catalytic hydrogenation of abietic acid gives tetrahydroabietic acid, $C_{20}H_{34}O_2$. Thus abietic acid contains two double bonds; also, since the parent hydrocarbon is $C_{19}H_{34}$ (regarding the carboxyl group as a substituent group), abietic acid is tricyclic (parent corresponds to C_nH_{2n-4}), which agrees with the evidence already given.

Oxidation of abietic acid with potassium permanganate gives a mixture of products, among which are two tricarboxylic acids, $C_{11}H_{16}O_6$ (II), and $C_{12}H_{18}O_6$ (III) [Ruzicka *et al.*, 1925, 1931]. II, on dehydrogenation with selenium, forms *m*-xylene, and III forms hemimellitene (1 : 2 : 3-trimethylbenzene) [Ruzicka *et al.*, 1931]. In both cases there is a loss of three carbon atoms, and if we assume that these were the three carboxyl groups, then two methyl groups in II and III must be in the *meta*-position. Furthermore, since II and III each contain the methyl group originally present in abietic acid (position 1), acids II and III must contain ring A of abietic acid. This suggests, therefore, that there is an angular methyl group at position 12, since it can be expected to be eliminated from this position in sulphur dehydrogenations of abietic acid (this 12-methyl group is *meta* to the 1-methyl group). Vocke (1932) showed that acid II evolves two molecules of carbon monoxide when warmed with concentrated sulphuric acid; this indicates that II contains two carboxyl groups attached to tertiary carbon atoms. These results can be explained by *assuming* that one carboxyl group in II is that in abietic acid, and since in both cases this carboxyl group is attached to a tertiary carbon atom, the most likely position of this group is 1 (in abietic acid). Accepting these assumptions, the oxidation of abietic acid may be formulated as follows, also assuming IV as the carbon



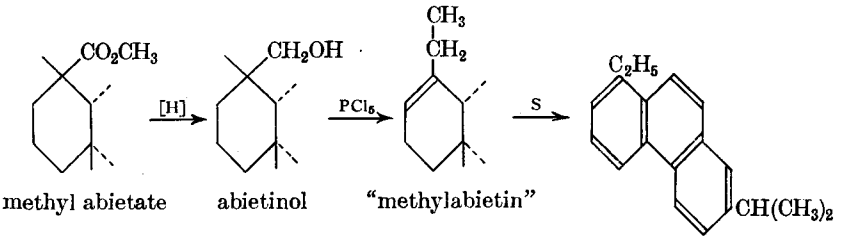
skeleton of abietic acid. Vocke subjected II to oxidative degradation, and obtained a dicarboxylic acid (V) which, on further oxidation, gave α-methylglutaric acid (VI). Vocke assumed that II had the structure shown, and formulated the reactions as below, *assuming* structure V as the best way of explaining the results.



II V VI

Structure V (assumed by Vocke) has been confirmed by synthesis (Rydon, 1937).

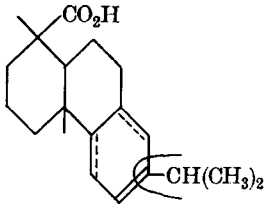
The position of the carboxyl group at position 1 in abietic acid (assumed above) has been confirmed by Ruzicka *et al.* (1922). Methyl abietate, $C_{19}H_{29}\cdot CO_2CH_3$, on reduction with sodium and ethanol, forms abietinol, $C_{19}H_{29}\cdot CH_2OH$, which, on treatment with phosphorus pentachloride, loses a molecule of water to form " methylabietin ", $C_{20}H_{30}$. This, on distillation with sulphur, forms homoretene, $C_{19}H_{20}$. Homoretene contains one CH_2 group more than retene, and on oxidation with alkaline potassium ferricyanide, gives phenanthrene-1 : 7-dicarboxylic acid, the identical product obtained from the oxidation of retene under similar conditions (Ruzicka *et al.*, 1932). These results can only be explained by assuming that homoretene has an ethyl group at position 1 (instead of the methyl group in retene), *i.e.*, homoretene is 1-ethyl-7-*iso*propylphenanthrene. This has been confirmed by synthesis (Haworth *et al.*, 1932; ethylmagnesium iodide was used instead of methylmagnesium iodide in the synthesis of retene). The formation of an ethyl group in homoretene can be explained by assuming that abietinol undergoes a Wagner–Meerwein rearrangement on dehydration (see §23d). Thus:



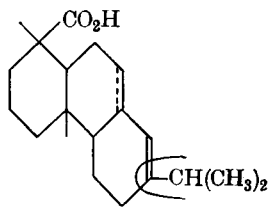
methyl abietate abietinol "methylabietin" homoretene

It has already been pointed out that abietic acid has two double bonds. Since abietic acid forms an adduct with maleic anhydride *at above* 100°, it was assumed that the two double bonds are conjugated (Ruzicka *et al.*, 1932). It was later shown, however, that levopimaric acid also forms the same adduct *at room temperature.* It thus appears that abietic acid isomerises to levopimaric acid at above 100°, and *then* forms the adduct. Thus this reaction cannot be accepted as evidence for conjugation in abietic acid. Nevertheless, the conjugation of the double bonds in abietic acid has been shown by means of the ultraviolet spectrum, which has not only shown the conjugation, but also indicates that the two double bonds are *not* in the same ring (Kraft, 1935; Sandermann, 1941).

Oxidation of abietic acid with potassium permanganate gives, among other products, *iso*butyric acid (Ruzicka *et al.*, 1925). This suggests that one double bond is in ring C and the 6 : 7- or 7 : 8-position. If the double bond is in the 6 : 7-position, then the other double bond, which is conjugated with it, must also be in the *same* ring (5 : 13 or 8 : 14); if 7 : 8, then the other double bond could be in the same ring C, but it could also



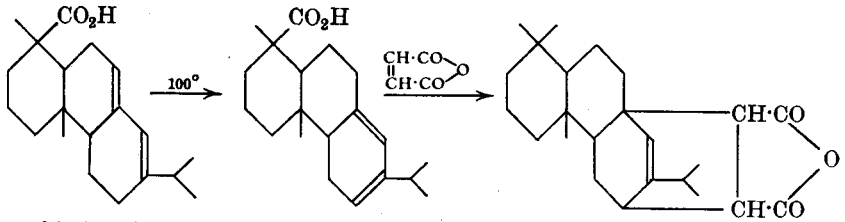
6:7-



7:8-

be in ring B. Since, as we have seen, the two double bonds are in different rings, their positions are *probably* 7 : 8 and 14 : 9. Further evidence for these positions is afforded by the fact that in the oxidation of abietic acid to give acids II and III (see above), *in which ring A is intact,* rings B and C are opened, and this can be readily explained only if rings B and C each have a double bond. Oxidative studies on abietic acid by Ruzicka *et al.* (1938–1941) have conclusively confirmed the positions 7 : 8 and 14 : 9.

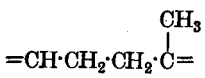
The only other point that will be mentioned here is the conversion of abietic acid into levopimaric acid. Since the latter was originally believed to be the enantiomorph of (+)-pimaric acid, it was called (−)-pimaric acid or lævopimaric acid. It is now known to be a *structural* isomer of dextropimaric acid, and so it has been suggested that levopimaric acid be called sapietic acid to avoid any confusion. The following equations show the formation of the adduct of abietic acid with maleic anhydride.



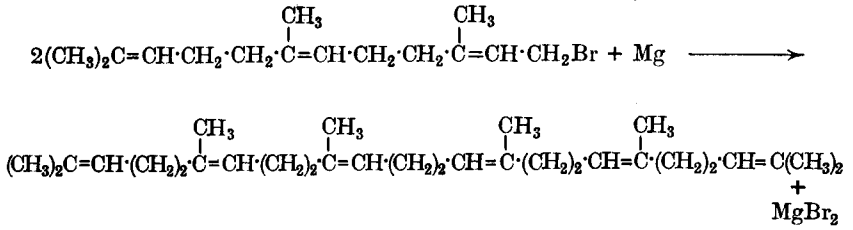
abietic acid — sapietic acid (levopimaric acid) — adduct

TRITERPENES

§32. Squalene, $C_{30}H_{50}$, b.p. 240–242°/4 mm., has been isolated from the liver oils of sharks. Other sources are olive oil and several other vegetable oils. Squalene has also been detected in leaves. Catalytic hydrogenation (nickel) converts squalene into perhydrosqualene, $C_{30}H_{62}$; therefore squalene has six double bonds, and is acyclic. Ozonolysis of squalene gives, among other products, lævulic acid; this suggests that the following group is present in squalene:



Since squalene cannot be reduced by sodium and amyl alcohol, there are no conjugated double bonds present in the molecule. Perhydrosqualene was found to be identical with the product obtained by subjecting hexahydrofarnesyl bromide to the Wurtz reaction. This led Karrer *et al.* (1931) to synthesise squalene itself from farnesyl bromide by a Wurtz reaction.



It should be noted that the centre portion of the squalene molecule has the two isoprene units joined tail to tail (*cf.* the carotenoids, Ch. IX). Squalene forms a thiourea inclusion complex, and hence it has been inferred that it is the all-*trans* stereoisomer (Schiessler *et al.*, 1952). This is supported by X-ray crystallographic studies of the thiourea inclusion complex (Nicolaides *et al.*, 1954).

§32a. Biosynthesis of terpenes. As more and more natural products were synthesised in the laboratory, so grew the interest in how these compounds are synthesised in the living organism (both animal and plant). The general approach to biosynthesis has been to break up the structure into units from which the compound could plausibly be derived. These units must, however, be known, or can be expected, to be available in the organism. Furthermore, this does not mean that the units chosen must necessarily be involved in the building-up of the compound. The general principle is that although a particular unit may itself be involved, it is also possible that its " equivalent " may act as a substitute, *i.e.*, any compound that can readily give rise to this unit (by means of various reactions such as reduction, oxidation, etc.) may be the actual compound involved in the biosynthesis. *E.g.*, the equivalent of formaldehyde could be formic acid, and that of acetone acetoacetic acid. One other point about the choice of units or their equivalents is to attempt to find some relationships between the various groups of natural products so that the units chosen are *common* precursors.

When the units have been chosen, the next problem is to consider the types of reactions whereby the natural products are synthesised in the organism. The general principle is to use reactions which have been developed in the laboratory. The difficulty here is that some types of laboratory reactions require conditions that cannot operate in the organism, *e.g.*, carboxylation and decarboxylation are known biological processes, but when carried out in the laboratory, these reactions normally require elevated temperatures. Deamination is also a known biological process, but in the laboratory this reaction is usually carried out under conditions of (*p*H) which would be lethal to the living organism. These differences between laboratory syntheses and biosyntheses are due to the action of enzymes in the latter. According to Schöpf (1932), syntheses in plants may take place through the agency of specific or non-specific enzymes (see §§12–17. XIII), or without enzymes at all. Chemical syntheses (these do not involve the use of enzymes) must therefore, from the point of biosynthetic studies, be carried out under conditions of *p*H and temperatures comparable with those operating in plants. Chemical syntheses performed in this way (with the suitable units) are said to be carried out under *physiological conditions* (which involve a *p*H of about 7 in aqueous media and ordinary temperatures).

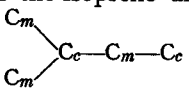
Reactions which are commonly postulated in biosynthesis are oxidation, hydrogenation, dehydrogenation, dehydration, esterification, hydrolysis, carboxylation, decarboxylation, amination, deamination, isomerisation, condensation and polymerisation. It might be noted here that the choice of

units and type of reaction are usually dependent on each other. Furthermore, other reactions which are known to occur in biological syntheses are *O*- and *N*-methylation or acylation. These may be described as *extra-skeletal processes*, and can occur at any suitable stage in the postulated biosynthesis. Another extra-skeletal process is *C*-methylation, but this is much rarer than those mentioned above.

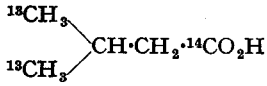
Now let us apply these principles to the biosynthesis of terpenes. As we have seen, according to the special isoprene rule, terpenes are built up of isoprene units joined head to tail (§1). Assuming then that the isoprene unit is the basic unit, the problem is: How is it formed, and how do these units join to form the various types of terpenes? At present it is believed that the fundamental units used in the cell in syntheses are water, carbon dioxide, formic acid (as " active formate "), and acetic acid (as " active acetate "). These " active " compounds are acyl derivatives of coenzyme A (written as CoA—H in the following equation); *e.g.*, acetoacetic acid is believed to be formed as follows:

$$2CH_3{\cdot}CO{\cdot}CoA + H_2O \longrightarrow CH_3{\cdot}CO{\cdot}CH_2{\cdot}CO_2H + 2CoA{-}H$$

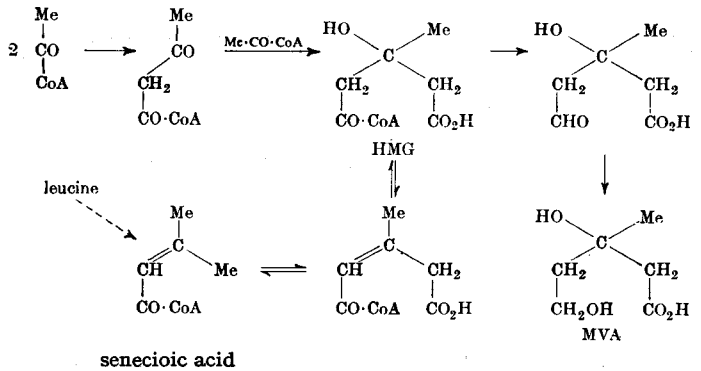
Now the biosynthesis of cholesterol (§7a. XI) from acetic acid labelled with ^{14}C in the methyl group (C_m) and in the carboxyl group (C_c) has led to the suggestion that the carbon atoms in the isoprene unit are distributed as follows:



This distribution is in agreement with a scheme in which senecioic acid (3-methylbut-2-enoic acid) is formed first, and this pathway was supported by the isolation of this acid from natural sources. Further support for the formation of this carbon skeleton is given by the fact that labelled *iso*valeric acid gives rise to cholesterol in which the *iso*propyl group and the carboxyl group have been incorporated.



Tavormina *et al.* (1956), however, have shown that the lactone of mevalonic acid (β-hydroxy-β-methyl-δ-valerolactone) is converted almost completely into cholesterol by rat liver, and is a much better precursor than senecioic acid. The following scheme has therefore been proposed for the early stages in the biosynthesis of terpenes; it is in agreement with the distribution of the carbon atoms in cholesterol (see above):



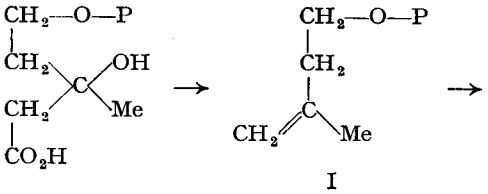
senecioic acid

Three molecules of active acetate form hydroxymethylglutaric acid, HMG (Lynen *et al.*, 1958; Rudney, 1959), and this is then converted into mevalonic

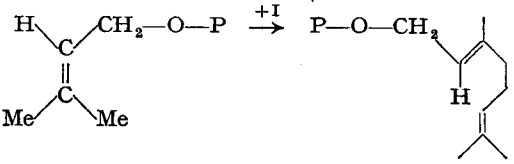
acid (MVA), possibly through the intermediate mevaldic acid (Rudney *et al.*, 1958; Lynen, 1959). Support for this sequence is afforded by the following facts. MVA has been isolated from natural sources (Wolf *et al.*, 1957), and it is also known that HMG may be formed from leucine by the route shown (Lynen *et al.*, 1958, 1959).

The biosynthesis of terpenes can be subdivided into three definite steps: (i) the formation of a biological *iso*pentane unit from acetate; (ii) the condensation of this unit to form acyclic terpenes; (iii) the conversion of acyclic into cyclic terpenes.

The stages leading to MVA have been discussed above. What happens after this is uncertain. One suggestion is that MVA forms a pyrophosphate (at the primary alcoholic group), and then the carboxyl and the tertiary hydroxyl group are eliminated simultaneously to form *iso*pentenyl pyrophosphate (I). This isomerises to the *iso*propylidene compound, $\beta:\beta$-dimethylallyl pyrophosphate, which combines with (I) to form the pyrophosphate of the acyclic terpene geraniol (in the following equations P represents the pyrophosphate residue, $P_2O_6H_3$):

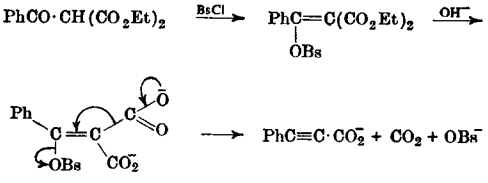


I

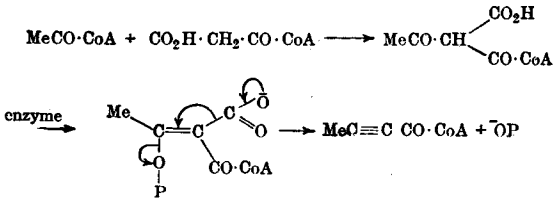


This is supported by the following work: Stanley (1958) has shown that labelled MVA (2-^{14}C-MVA) is incorporated into α-pinene. Park *et al.* (1958) have observed the incorporation of labelled MVA into rubber (§33) by an enzyme system from latex, and Lynen *et al.* (1961) have also demonstrated the conversion of *iso*pentenyl pyrophosphate into rubber (see also §7a. XI). Geranyl pyrophosphate has also been shown to be a precursor for farnesyl pyrophosphate, which then gives squalene.

A point of interest here is that Harley-Mason *et al.* (1961) have prepared phenylpropiolic acid by the action of brosyl chloride on the sodium derivative of diethyl benzoylmalonate and treating the product with sodium hydroxide in aqueous dioxan at room temperature. The reaction has been formulated as follows:



This provides one of the mildest known methods for making an acetylenic bond, and this reaction may be regarded as support for the mechanism proposed by Jones (1961) as a possible route for the biosynthesis of acetylenic bonds:



POLYTERPENES

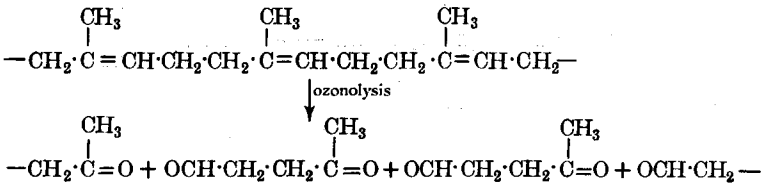
§33. Rubber. *Rubber* (*caoutchouc*) is obtained from latex, which is an emulsion of rubber particles in water that is obtained from the inner bark of many types of trees which grow in the tropics and sub-tropics. When the bark of the rubber tree is cut, latex slowly exudes from the cut. Addition of acetic acid coagulates the rubber, which is then separated from the liquor and either pressed into blocks or rolled into sheets, and finally dried in a current of warm air, or smoked.

Crude latex rubber contains, in addition to the actual rubber hydrocarbons (90–95 per cent.), proteins, sugars, fatty acids and resins, the amounts of these substances depending on the source. Crude rubber is soft and sticky, becoming more so as the temperature rises. It has a low tensile strength and its elasticity is exhibited only over a narrow range of temperature. When treated with solvents such as benzene, ether, light petrol, a large part of the crude rubber dissolves; the rest swells but does not dissolve. This insoluble fraction apparently contains almost all of the protein impurity. On the other hand, rubber is insoluble in acetone, methanol, etc. When unstretched, rubber is amorphous; stretching or prolonged cooling causes rubber to crystallise.

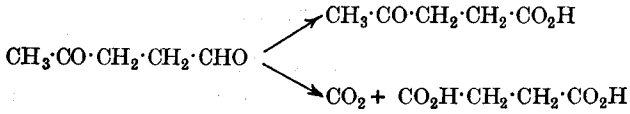
Structure of rubber. The destructive distillation of rubber gives isoprene as one of the main products; this led to the suggestion that rubber is a polymer of isoprene, and therefore to the molecular formula $(C_5H_8)_n$. This molecular formula has been confirmed by the analysis of pure rubber. Crude rubber may be purified by fractional precipitation from benzene solution by the addition of acetone. This fractional precipitation, however, produces molecules of different sizes, as shown by the determination of the molecular weights of the various fractions by osmotic pressure, viscosity and ultracentrifuge measurements; molecular weights of the order of 300,000 have been obtained.

The halogens and the halogen acids readily add on to rubber, *e.g.*, bromine gives an addition product of formula $(C_5H_8Br_2)_n$, and hydrogen chloride the addition product $(C_5H_9Cl)_n$. Pure rubber has been hydrogenated to the fully saturated hydrocarbon $(C_5H_{10})_n$—this is known as *hydrorubber*—by heating with hydrogen in the presence of platinum as catalyst (Pummerer *et al.*, 1922). Rubber also forms an ozonide of formula $(C_5H_8O_3)_n$. All these addition reactions clearly indicate that rubber is an unsaturated compound, and the formulæ of the addition products show that there is one double bond for each isoprene unit present.

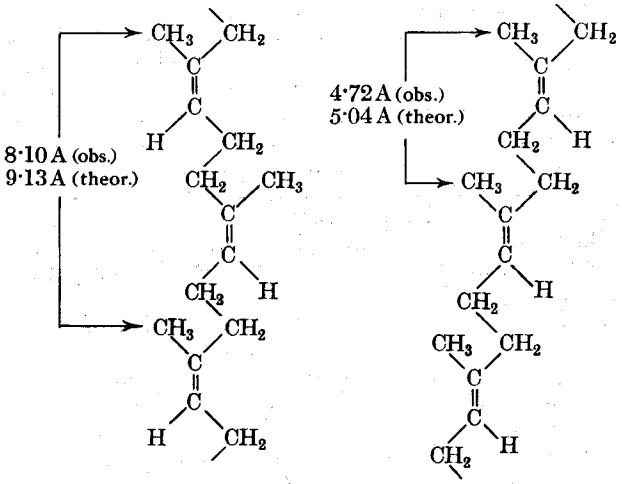
Ozonolysis of rubber produces lævulaldehyde and its peroxide, lævulic acid and small amounts of carbon dioxide, formic acid and succinic acid (Harries, 1905–1912). Pummerer (1931) showed that the lævulic derivatives comprised about 90 per cent. of the products formed by the ozonolysis. This observation led to the suggestion that rubber is composed of isoprene units joined head to tail. Thus, if rubber has the following structure, the formation of the products of ozonolysis can be explained:



Some of the lævulaldehyde is further oxidised to lævulic and succinic acids.



Gutta-percha (which is also obtained from the bark of various trees) is isomeric with rubber; their structures are the same, as shown by the methods of analysis that were used for rubber. X-ray diffraction studies (Bunn,



rubber
cis-form

gutta-percha
trans-form

1942) have shown that rubber is composed of long chains built up of isoprene units arranged in the *cis*-form, whereas gutta-percha is the *trans*-form. Gutta-percha is hard and has a very low elasticity.

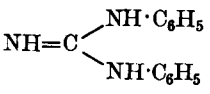
In rubber, the chain repeat unit is 8·10 A, whereas in gutta-percha it is 4·72 A. Both of these values are shorter than the theoretical values of the repeat distances (9·13 A and 5·04 A respectively) calculated from models. The reasons for these discrepancies are not clear, but for gutta-percha it has been explained by assuming that the isoprene units are not coplanar. The infra-red absorption spectrum of rubber has bands which are in keeping with the structure that has been proposed. Also, the *linear* shape of the molecule is indicated by viscosity measurements of rubber solutions. Schulz *et al.* have examined *cyclo*hexane solutions of rubber by light-scattering methods, and obtained a value of 1,300,000 for the molecular weight. Their other work also supports the linear nature of the chain.

§33a. Vulcanisation of rubber. When crude rubber is heated with a few per cent. of sulphur, the rubber becomes *vulcanised*. Vulcanised rubber

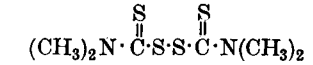
is less sticky than crude rubber, and is not so soluble and does not swell so much in organic solvents. Furthermore, vulcanised rubber has greater tensile strength and elasticity than crude rubber.

The mechanism of vulcanisation is still not clear. Vulcanised rubber is not so unsaturated as rubber itself, the loss of one double bond corresponding approximately to each sulphur atom introduced. It therefore appears that *some* sulphur atoms enter the chain, vulcanisation thus occurring through intramolecular and intermolecular cross-links; it is the latter type of reaction that is desirable in vulcanisation. It should be noted that not all the sulphur is in a combined state; some is *free*, and this can be readily extracted.

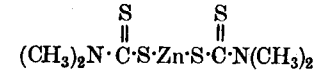
Vulcanisation may be accelerated and carried out at lower temperatures in the presence of certain organic compounds. These compounds are consequently known as *accelerators*, and all of them contain nitrogen or sulphur, or both, *e.g.*,



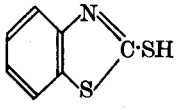
diphenylguanidine



tetramethylthiuram disulphide



zinc dimethyldithiocarbamate



mercaptobenzothiazole

Mercaptobenzothiazole is the most widely used accelerator. Many inorganic compounds can also act as accelerators, *e.g.*, zinc oxide. Organic accelerators are promoted by these inorganic compounds, and current practice is to vulcanise rubber with, *e.g.*, mercaptobenzothiazole in the presence of zinc oxide.

The actual properties of vulcanised rubber depend on the amount of sulphur used, the best physical properties apparently being achieved by using about 3 per cent. sulphur, 5 per cent. zinc oxide and about 1 per cent. of the accelerator. When 30–50 per cent. sulphur is used, the product is *ebonite*.

The elasticity of rubber is believed to be due to the existence of rubber as long-chain molecules which are highly " kinked " in the normal state. When subjected to a stretching force, these chains " unkink ", and return to their normal condition when the force is removed.

§33b. Synthetic rubbers. There are many synthetic rubbers in use, each type possessing certain desirable properties. A great deal of work has been done on the synthesis of *natural* rubber, but the difficulty has been to obtain the isoprene units in the all-*cis* configuration. Wilson *et al.* (1956) have achieved this by using stereospecific catalysts.

Buna rubbers. Under the influence of sodium, butadiene polymerises to a substance which has been used as a rubber substitute under the name of *Buna* (see Vol. I). *Buna N* is a synthetic rubber which is produced by the copolymerisation of butadiene and vinyl cyanide. *Buna S* or *Perbunan* is a copolymer of butadiene and styrene.

Butyl rubber. Copolymerisation of *iso*butylene with a small amount of isoprene produces a poly*iso*butylene known as *Butyl rubber*.

Neoprene. When passed into a solution of cuprous chloride in ammonium chloride, acetylene dimerises to vinylacetylene. This dimer can

add on one molecule of hydrogen chloride to form *Chloroprene* (2-chlorobuta-1 : 3-diene), the addition taking place in accordance with Markownikoff's rule (see also Vol. I).

$$2CH{\equiv}CH \longrightarrow CH_2{=}CH{-}C{\equiv}CH \xrightarrow{HCl} CH_2{=}CH{-}CCl{=}CH_2$$

Chloroprene readily polymerises to a rubber-like substance known as *Neoprene.* Actually, the nature of the polychloroprene depends on the conditions of the polymerisation.

Silicone rubbers. These are chemically similar to the silicone resins. The chief silicone rubber is prepared by treating the hydrolysis product of dimethyldichlorosilane, $(CH_3)_2SiCl_2$, with various compounds capable of increasing the molecular weight without the formation of cross-links, *i.e.*, they produce long-chain molecules.

$$-Si(CH_3)_2-O-Si(CH_3)_2-O-Si(CH_3)_2-O-$$

Silicone rubbers have very high electrical insulating properties, and do not deteriorate on exposure to light and air, and are resistant to the action of acids and alkalis.

READING REFERENCES

The Terpenes, Cambridge University Press (2nd ed.). Sir John Simonsen and Owen. Vol. I (1947); Vol. II (1949). Sir John Simonsen and Barton. Vol. III (1952). Sir John Simonsen and Ross. Vol. IV. (1957); Vol. V. (1957).

Gilman (Ed.), *Advanced Organic Chemistry,* Wiley (1953). Vol. IV, Ch. 7. The Terpenes.

Rodd (Ed.), *Chemistry of the Carbon Compounds,* Elsevier. (i) Vol. IIA (1953). Ch. 11. Rubber and Rubber-like Compounds (p. 407). (ii) Vol. IIB (1953). Chh. 12–16. Terpenoids.

Mayo, Vol. I. *Mono-* and *Sesquiterpenoids.* Vol. II. *The Higher Terpenoids.* Interscience (1959).

Pinder, *The Chemistry of the Terpenes,* Chapman and Hall (1960).

Ruzicka, History of the Isoprene Rule, *Proc. Chem. Soc.,* 1959, 341.

Ginsburg (Ed.), *Non-Benzenoid Aromatic Compounds,* Interscience (1959). Chh. V, VI. Azulenes.

Streitwieser, Solvolytic Displacement Reactions at Saturated Carbon Atoms, *Chem. Reviews,* 1956, **56,** p. 698 (Wagner–Meerwein Rearrangements).

Barton, The Chemistry of the Diterpenoids, *Quart. Reviews* (*Chem. Soc.*), 1949, **3,** 36.

Gascoigne and Simes, The Tetracyclic Terpenes, *Quart. Reviews* (*Chem. Soc.*), 1955, **9,** 328.

Barton and Mayo, Recent Advances in Sesquiterpenoid Chemistry, *Quart. Reviews* (*Chem. Soc.*), 1957, **11,** 189.

Halsall and Theobald, Recent Aspects of Sesquiterpenoid Chemistry, *Quart. Reviews* (*Chem. Soc.*), 1962, **16,** 101.

Progress in Organic Chemistry, Butterworths. Vol. 5 (1961). Ch. 4. The Chemistry of the Higher Terpenoids.

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

Sir Robert Robinson, *The Structural Relations of Natural Products,* Oxford Press (1955).

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

Birch, Some Pathways in Biosynthesis, *Proc. Chem. Soc.,* 1962, 3.

Gee, Some Thermodynamic Properties of High Polymers and their Molecular Interpretation, *Quart. Reviews* (*Chem. Soc.*), 1947, **1,** 265.

Hardy and Megson, The Chemistry of Silicon Polymers, *Quart. Reviews* (*Chem. Soc.*), 1948, **2,** 25.

Flory, *Principles of Polymer Chemistry,* Cornell University Press (1953).

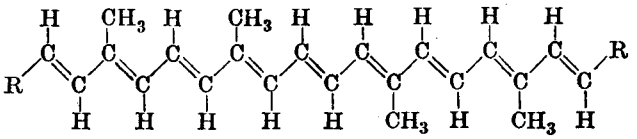
CHAPTER IX

CAROTENOIDS

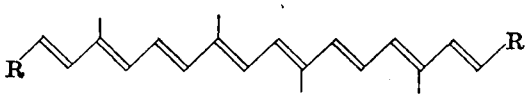
§1. Introduction. The carotenoids are yellow or orange pigments which are widely distributed in plants and animals. Chlorophyll is always associated with the carotenoids carotene and lutein; the carotenoids act as photosensitisers in conjunction with chlorophyll. When chlorophyll is absent, *e.g.*, in fungi, then the carotenoids are mainly responsible for colour. Carotenoids are also known as lipochromes or chromolipids because they are fat-soluble pigments. They give a deep blue colour with concentrated sulphuric acid and with a chloroform solution of antimony trichloride (the Carr–Price reaction); this Carr–Price reaction is the basis of one method of the quantitative estimation of carotenoids. Some carotenoids are hydrocarbons; these are known as the *carotenes*. Other carotenoids are oxygenated derivatives of the carotenes; these are the *xanthophylls*. There are also acids, the *carotenoid acids*, and esters, the *xanthophyll esters*.

Chemically, the carotenoids are polyenes, and almost all the carotenoid hydrocarbons have the molecular formula $C_{40}H_{56}$. Also, since the carbon skeleton of these compounds has a polyisoprene structure, they may be regarded as tetraterpenes (*cf.* §1. VIII).

In most of the carotenoids, the central portion of the molecule is composed of a long conjugated chain comprised of four isoprene units, the centre two of which are joined tail to tail. The ends of the chain may be two open-chain structures, or one open-chain structure and one ring, or two rings. The colour of the carotenoids is attributed to the extended conjugation of the central chain (see Vol. I). X-ray analysis has shown that in the majority of natural carotenoids, the double bonds are in the *trans*-position; a few natural carotenoids are *cis*-. Thus, if we represent the ends of the chain by R (where R may be an open-chain structure or a ring system), *trans*-carotenes may be written:



If we use the conventional formulæ of terpenes (§4. VIII), the above formula will be the following (the reader should write out in this way the various formulæ given in the text; see §6 for an example):



§2. Carotenes. Carotene was first isolated by Wackenroder (1831) from carrots (this was the origin of the name *carotin*, which was later changed to *carotene*). The molecular formula of carotene, however, was not determined until 1907, when Willstätter showed it was $C_{40}H_{56}$. Carotene was shown to be unsaturated, and when treated with a *small* amount of iodine, it forms a crystalline di-iodide, $C_{40}H_{56}I_2$. Kuhn (1929) separated this di-iodide into two fractions by means of fractional crystallisation. Treatment of each fraction with thiosulphate regenerated the corresponding carotenes, which were designated α- and β-carotene. Kuhn *et al.* (1933) then found that

chromatography gives a much better separation of the carotenes themselves, and in this way isolated a third isomer, which he designated γ-carotene.

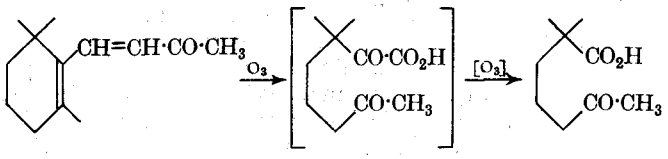
α-Carotene, m.p. 187–187·5°; optically active (dextrorotatory).
β-Carotene, m.p. 184·5°; optically inactive.
γ-Carotene, m.p. 176·5°; optically inactive.

It appears that all three carotenes occur together in nature, but their relative proportions vary with the source, *e.g.*, carrots contain 15 per cent. α, 85 per cent. β and 0·1 per cent. γ. Carotenes are obtained commercially by chromatography, two of the best sources being carrots and alfalfa.

Biosynthetic studies of the carotenes have been carried out, and the pathways are those for the terpenes (§32a. VIII). Thus Braithwaite *et al.* (1957) and Grob (1957) have shown that labelled mevalonic acid is incorporated into β-carotene. Scheuer *et al.* (1959) have also shown that this acid is incorporated into lycopene. Furthermore, Modi *et al.* (1961) have isolated mevalonic acid from carrots.

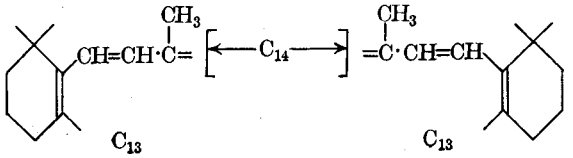
§3. β-Carotene, $C_{40}H_{56}$. When catalytically hydrogenated (platinum), β-carotene forms perhydro-β-carotene, $C_{40}H_{78}$. Thus β-carotene contains eleven double bonds, and since the formula of perhydro-β-carotene corresponds to the general formula C_nH_{2n-2}, it follows that the compound contains two rings.

When exposed to air, β-carotene develops the odour of violets. Since this odour is characteristic of β-ionone, it was thought that this residue is present in β-carotene (see §6. VIII). This was confirmed by the fact that the oxidation of a benzene solution of β-carotene with cold aqueous potassium permanganate gives β-ionone. Now β-ionone, I, on ozonolysis, gives, among other things, geronic acid, II (Karrer *et al.*, 1929).



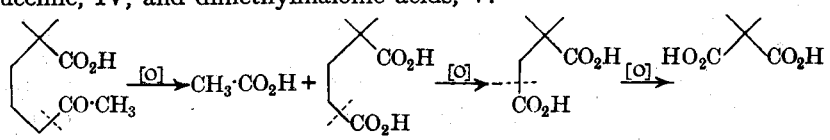
I II

β-Carotene, on ozonolysis, gives geronic acid in an amount that corresponds to the presence of two β-ionone residues (Karrer *et al.*, 1930). Thus a tentative structure for β-carotene is:



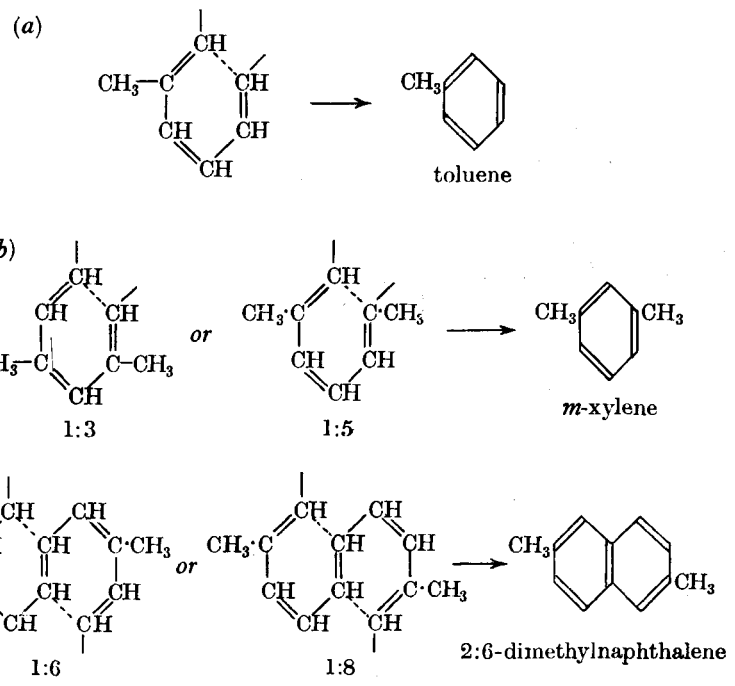
Since the colour of β-carotene is due to extended conjugation (§1), the C_{14} portion of the molecule will be conjugated. The presence of conjugation in this central portion is confirmed by the fact that β-carotene forms an adduct with five molecules of maleic anhydride (Nakamiya, 1936).

Geronic acid, on oxidation with cold aqueous potassium permanganate, forms a mixture of acetic acid, α : α-dimethylglutaric, III, α : α-dimethylsuccinic, IV, and dimethylmalonic acids, V.

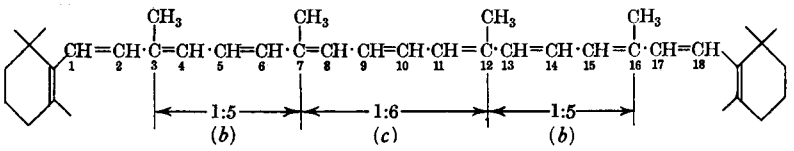


II III IV V

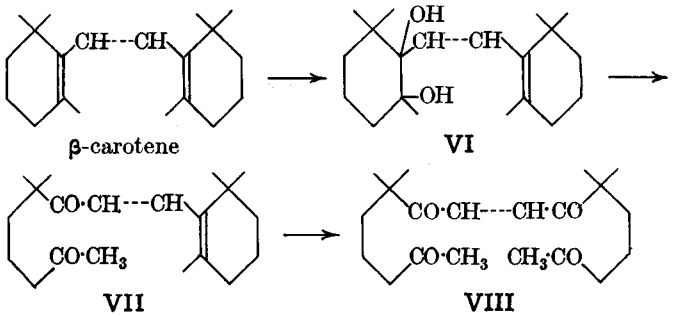
Oxidation of β-carotene in benzene solution with cold aqueous permanganate gives a mixture of β-ionone, III, IV, V, and acetic acid, the amount of acetic acid being more than can be accounted for by the presence of two β-ionone residues. Thus there must be some methyl side-chains in the central C_{14} portion of the molecule. Since it is essential to know the exact number of these methyl side-chains, this led to the development of the **Kuhn–Roth methyl side-chain determination** (1931). The first method used was to oxidise the carotenoid with alkaline permanganate, but later chromic acid (chromium trioxide in sulphuric acid) was found to be more reliable, the methyl group in the fragment $—C(CH_3)=$ being always oxidised to acetic acid. It was found that alkaline permanganate only oxidises the fragment $=C(CH_3)—CH=$ to acetic acid, and fragments such as $=C(CH_3)—CH_2—$ are incompletely oxidised to acetic acid, or not attacked at all (Karrer *et al.*, 1930). Since a molecule ending in an *iso*propylidene group also gives acetic acid on oxidation with chromic acid, this end group is determined by ozonolysis, the acetone so formed being estimated volumetrically. Application of the Kuhn–Roth methyl side-chain determination to β-carotene gave four molecules of acetic acid, thus indicating that there are four $—C(CH_3)=$ groups in the *chain*. The positions of two of these have already been tentatively placed in the two end β-ionone residues (see tentative structure above), and so the problem is now to find the positions of the remaining two. This was done as follows. Distillation of carotenoids under normal conditions brings about decomposition with the formation of aromatic compounds. Thus the distillation of β-carotene produces toluene, *m*-xylene and 2 : 6-dimethylnaphthalene (Kuhn *et al.*, 1933). The formation of these compounds may be explained by the cyclisation of fragments of the polyene chain, without the β-ionone rings being involved. The following types of chain fragments would give the desired aromatic products:



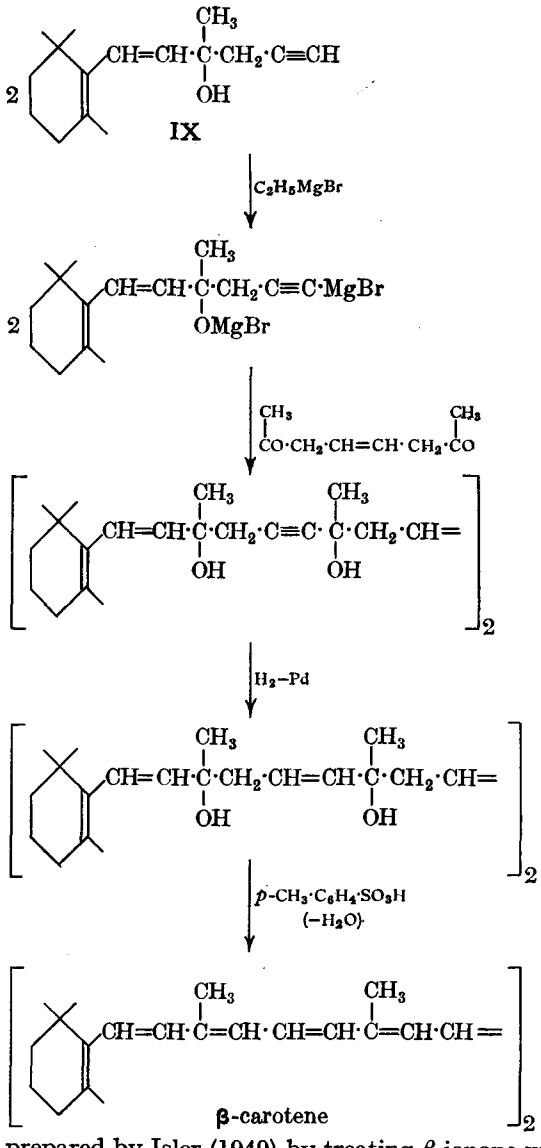
The following *symmetrical* structure for β-carotene would satisfy the requirements of (*a*), (*b*) and (*c*); the tail to tail union of the two isoprene units at the centre should be noted.



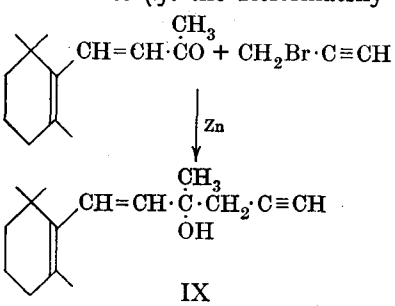
This symmetrical formula for β-carotene has been confirmed by the following oxidation experiments (Kuhn *et al.*, 1932–1935). When β-carotene is oxidised *rapidly* with potassium dichromate, dihydroxy-β-carotene, VI, is obtained and this, on oxidation with lead tetra-acetate, gives semi-β-carotenone, VII, a diketone. Since both VI and VII contain the *same* number of carbon atoms as β-carotene, it follows that the *double bond in one of the β-ionone rings* has been oxidised; otherwise there would have been chain scission had the chain been oxidised. Oxidation of semi-β-carotenone with chromium trioxide produces β-carotenone, VIII, a tetraketone which also has the same number of carbon atoms as β-carotene. Thus, in this compound, the *other* β-ionone ring is opened. Now only *one* dihydroxy-β-carotene and *one* semi-β-carotenone are obtained, and this can be explained only by assuming a symmetrical structure for β-carotene. Thus the oxidations may be formulated:



This structure for β-carotene has been confirmed by synthesis, *e.g.*, that of Karrer *et al.* (1950). The acetylenic carbinol IX is treated with ethylmagnesium bromide and the product is treated as shown on opposite page.



IX has been prepared by Isler (1949) by treating β-ionone with propargyl bromide in the presence of zinc (*cf.* the Reformatsky reaction):



The most convenient way of preparing the diketone (oct-4-ene-2 : 7-dione) starts with but-1-yn-3-ol (Inhoffen *et al.*, 1951):

$$2CH_3{\cdot}CHOH{\cdot}C{\equiv}CH \xrightarrow[CuCl-NH_4Cl]{O_2} CH_3{\cdot}CHOH{\cdot}C{\equiv}C{\cdot}C{\equiv}C{\cdot}CHOH{\cdot}CH_3 \xrightarrow{LiAlH_4}$$

$$CH_3{\cdot}CHOH{\cdot}CH{=}CH{\cdot}CH{=}CH{\cdot}CHOH{\cdot}CH_3 \xrightarrow{MnO_2}$$

$$CH_3{\cdot}CO{\cdot}CH{=}CH{\cdot}CH{=}CH{\cdot}CO{\cdot}CH_3 \xrightarrow[CH_3{\cdot}CO_2H]{Zn}$$

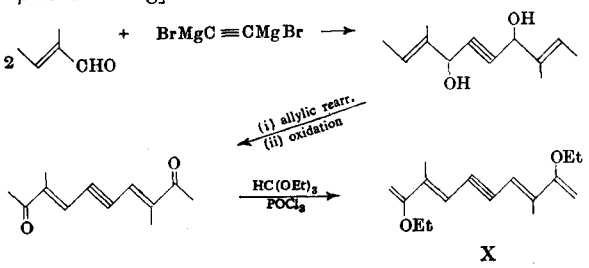
$$CH_3{\cdot}CO{\cdot}CH_2{\cdot}CH{=}CH{\cdot}CH_2{\cdot}CO{\cdot}CH_3$$

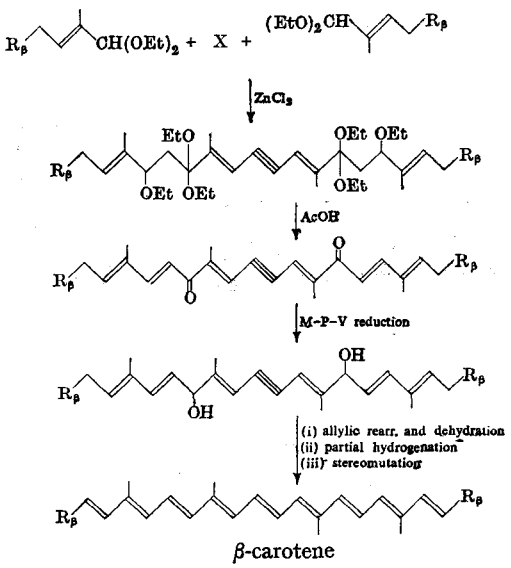
An important point to note in this synthesis is that lithium aluminium hydride will reduce a triple bond to a double bond when the former is adjacent to a propargylic hydroxyl group, *i.e.*,

$$-C(OH){\cdot}C{\equiv}C- \xrightarrow{LiAlH_4} -C(OH){\cdot}CH{=}CH-$$

It is worth while at this point to consider the general aspects of carotene syntheses. All syntheses have used the union of a bifunctional unit, which forms the central part of the carotene molecule, with two molecules (identical as for, *e.g.*, β-carotene, or not identical as for, *e.g.*, α-carotene). The various methods have been divided into four groups according to the carbon content of the three units used in the synthesis: $C_{19} + C_2 + C_{19}$; $C_{16} + C_8 + C_{16}$; $C_{14} + C_{12} + C_{14}$; $C_{10} + C_{20} + C_{10}$. The second group has been used in the above synthesis of β-carotene.

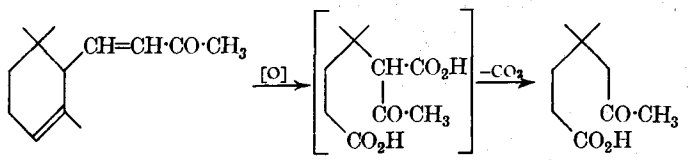
An example of the synthesis of β-carotene by the third is that of Isler *et al.* (1957) [R_β = β-ionine ring]:





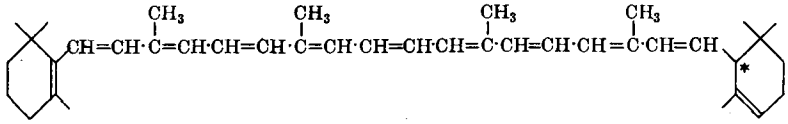
An example of the fourth group makes use of the Wittig reaction (see crocetin, §9 for an illustration of this method).

§4. α-Carotene, $C_{40}H_{56}$. This is isomeric with β-carotene, and oxidation experiments on α-carotene have led to results similar to those obtained for β-carotene, except that *iso*geronic acid is obtained as well as geronic acid. Since *iso*geronic acid is an oxidation product of α-ionone, the conclusion is that α-carotene contains one β-ionone ring and one α-ionone ring (§6. VIII) [Karrer *et al.*, 1933].

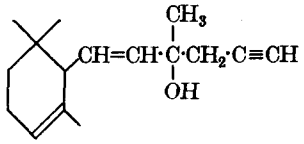


α-ionone *iso*geronic acid

Thus the structure of α-carotene is:



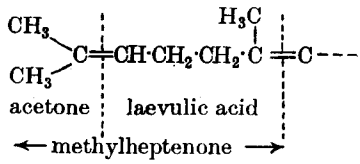
As we have seen, α-carotene is optically active (§1), and this is due to the presence of the asymmetric carbon atom (*) in the α-ionone ring. The structure given for α-carotene has been confirmed by synthesis (Karrer *et al.*, 1950). The method is the same as that described for β-carotene, except that *one* molecule of the acetylenic alcohol (structure IX, §3) is used together with one molecule of the corresponding α-ionone derivative:



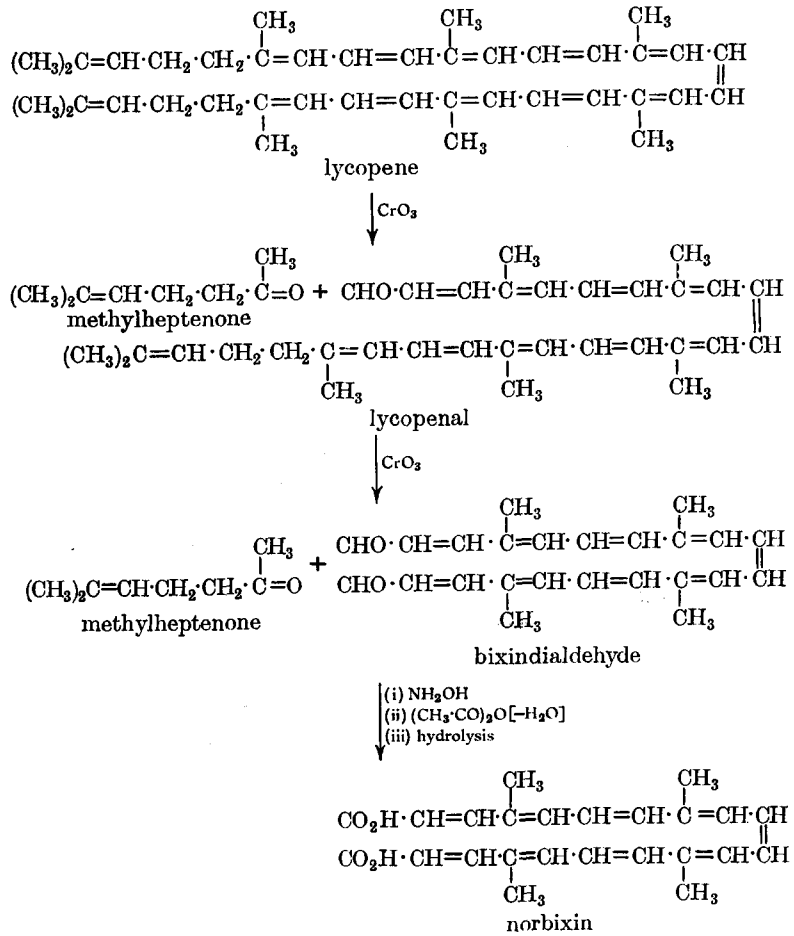
It is interesting to note that α-carotene has been converted into the β-isomer by heating the α-compound with ethanolic sodium ethoxide and benzene at 100–110° for some time (Karrer *et al.*, 1947); this is an example of *three carbon prototropy*.

§5. Lycopene, $C_{40}H_{56}$, m.p. 175°, is a carotenoid that is the tomato pigment. Since the structure of γ-carotene depends on that of lycopene, the latter will be discussed here, and the former in the next section.

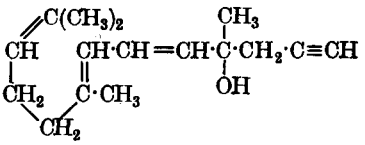
On catalytic hydrogenation (platinum), lycopene is converted into perhydrolycopene, $C_{40}H_{82}$. Therefore lycopene has thirteen double bonds, and is an acyclic compound (Karrer *et al.*, 1928). Ozonolysis of lycopene gives, among other products, acetone and lævulic acid; this suggests that lycopene contains the terminal residue:



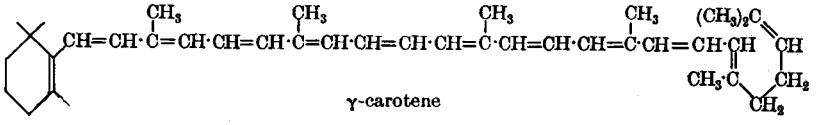
This is supported by the fact that controlled oxidation of lycopene with chromic acid produces 6-methylhept-5-en-2-one (*cf.* §5. VIII). Quantitative oxidation experiments (ozonolysis) indicate that this grouping occurs at each end of the molecule (Karrer *et al.*, 1929, 1931). Also, the quantitative oxidation of lycopene with chromic acid gives six molecules of acetic acid per molecule of lycopene, thereby suggesting that there are six $-C(CH_3)=$ groups present in the chain (*cf.* §3). Controlled oxidation of lycopene with chromic acid gives one molecule of methylheptenone and one molecule of lycopenal, $C_{32}H_{42}O$, and the latter may be further oxidised with chromic acid to another molecule of methylheptenone and one molecule of a dialdehyde, $C_{24}H_{28}O_2$ (Kuhn *et al.*, 1932). Thus this dialdehyde constitutes the central part of the chain, and the two molecules of methylheptenone must have been produced by the oxidation of each end of the chain in lycopene. The dialdehyde may be converted into the corresponding dioxime, and this, on dehydration to the dicyanide, followed by hydrolysis, forms the dicarboxylic acid $C_{24}H_{28}O_4$, which is identical with norbixin (§9). Thus the dialdehyde must be bixindialdehyde, and so it may be inferred that the structure of lycopene is the following symmetrical one, since it accounts for all the above facts.



The structure assigned to lycopene has been confirmed by synthesis (Karrer *et al.*, 1950). Instead of the acetylenic carbinol IX in §3, two molecules of the following compound were used.

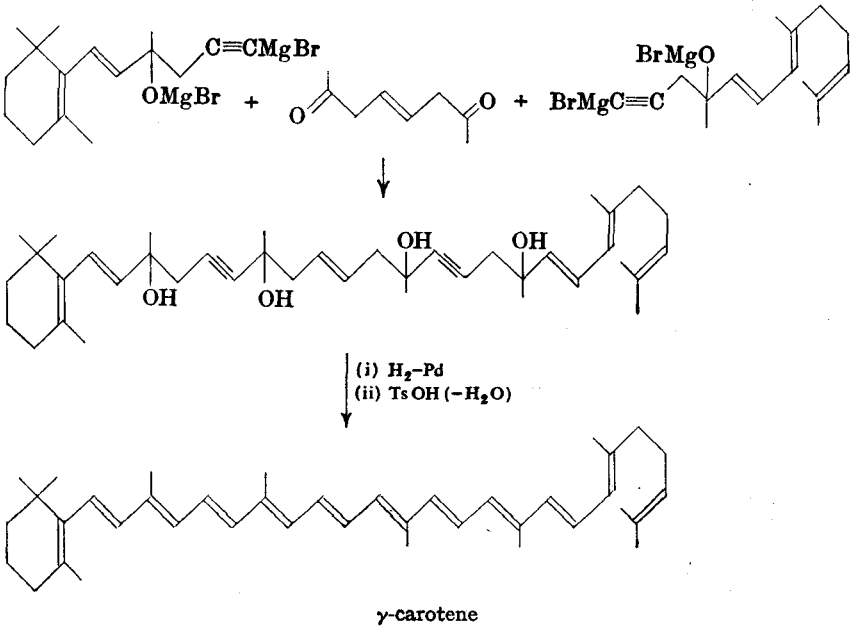


§6. γ-Carotene, $C_{40}H_{56}$. Catalytic hydrogenation converts γ-carotene into perhydro-γ-carotene, $C_{40}H_{80}$. Thus there are twelve double bonds present, and the compound contains one ring. Ozonolysis of γ-carotene gives, among other products, acetone, lævulic acid and geronic acid. The formation of acetone and lævulic acid indicates the structural relationship of γ-carotene to lycopene, and the formation of geronic acid indicates the presence of a β-ionone ring (Kuhn *et al.*, 1933). On this evidence, and also on the fact that the growth-promoting response in rats was found to be half that of β-carotene, Kuhn suggested that γ-carotene consists of half a molecule of β-carotene joined to half a molecule of lycopene; thus:



γ-carotene

This structure for γ-carotene is supported by the fact that the absorption maximum of γ-carotene in the visible region lies between that of β-carotene and that of lycopene. Final proof for this structure has been obtained by the synthesis of γ-carotene (Karrer *et al.*, 1953); the following reactions are written with the conventional formulæ (see §1):



γ-carotene

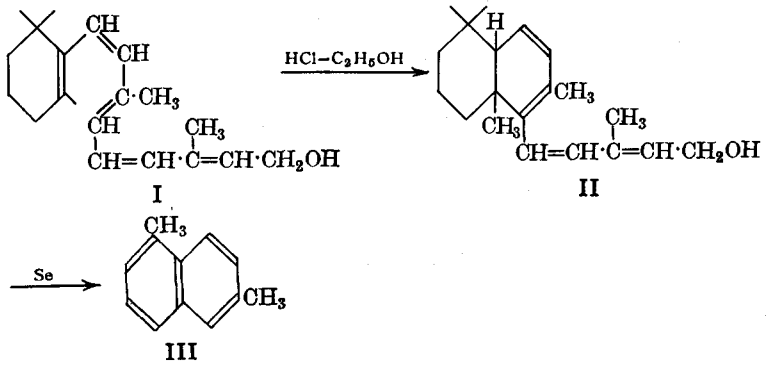
A δ-carotene has also been isolated, and this has been shown to be the α-ionone analogue of γ-carotene (Kargel *et al.*, 1960).

§7. **Vitamin A,** $C_{20}H_{30}O$. Vitamin A is also known as **Axerophthol,** and is also usually referred to as vitamin A_1 since a second compound, known as vitamin A_2, has been isolated.

Vitamin A_1 influences growth in animals, and also apparently increases resistance to disease. Night blindness is due to vitamin A_1 deficiency in the human diet, and a prolonged deficiency leads to xerophthalmia (hardening of the cornea, etc.). Vitamin A_1 occurs free and as esters in fats, in fish livers and in blood. It was originally isolated as a viscous yellow oil, but later it was obtained as a crystalline solid, m.p. 63–64° (Baxter *et al.*, 1940). Vitamin A_1 is estimated by the blue colour reaction it gives with a solution of antimony trichloride in chloroform (the Carr–Price reaction; *cf.* §1); it is also estimated by light absorption (vitamin A_1 has a maximum at 328 mμ).

Carotenoids are converted into vitamin A_1 in the intestinal mucosa, and feeding experiments showed that the potency of α- and γ-carotenes is half that of β-carotene. This provitamin nature of β-carotene led to the suggestion that vitamin A_1 is half the molecule of β-carotene.

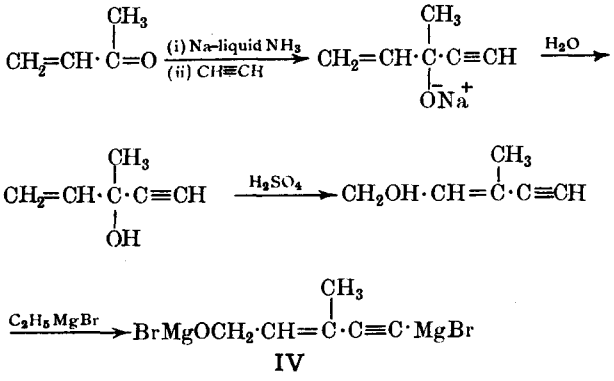
On catalytic hydrogenation, vitamin A_1 is converted into perhydrovitamin A_1, $C_{20}H_{40}O$; thus vitamin A_1 contains five double bonds. Since vitamin A_1 forms an ester with *p*-nitrobenzoic acid (this ester is not crystallisable), it follows that vitamin A_1 contains a hydroxyl group. Thus the parent hydrocarbon of vitamin A_1 is $C_{20}H_{40}$, and consequently the molecule contains one ring. Ozonolysis of vitamin A_1 produces one molecule of geronic acid (§3) per molecule of vitamin A_1, and so there must be one β-ionone nucleus present (Karrer, 1931, 1932). Oxidation of vitamin A_1 with permanganate produces acetic acid; this suggests that there are some $-C(CH_3)=$ groups in the chain. All of the foregoing facts are in keeping with the suggestion that vitamin A_1 is half the β-carotene structure. When heated with an ethanolic solution of hydrogen chloride, vitamin A_1 is converted into some compound (II) which, on dehydrogenation with selenium forms 1 : 6-dimethylnaphthalene, III (Heilbron *et al.*, 1932). Heilbron *assumed* I as the structure of vitamin A_1, and explained the course of the reaction as follows:



Perhydrovitamin A_1 has been synthesised from β-ionone (Karrer, 1933), and was shown to be identical with the compound obtained by reducing vitamin A_1; thus there is evidence to support the structure assigned to vitamin A_1. Final proof of structure must rest with a synthesis of vitamin A_1 itself, and this has now been accomplished by several groups of workers.

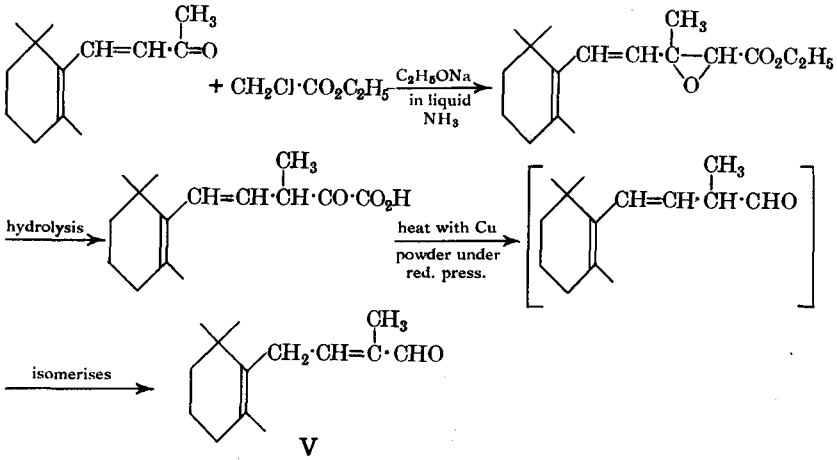
The following synthesis is that of Isler *et al.* (1947). This starts with methyl vinyl ketone to produce compound IV, one stage of the reactions involving

Preparation of IV.



IV

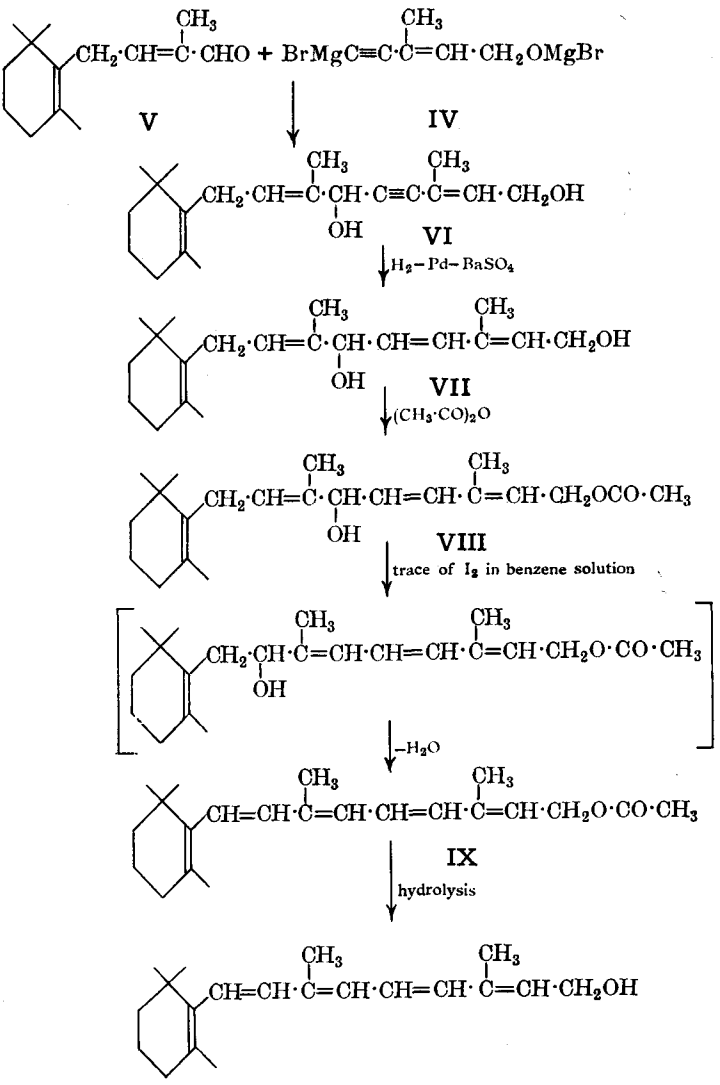
Preparation of V.



V

an allylic rearrangement (*cf.* §8. VIII). Compound V is prepared from β-ionone by means of the Darzens glycidic ester reaction (see also Vol. I). The following chart shows the steps of the synthesis, and it should be noted that another allylic rearrangement is involved in one of the later steps.

Combination of IV and V, etc.

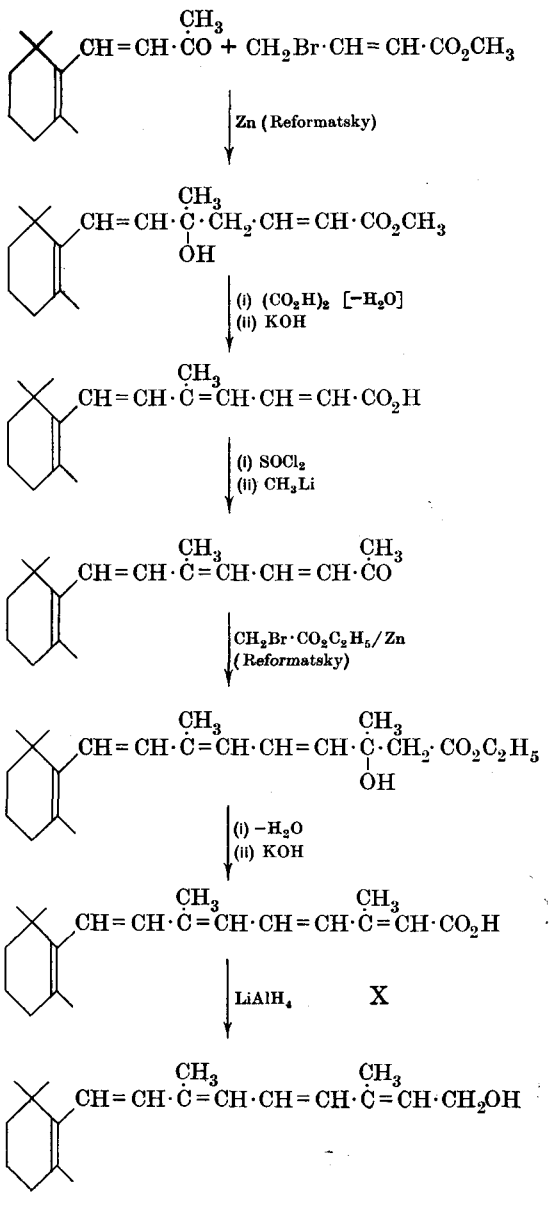


In the hydrogenation of VI to VII, barium sulphate is used to act as a poison to the catalyst to prevent hydrogenation of the *double* bonds. Partial acetylation of VII (primary alcoholic groups are more readily acetylated than secondary) protects the terminal group from an allylic rearrangement in the conversion of VIII to IX.

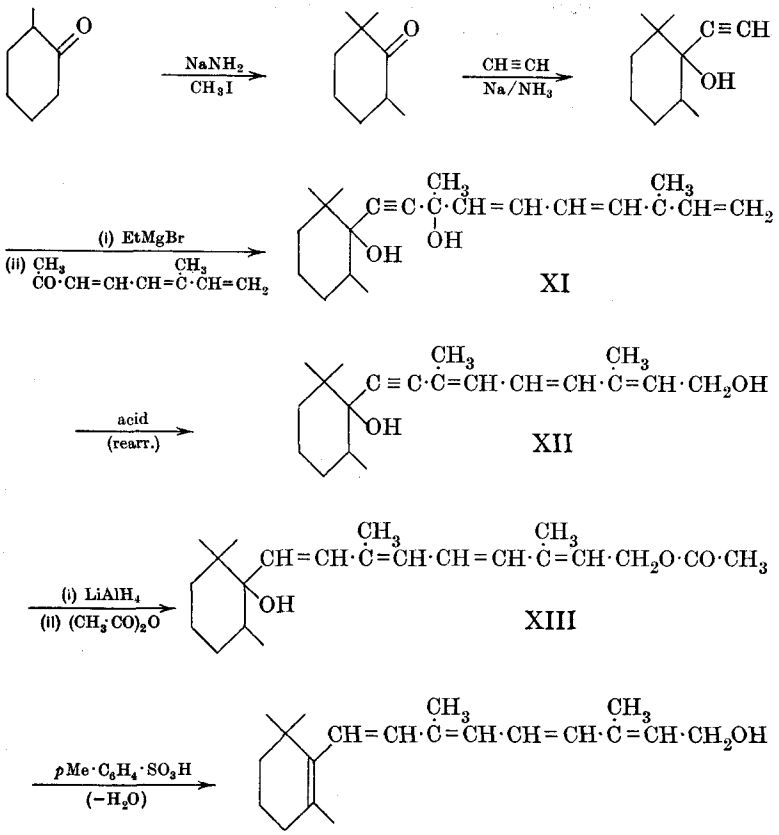
The crude vitamin A_1 obtained in the above synthesis was purified *via* its ester with anthraquinone-2-carboxylic acid, and was thereby obtained in a crystalline form which was shown to be identical with natural vitamin A_1.

Lindlar (1952) has shown that triple bonds may be partially hydrogenated in the presence of a Pd—$CaCO_3$ catalyst that has been partially inactivated by treatment with lead acetate; better results are obtained by the addition of quinoline. Thus the hydrogenation of VI gives VII in 86 per cent. yield when the Lindlar catalyst is used.

Another method of synthesising vitamin A_1 is due to van Dorp *et al.* (1946) who prepared vitamin A_1 acid (X), which was then reduced by means of lithium aluminium hydride to vitamin A_1 by Tishler (1949); β-ionone and methyl γ-bromocrotonate are the starting materials.



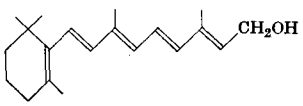
Attenburrow *et al.* (1952) have also synthesised vitamin A_1 starting from 2-methyl*cyclo*hexanone.



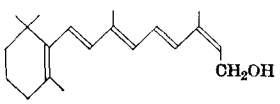
Acid causes rearrangement of XI to XII in which all multiple bonds are in complete conjugation, and the reduction of XII to XIII by lithium aluminium hydride is possible because of the presence of the propargylic hydroxyl grouping (§3).

Synthetic vitamin A_1 is now a commercial product.

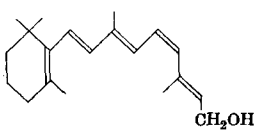
Two biologically active geometrical isomers of Vitamin A_1 (all-*trans*) have also been isolated: **neovitamin a** from rat liver (Robeson *et al.*, 1947) and **neovitamin b** from the eye (Oroshnik *et al.*, 1956). Vitamin A_1 is the most active form in curing " vitamin A " deficiency.



vitamin A_1

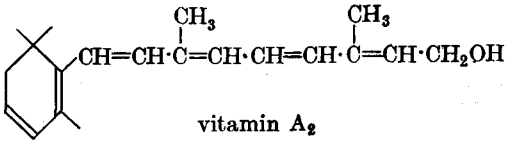


neovitamin a



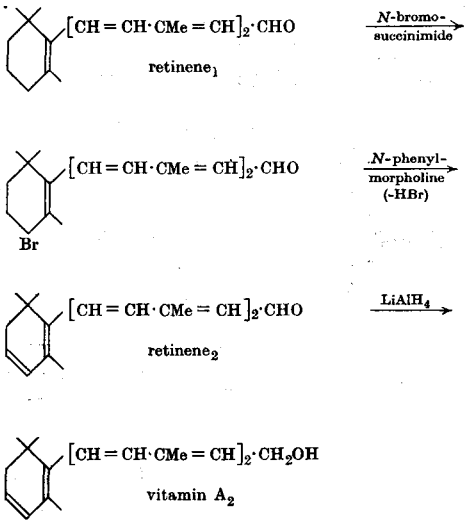
neovitamin b

Vitamin A_2. A second vitamin A, vitamin A_2, has been isolated from natural sources, and has been synthesised by Jones *et al.* (1951, 1952); it is dehydrovitamin A_1.



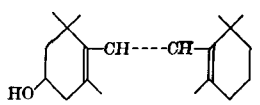
vitamin A_2

Jones *et al.* (1955) have also introduced a method for converting vitamin A_1 into vitamin A_2. Vitamin A_1 may be oxidised to vitamin A_1 aldehyde (**retinene$_1$**) by means of manganese dioxide in acetone solution (Morton *et al.*, 1948), and then treated as follows:

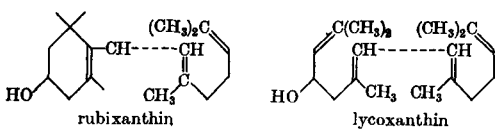


§8. Xanthophylls. The xanthophylls occur naturally, and all have the same carbon skeletons as the carotenes or lycopene (except flavoxanthin).

Cryptoxanthin, $C_{40}H_{56}O$, m.p. 169°, is monohydroxy-β-carotene; it has provitamin-A activity.

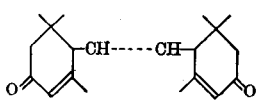


Rubixanthin, $C_{40}H_{56}O$, m.p. 160°, is monohydroxy-γ-carotene, and **lycoxanthin,** $C_{40}H_{56}O$, m.p. 168°, appears to be monohydroxylycopene.

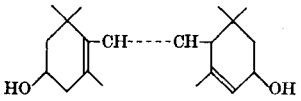


rubixanthin lycoxanthin

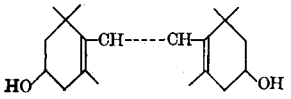
Rhodoxanthin, $C_{40}H_{52}O_2$, m.p. 219°, is believed to be the following diketone.



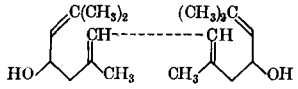
Lutein, $C_{40}H_{56}O_2$, m.p. 193°, was formerly known as xanthophyll; it is dihydroxy-α-carotene.



Zeaxanthin, m.p. 205°, and **lycophyll,** m.p. 179°, are the corresponding dihydroxy derivatives of β-carotene and lycopene, respectively.



zeaxanthin



lycophyll

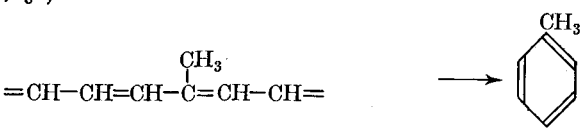
§9. Carotenoid acids. These are compounds which do not contain 40 carbon atoms.

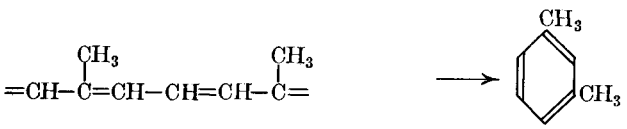
Bixin, $C_{25}H_{30}O_4$. Natural bixin is a brown solid, m.p. 198°, and is the *cis*-form; it is readily converted into the more stable *trans*-form, m.p. 216–217°.

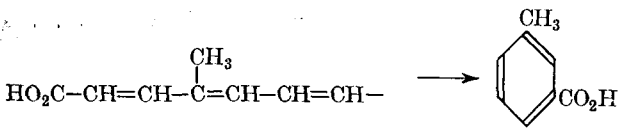
When boiled with potassium hydroxide solution, bixin produces one molecule of methanol and a dipotassium salt which, on acidification, gives the dibasic acid **norbixin,** $C_{24}H_{28}O_4$. Thus bixin is a monomethyl ester, and can be esterified to give methylbixin.

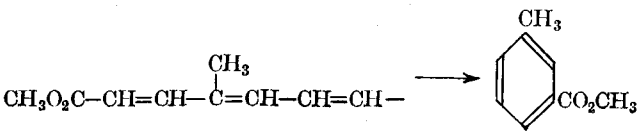
On catalytic hydrogenation, bixin is converted into perhydrobixin, $C_{25}H_{48}O_4$; thus there are 9 double bonds present in the molecule (Liebermann *et al.*, 1915). Perhydrobixin, on hydrolysis, forms perhydronorbixin. Oxidation of bixin with permanganate produces four molecules of acetic acid (Kuhn *et al.*, 1929); thus there are four $-C(CH_3)=$ groups in the chain. Furthermore, since the parent hydrocarbon of perhydronorbixin, $C_{24}H_{46}O_4$, is $C_{22}H_{46}$ (the two carboxyl groups are regarded as substituents), the molecule is acyclic.

The thermal decomposition of bixin produces toluene, *m*-xylene, *m*-toluic acid and the methyl ester of this acid (Kuhn *et al.*, 1932). Hence the following assumptions may be made regarding the nature of the chain (*cf.* β-carotene, §3).

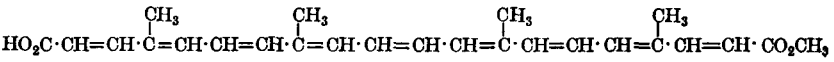






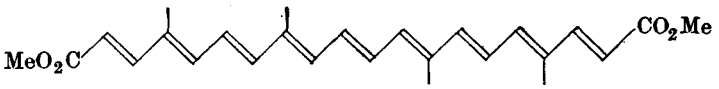


The foregoing facts may be explained by assuming the following structure for bixin (Kuhn *et al.*, 1932):



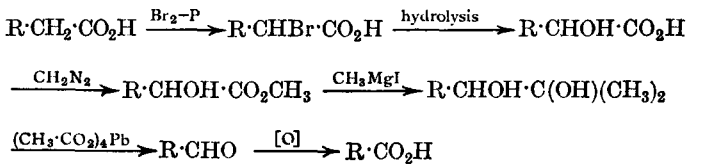
This structure is supported by the fact that perhydronorbixin has been synthesised, and shown to be identical with the compound obtained from the reduction of bixin (Karrer *et al.*, 1933). Further proof is the synthesis of norbixin (Isler *et al.*, 1957).

Jackman *et al.* (1960) have shown, from an examination of the NMR spectra (§19a. I) of many carotenoids, that the positions of the absorption bands resulting from the methyl groups give some indication of the molecular environment of these groups. "Natural" methylbixin is the *cis*-isomer of the following *trans*-isomer:

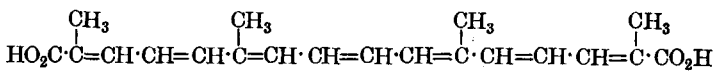


The methyl ester of crocetin (see below) also probably has the *cis*-configuration at the corresponding 2,3-position.

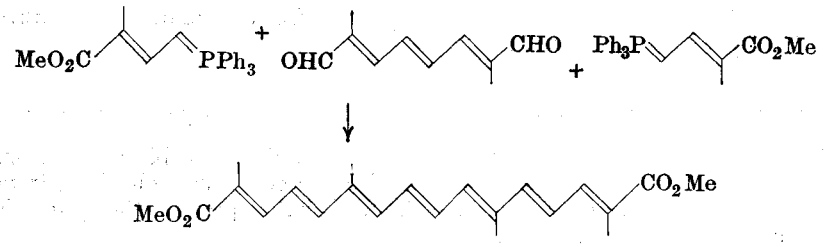
Crocetin, $C_{20}H_{24}O_4$. Crocetin occurs in saffron as the digentiobioside, *crocin*. The structure of crocetin was elucidated by Karrer *et al.* (1928) and Kuhn *et al.* (1931). Crocetin behaves as a dicarboxylic acid and has seven double bonds (as shown by catalytic hydrogenation to perhydrocrocetin, $C_{20}H_{38}O_4$). On oxidation with chromic acid, crocetin gives 3–4 molecules of acetic acid per molecule of crocetin; thus there are 3–4 methyl side-chains. The structure of crocetin was finally shown by the degradation of perhydronorbixin, $C_{24}H_{46}O_4$, by means of the following method:



This set of reactions was performed *twice* on perhydronorbixin, thereby resulting in the loss of four carbon atoms (two from each end); the product so obtained was perhydrocrocetin, $C_{20}H_{38}O_4$. On these results, crocetin is therefore:



This structure is supported by the fact that the removal of two carbon atoms from perhydrocrocetin by the above technique (one carbon atom is lost from each end) resulted in the formation of a *diketone*. The formation of this compound shows the presence of an α-methyl group at each end of the molecule. The structure of crocetin is further supported by the synthesis of perhydrocrocetin, and by the synthesis of crocetin diesters by Isler *et al.* (1957). These diesters probably have the *cis*-configuration at the 2,3-position (see bixin, above). The *trans*-crocetin dimethyl ester has been synthesised by the Wittig reaction (a carbonyl group is exchanged for a methylene group; Vol. I) between the dialdehyde and two molecules of the phosphorane (Buchta *et al.*, 1959, 1960).



READING REFERENCES

Karrer and Jucker, *Carotenoids*, Elsevier (translated and revised by Braude, 1950).
Rodd (Ed.), *Chemistry of the Carbon Compounds*, Elsevier. Vol. 11A (1953). Ch. 10. The Carotenoid Group.
Gilman (Ed.), *Advanced Organic Chemistry*, Wiley. Vol. IV (1953). Ch. 7. The Terpenes (see the section on Tetraterpenes).
Bentley, *The Natural Pigments*, Interscience (1960).

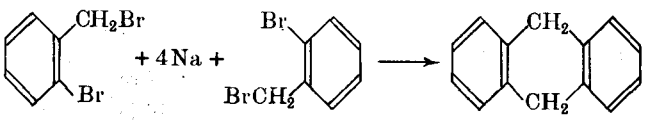
CHAPTER X

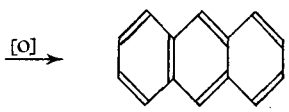
POLYCYCLIC AROMATIC HYDROCARBONS

§1. Introduction. Naphthalene, anthracene, phenanthrene, fluorene, etc., have been described in Volume I. All these compounds occur in coal-tar, but also present are many polycyclic hydrocarbons containing four or more rings, and others of this type have been synthesised.

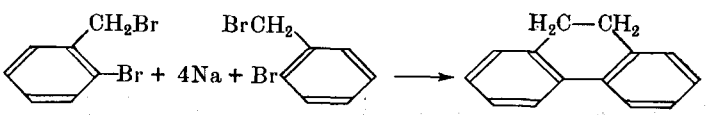
§2. General methods of preparation of polycyclic hydrocarbons. Before dealing with a number of individual hydrocarbons, it is instructive to review some of the general methods whereby these polycyclic hydrocarbons may be prepared (see also Vol. I).

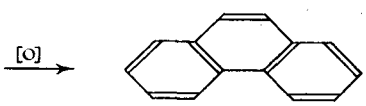
(i) **Fittig reaction,** *e.g.*, anthracene and phenanthrene may be prepared by the action of sodium on *o*-bromobenzyl bromide.





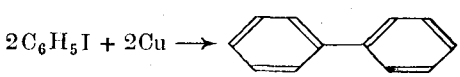
anthracene





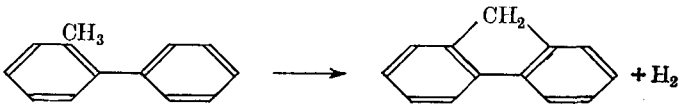
phenanthrene

(ii) **Ullmann diaryl synthesis.** This method results in the formation of isolated polynuclear compounds, *e.g.*, heating iodobenzene with copper powder in a sealed tube produces diphenyl.

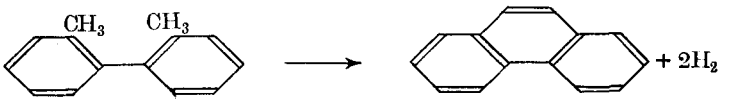


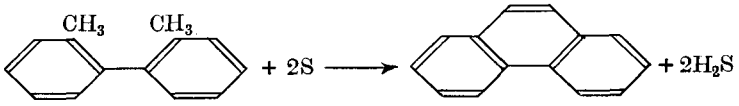
Compounds of the isolated system type can, under suitable conditions, be converted into condensed polycyclic compounds (see method iii). In certain cases, the Ullmann synthesis leads to condensed systems (see §4c).

(iii) Many compounds of the isolated system type can be converted into condensed systems by strong heating, *e.g.*, *o*-methyldiphenyl forms fluorene. 2 : 2′-Dimethyldiphenyl forms phenanthrene when passed through a red-hot

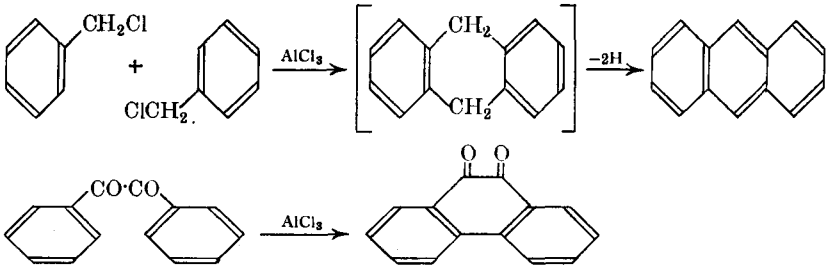


tube, but a much better yield is obtained when the dimethyldiphenyl is heated with sulphur. The latter is an example of cyclodehydrogenation (see also method vii).

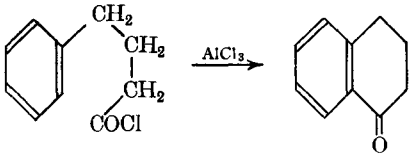




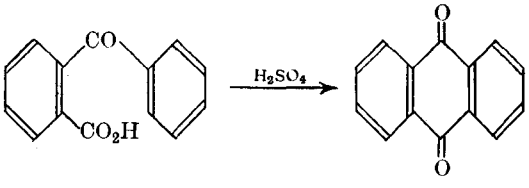
(iv) **Friedel–Crafts reaction.** Condensed polycyclic compounds may be prepared *via* an external or an internal Friedel–Crafts reaction. An example of the former is the preparation of anthracene from benzyl chloride; an example of the latter is the preparation of phenanthraquinone from benzil.



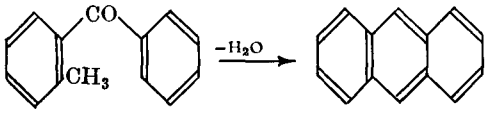
A very important case of the internal Friedel–Crafts reaction is that in which ring closure is effected on acid chlorides, *e.g.*, the conversion of γ-phenylbutyryl chloride to α-tetralone.



This type of ring closure may be effected by the action of concentrated sulphuric acid on the carboxylic acid itself, *e.g.*,

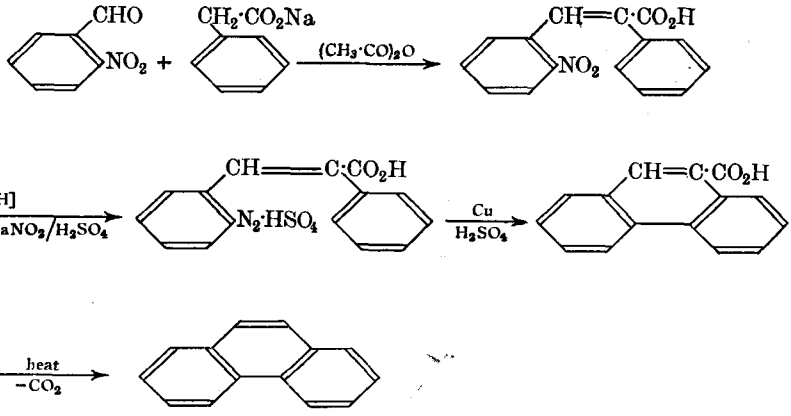


(v) **Elbs reaction.** In this method, polynuclear hydrocarbons are produced from a diaryl ketone containing a methyl group in the *o*-position to the keto group. The reaction is usually carried out by heating the ketone under reflux or at 400–450° until water is no longer evolved, *e.g.*, *o*-methylbenzophenone forms anthracene.

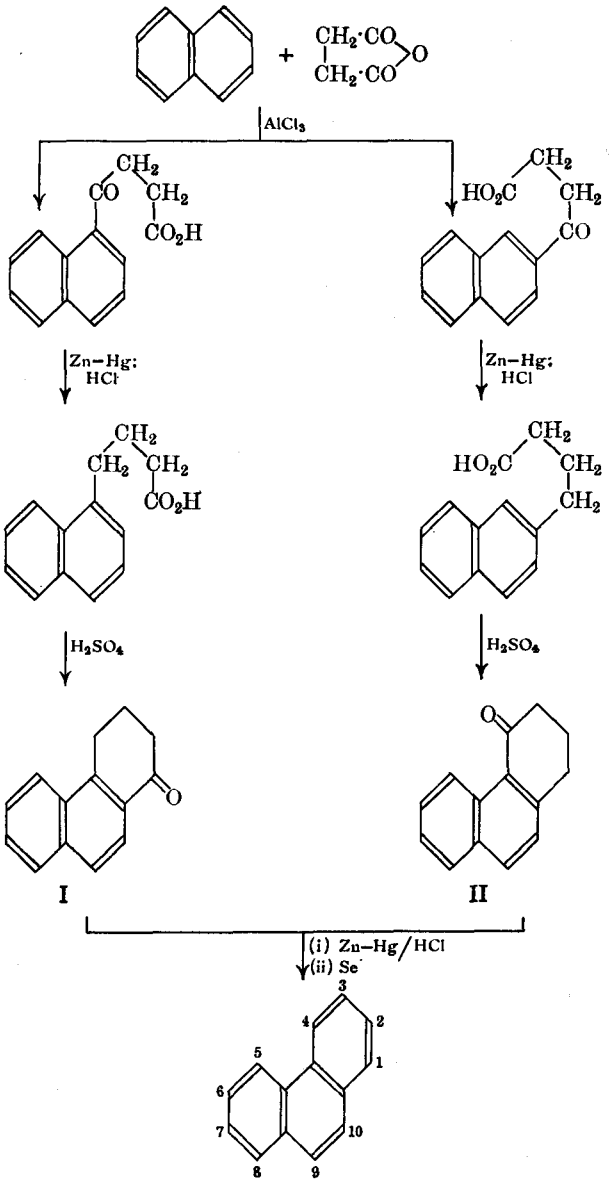


(vi) **Phenanthrene syntheses.** The phenanthrene nucleus is particularly important in steroid chemistry, and so a number of methods for synthesising phenanthrene are dealt with in some detail.

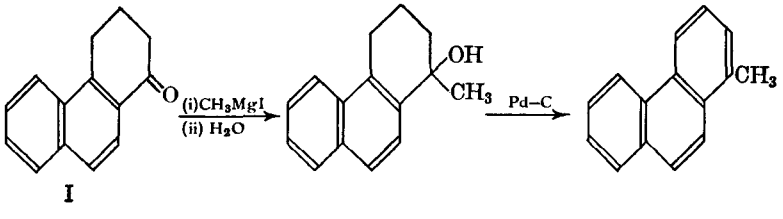
(*a*) **Pschorr synthesis** (1896). This method offers a means of preparing phenanthrene and substituted phenanthrenes with the substituents in known positions. Phenanthrene may be prepared as follows, starting with *o*-nitrobenzaldehyde and sodium *β*-phenylacetate.



(*b*) **Haworth synthesis** (1932). Naphthalene is condensed with succinic anhydride in the presence of aluminium chloride in nitrobenzene solution. Two naphthoylpropionic acids are obtained, and these may be separated (see next page).

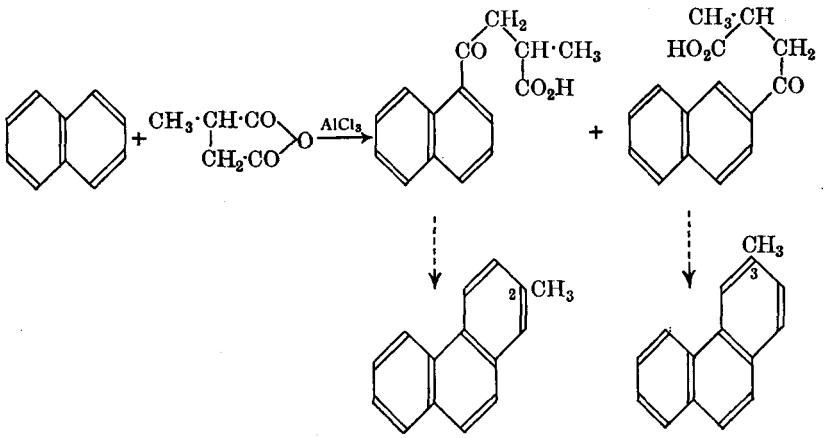


The Haworth synthesis is very useful for preparing alkylphenanthrenes with the alkyl group in position 1 (from I) or position 4 (from II); *e.g.*,

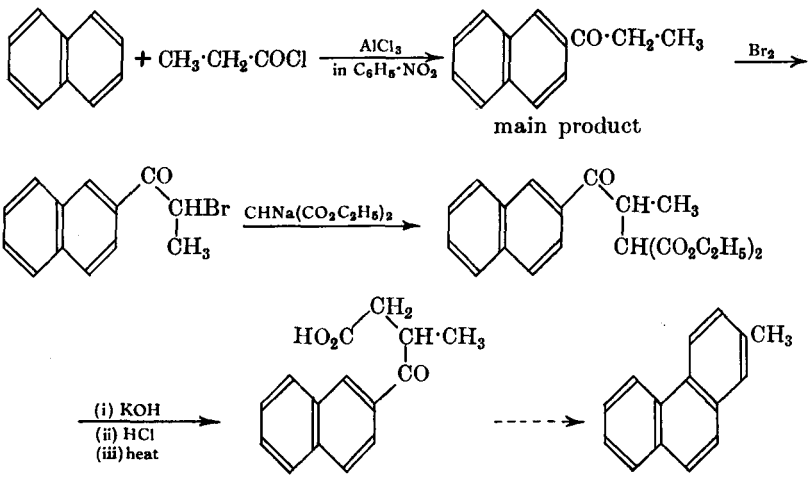


I

By using methylsuccinic anhydride instead of succinic anhydride, a methyl group can be introduced into the 2- or 3-position; in this case the condensation occurs at the less hindered keto group, *i.e.*, at the one which is farther removed from the methyl substituent.

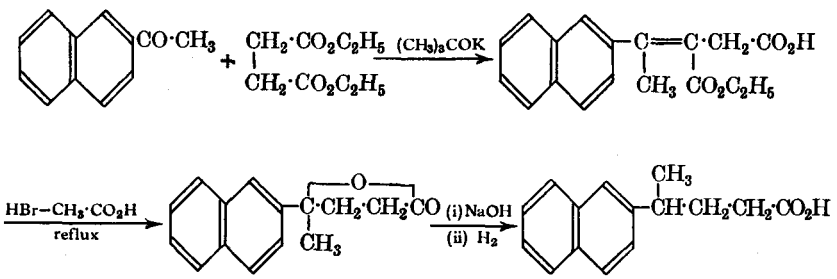


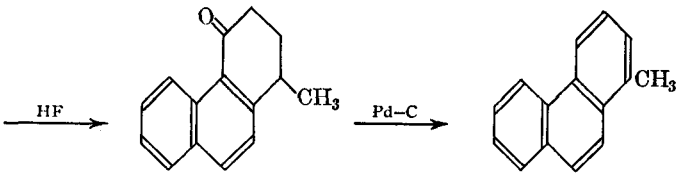
α-Bromoketone derivatives of naphthalene may be used in the malonic ester synthesis to prepare alkylphenanthrenes, *e.g.*,



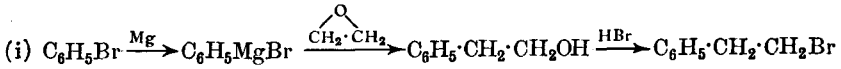
main product

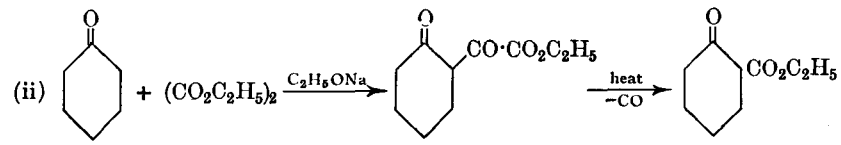
(*c*) **Stobbe condensation** (1893). This method has been improved by Johnson (1944), and has been used to prepare phenanthrene derivatives (see Vol. I); *e.g.*,



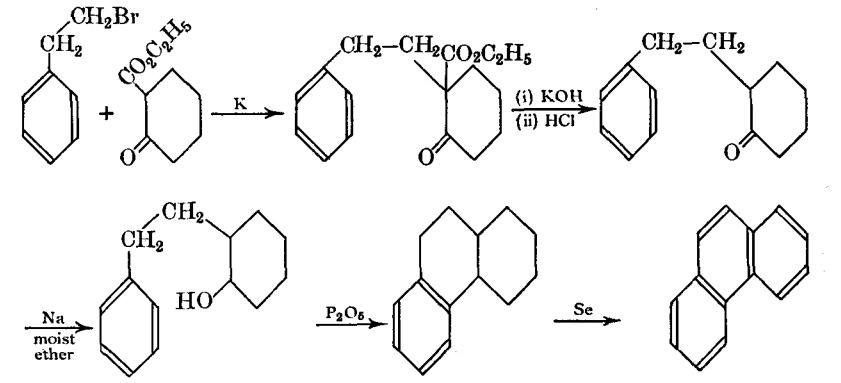


(*d*) **Bardhan–Sengupta synthesis** (1932). In this synthesis the starting materials are 2-phenylethyl bromide and ethyl *cyclo*hexane-2-carboxylate; these may be prepared as follows:

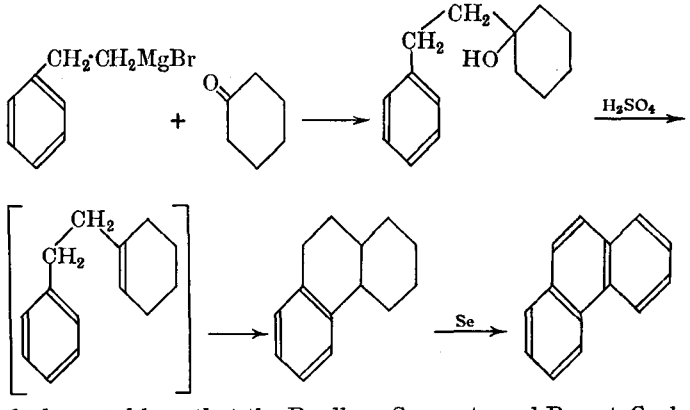




These two compounds are then treated as shown:

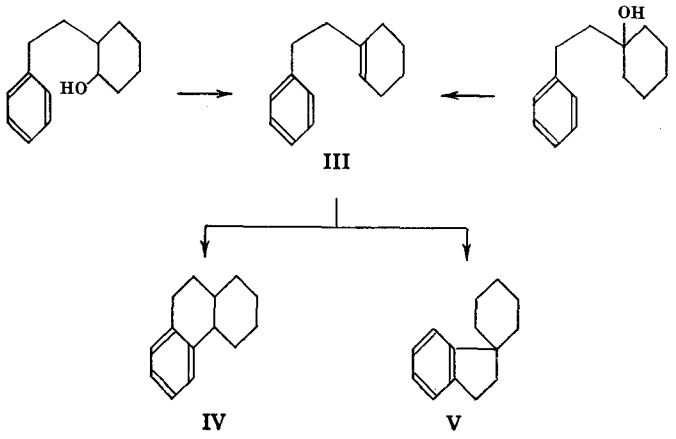


(*e*) **Bogert–Cook synthesis** (1933). The following chart shows the preparation of phenanthrene.



It might be noted here that the Bardhan–Sengupta and Bogert–Cook methods

both proceed *via* the formation of olefin III, which then gives a mixture of octahydrophenanthrene IV and the spiran V.

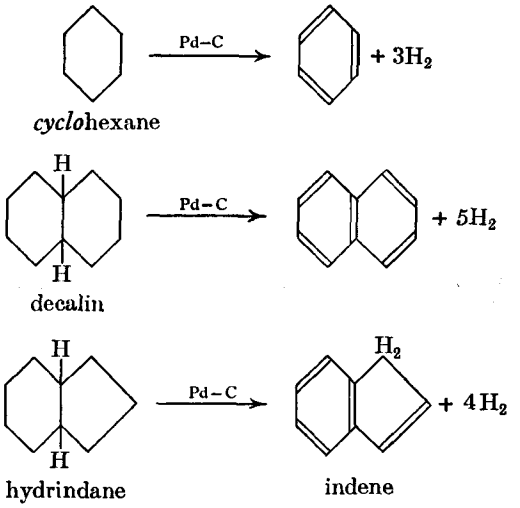


(vii) **Dehydrogenation of hydroaromatic compounds with sulphur, selenium or palladised charcoal.** This method is mainly confined to the dehydrogenation of six-membered rings, but five-membered rings may sometimes be dehydrogenated when they are fused to a six-membered ring. The general methods are as follows:

(*a*) Heating the compound with the calculated amount of sulphur at 200–220°; hydrogen is eliminated as hydrogen sulphide (Vesterberg, 1903).

(*b*) Heating the compound with the calculated amount of selenium at 250–280°; hydrogen is eliminated as hydrogen selenide (Diels, 1927).

(*c*) Heating the compound with palladium-charcoal up to about 300°, or passing the vapour of the compound over the catalyst heated at 180–350°; hydrogen is eliminated catalytically. Simple examples of catalytic dehydrogenation are:



Perhydro-compounds, *i.e.*, fully hydrogenated compounds, are readily dehydrogenated catalytically, but are very little affected, if at all, by the chemical reagents sulphur and selenium. Partially unsaturated compounds, however, are readily dehydrogenated by sulphur and selenium.

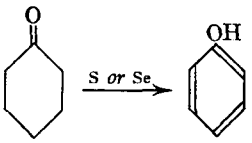
The method of dehydrogenation has been very useful in the elucidation of structure in terpene and steroid chemistry; specific examples are described in these two chapters. The following is an account of some of the general problems involved in dehydrogenation.

Originally, dehydrogenation was applied almost entirely to hydrocarbons, but subsequently it was found that many compounds containing certain functional groups could also be dehydrogenated, the nature of the products depending on the nature of the functional group.

(i) Alcoholic groups may be eliminated with the formation of unsaturated hydrocarbons, *e.g.*, eudesmol gives eudalene (§28b. VIII); cholesterol gives Diels' hydrocarbon (§1. XI).

(ii) Phenolic hydroxyl groups and methylated phenolic groups are usually unaffected by dehydrogenation with sulphur. With selenium, these groups may or may not be eliminated, but the higher the temperature at which the dehydrogenation is carried out (particularly above 300°), the greater the likelihood of these groups being eliminated.

(iii) The products obtained from ketones depend on whether the keto group is in a ring or in an open chain. Thus cyclic ketones are dehydrogenated to phenols, *e.g.*,

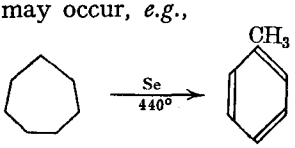


When the keto group is in a side-chain, then it is often unaffected.

(iv) Carboxyl (or carboalkoxyl) groups are eliminated when attached to a tertiary carbon atom, *e.g.*, abietic acid gives retene (§31. VIII). If, however, the carboxyl group is attached to a primary or secondary carbon atom, it is usually unaffected when the dehydrogenation is carried out with sulphur or palladium-charcoal. On the other hand, the carboxyl group is usually eliminated (decarboxylation) when selenium is used, but in some cases it is converted into a methyl group (see, *e.g.*, vitamin D, §6. XI).

(v) In a number of cases, dehydrogenation is accompanied by a rearrangement of the carbon skeleton, this tending to occur at higher temperatures and when the heating is prolonged.

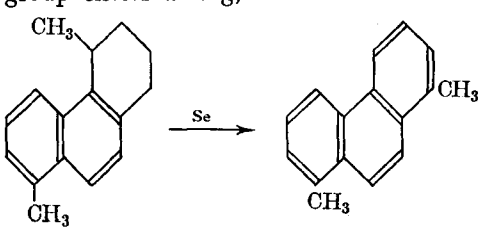
(*a*) Ring contraction may occur, *e.g.*,



*cyclo*heptane

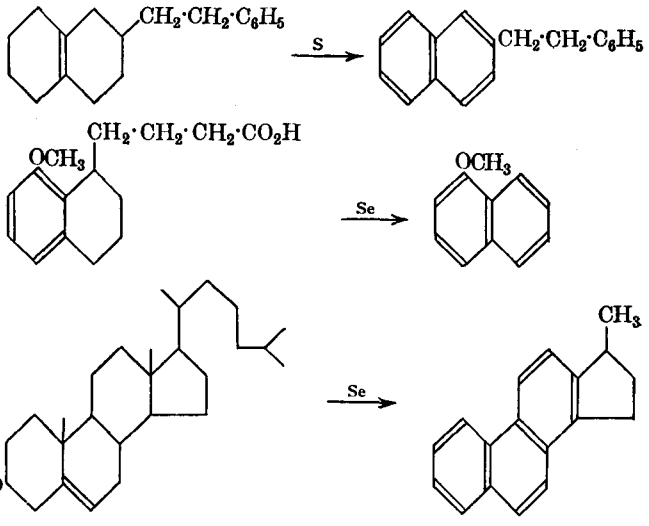
(*b*) Ring expansion may occur, *e.g.*, cholesterol gives chrysene (see §1. XI).

(*c*) Compounds containing an *angular* methyl group tend to eliminate this methyl group as CH_3SH or CH_3SeH, *e.g.*, eudesmol gives eudalene (§28b. VIII), cholesterol gives Diels' hydrocarbon (§1. XI). In some cases, the angular methyl group enters a ring,



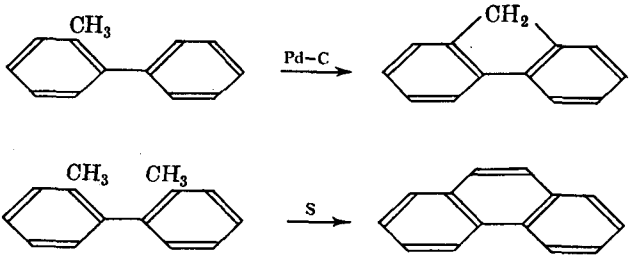
thereby bringing about ring expansion [*cf.* (*b*) above]. On the other hand, a normal substituent methyl group may migrate to another position, *e.g.*, 5 : 6 : 7 : 8-tetrahydro-1 : 5-dimethylphenanthrene gives 1 : 8-dimethylphenanthrene on dehydrogenation with selenium.

(*d*) Side-chains larger than methyl may remain intact, or be eliminated or be degraded, *e.g.*,



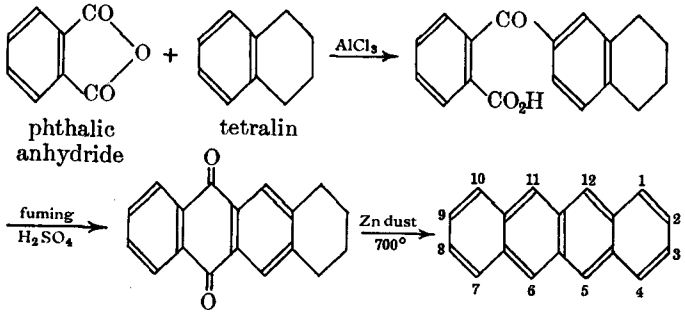
cholesterol Diels' hydrocarbon

(*e*) Dehydrogenation may produce new rings (*cf.* method iii); *e.g.*,

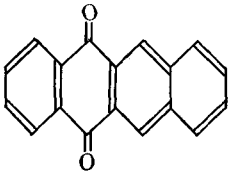


BENZANTHRACENES

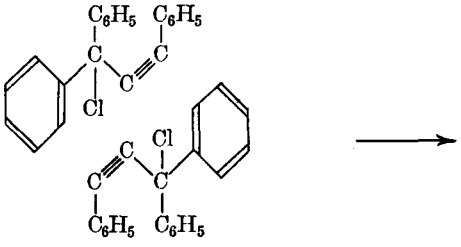
§3. Naphthacene (2 : 3-Benzanthracene), $C_{18}H_{12}$, is an orange solid, m.p. 357°. It occurs in coal-tar, and has been synthesised as follows (Fieser, 1931).

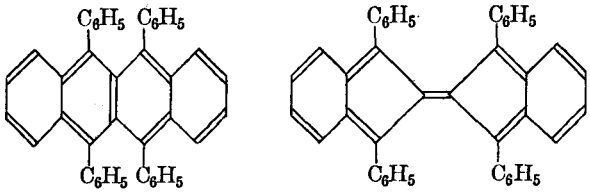


When oxidised with fuming nitric acid, naphthacene forms naphthacene-quinone.



§**3a. Rubrene** (5 : 6 : 11 : 12-tetraphenylnaphthacene) may be prepared by heating **3**-chloro-**1** : **3** : 3-triphenylprop-**1**-yne alone, or better, with quinoline at 120° *in vacuo* (Dufraisse *et al.*, 1926).

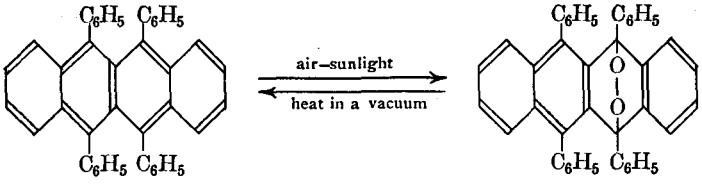




I II

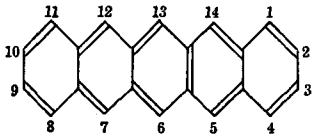
It is interesting to note that Dufraisse originally gave rubrene structure II, but changed it to I in 1935. The mechanism of the reaction is uncertain.

Rubrene is an orange-red solid, m.p. 334°. Its solution in benzene has a yellow fluorescence, but when this solution is shaken with air in sunlight, the fluorescence slowly disappears, and a white solid can now be isolated. This is rubrene peroxide, and when heated to 100–140° in a high vacuum, it emits yellow-green light and evolves oxygen, reforming rubrene.

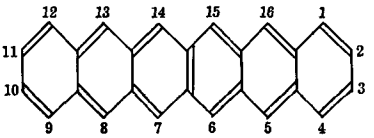


Rubrene peroxide is actually a derivative of 5 : 12-dihydronaphthacene, and so the molecule is not flat but folded about the O–O axis (the carbon atoms at 5 and 12 are tetrahedrally hybridised).

§**3b.** Two linear benzene derivatives of naphthacene have been prepared, *viz.*, **pentacene** (a deep violet-blue solid) and **hexacene** (a deep-green solid) [Clar, 1930, 1939].

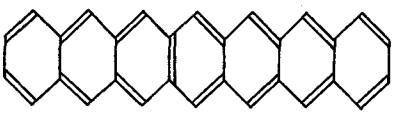


pentacene



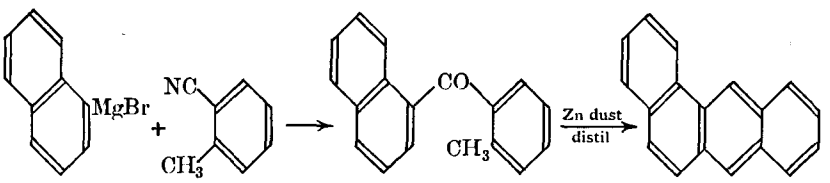
hexacene

Clar (1942) thought he had prepared heptacene, but in 1950 he showed that the compound he had isolated was 1 : 2-benzohexacene. Bailey *et al.* (1955) have synthesised heptacene.



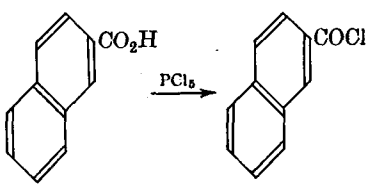
heptacene

§3c. 1 : 2-Benzanthracene, m.p. 160°, occurs in coal-tar, and has been synthesised as follows (Bachmann, 1937).

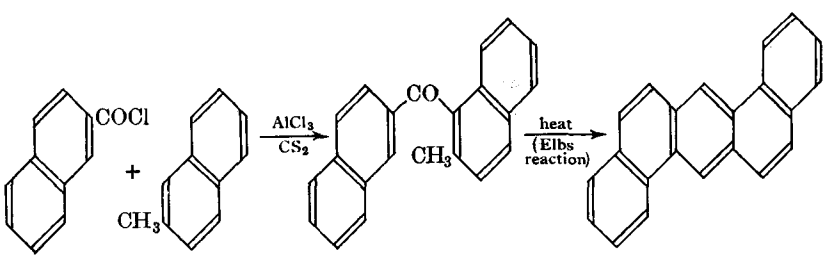


1-naphthalene-magnesium bromide

§3d. 1 : 2 : 5 : 6-Dibenzanthracene, m.p. 266°, has been synthesised by Cook *et al.* (1931), who showed that it had strong carcinogenic activity.



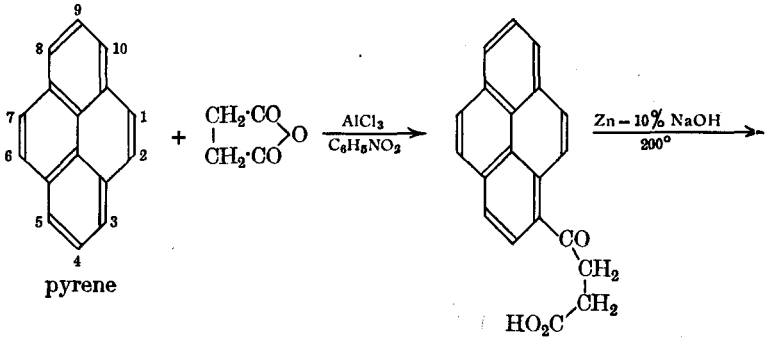
2-naphthoic acid

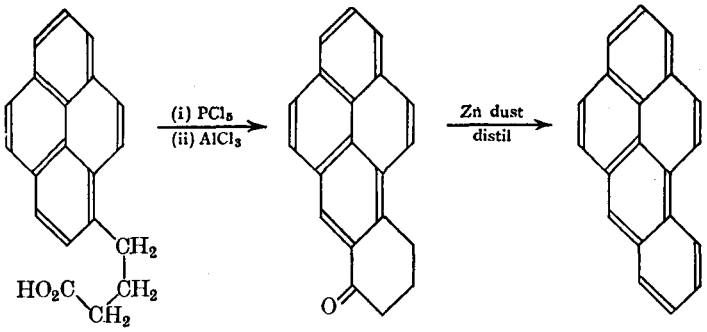


2-methyl-naphthalene

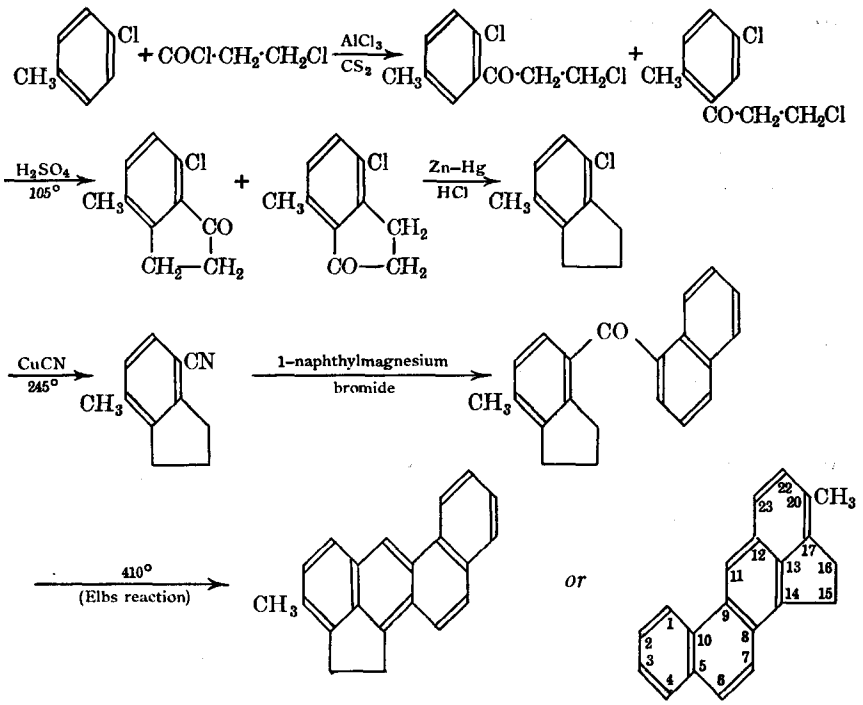
Buu-Hoï *et al.* (1960) have shown that picene (§4a) is converted into 1 : 2 : 5 : 6-dibenzanthracene by aluminium chloride in benzene.

§3e. 3 : 4-Benzpyrene is a pale yellow solid, m.p. 179°, which is very strongly carcinogenic. It occurs in coal-tar, and has been synthesised as follows from pyrene (see §4b).





§3f. 20-Methylcholanthrene is a pale yellow solid, m.p. 180°. It is a steroid derivative, and has been prepared by the degradation of, *e.g.*, cholesterol (see §3 iii. XI). Cook (1934) showed that methylcholanthrene has powerful carcinogenic properties, and Fieser *et al.* (1935) synthesised it in the following way:

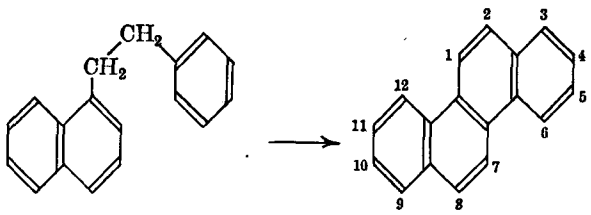


The alternative way of writing the formula shows more clearly the relationship of methylcholanthrene to the steroids (see §3. XI for the method of numbering in cholesterol). The steroids are phenanthrene derivatives, and so methylcholanthrene may also be regarded as a phenanthrene derivative (instead of an anthracene derivative).

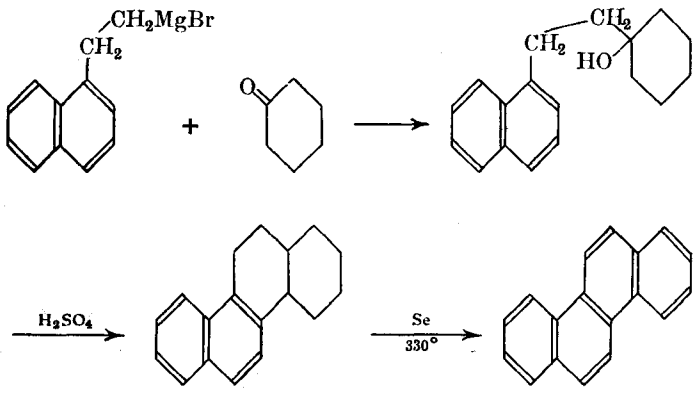
PHENANTHRENE DERIVATIVES

§4. Chrysene (1 : 2-benzphenanthrene) is a colourless solid, m.p. 251°. It occurs in coal-tar, and has been synthesised in several ways:

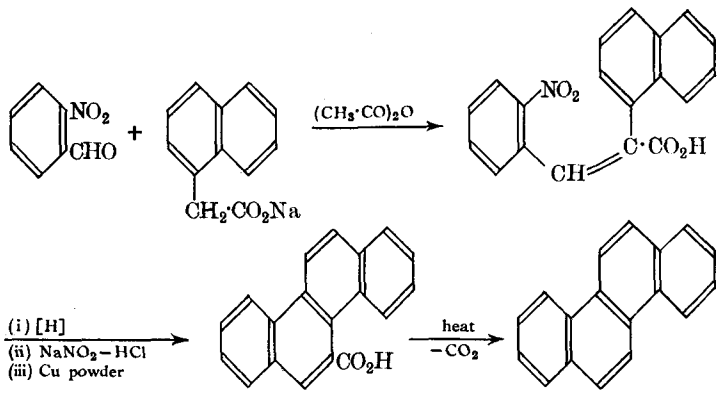
(i) By strongly heating 2-[1-naphthyl]-1-phenylethane.



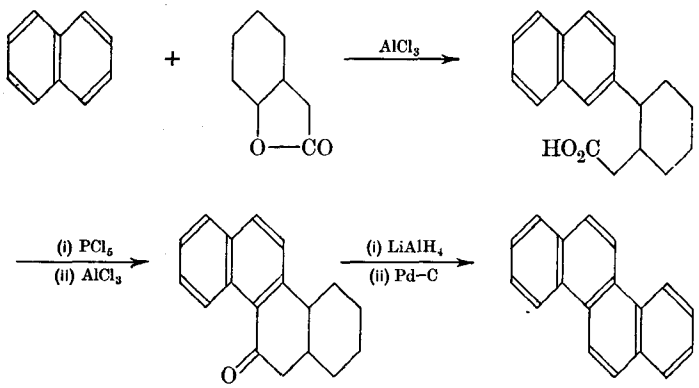
(ii) By a Bogert–Cook synthesis (*cf.* §2. (vi) *e*).



(iii) By a Pschorr synthesis [*cf.* §2 (vi) *a*].



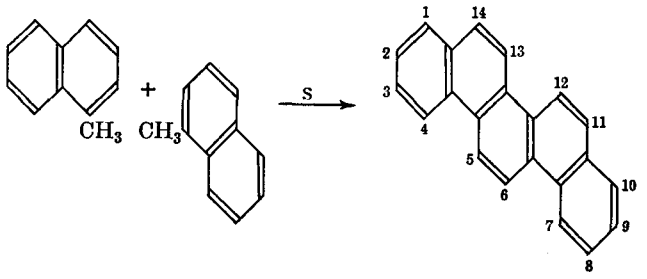
(iv) Phillips (1956) has prepared chrysene from naphthalene and the lactone of *trans* 2-hydroxy*cyclo*hexaneacetic acid:



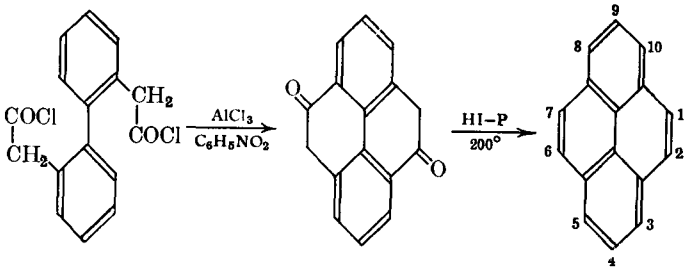
Chrysene is produced by the pyrolysis of indene, and also by the dehydrogenation of steroids with selenium.

§**4a. Picene** (1 : 2 : 7 : 8-dibenzphenanthrene), m.p. 365°, is obtained when cholesterol or cholic acid is dehydrogenated with selenium. It has been

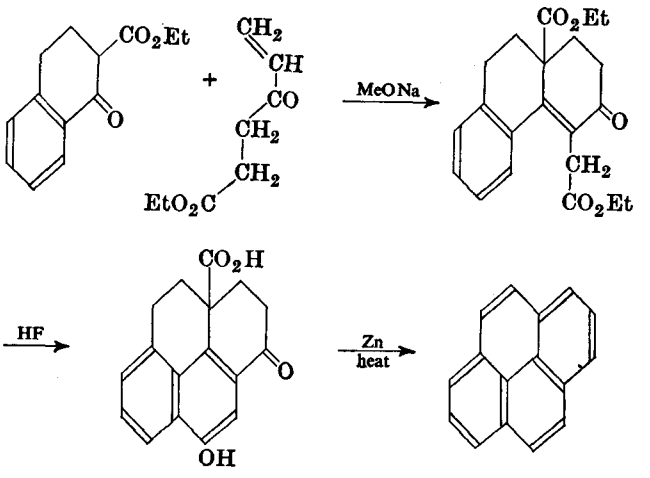
synthesised by heating 1-methylnaphthalene with sulphur at 300° (see also §3d).



§4b. Pyrene is a colourless solid, m.p. 150°. It occurs in coal-tar, and has been synthesised from diphenyl-2 : 2′-diacetyl chloride as follows:

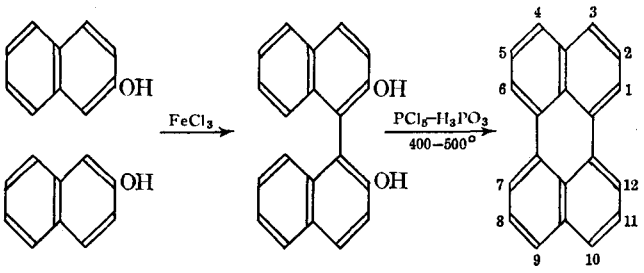


Buchta *et al.* (1958) have synthesised pyrene using an internal Stobbe reaction [§2 (vi) *c*]:

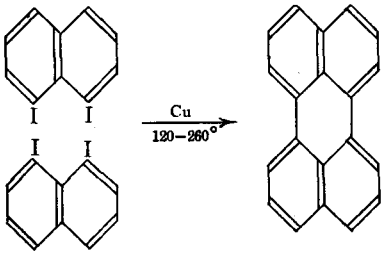


§4c. Perylene is a very pale yellow solid, m.p. 273°. It occurs in coal-tar, and has been synthesised in several ways.

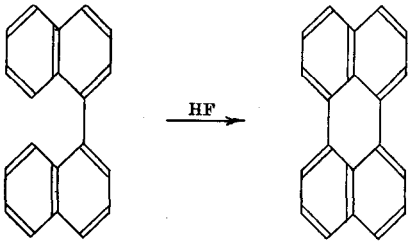
(i) 2-Naphthol, on treatment with ferric chloride solution, forms 1 : 1′-dinaphthol, and this, on heating with a mixture of phosphorus pentachloride and phosphorous acid, gives perylene.



(ii) Perylene may also be prepared by heating 1 : 8-di-iodonaphthalene with copper powder (*i.e.*, by an Ullmann synthesis; *cf.* §2. ii).

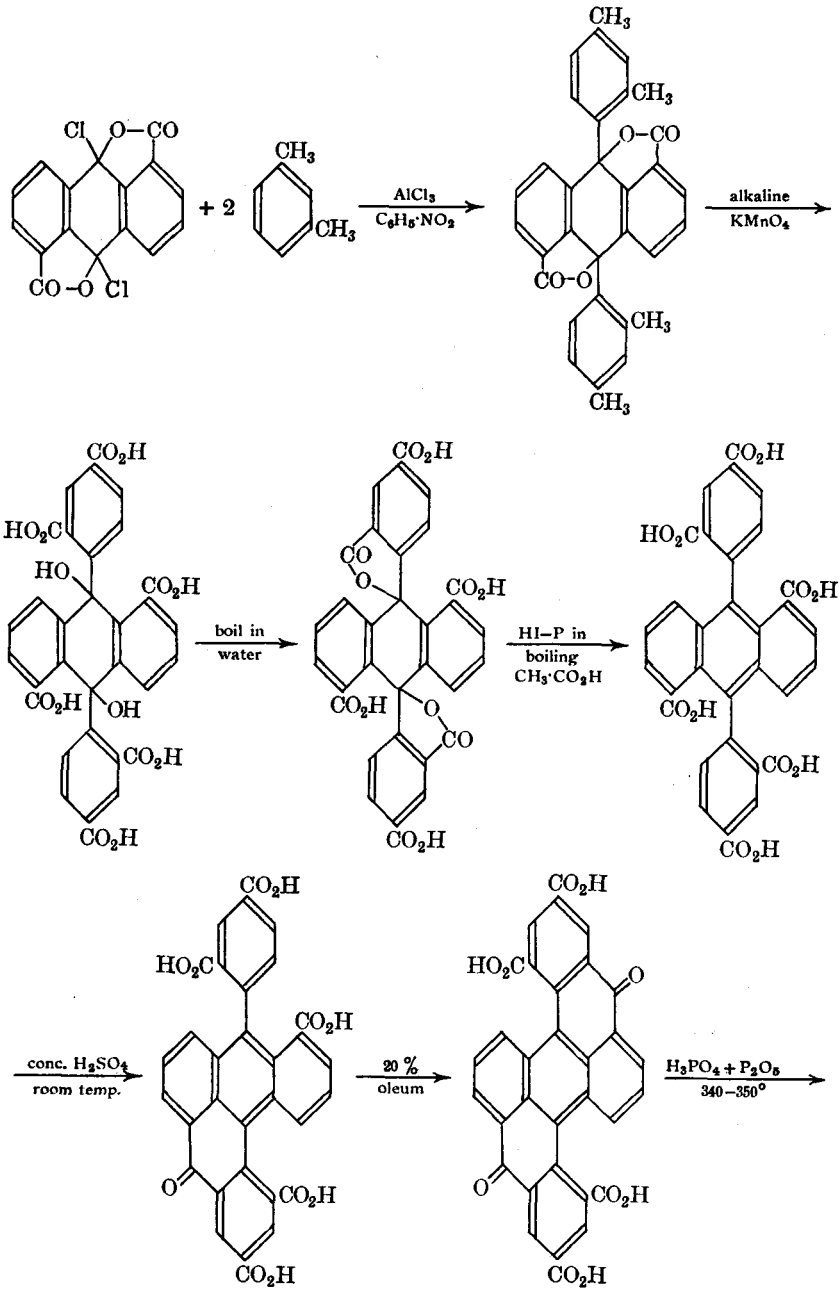


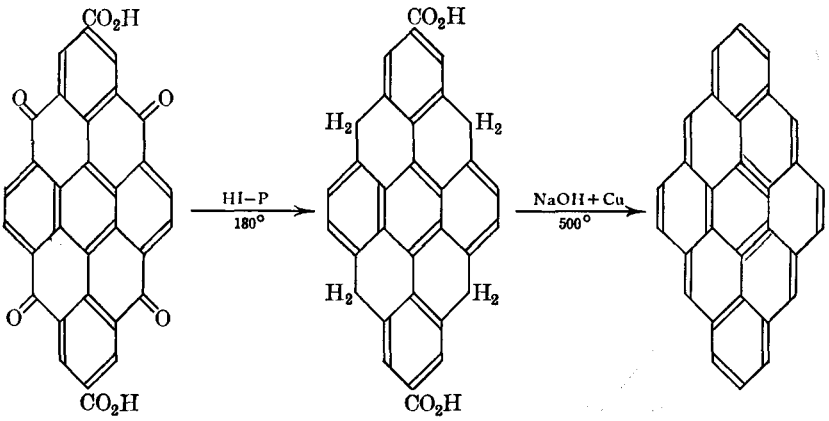
(iii) Perylene is formed when 1 : 1′-dinaphthyl is heated with hydrogen fluoride under pressure.

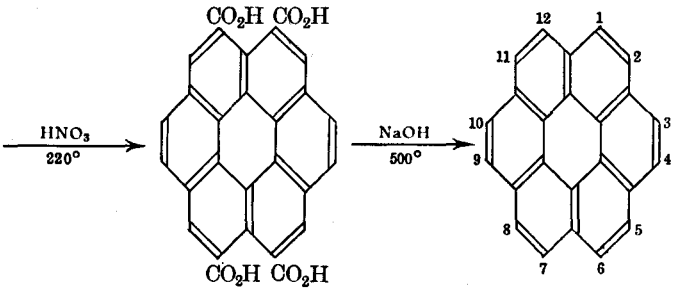


Robertson *et al.* (1953), by X-ray analysis of perylene, have shown that the two bonds connecting the two naphthalene units are longer (1·50 A) than the usual aromatic C—C bond (1·38–1·44 A). The existence of these long bonds is supported by some magnetic susceptibility measurements (Hazato, 1949).

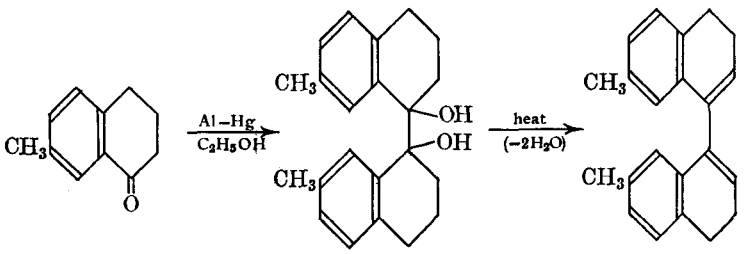
§4d. Coronene, m.p. 430°, is a yellow solid with a blue fluorescence in benzene solution; it has been found in coal-gas (Lindsay *et al.*, 1956). It was synthesised by Scholl *et al.* (1932), starting from *m*-xylene and anthraquinone-1 : 5-dicarbonyl chloride, the latter behaving in the tautomeric form shown in the following chart.

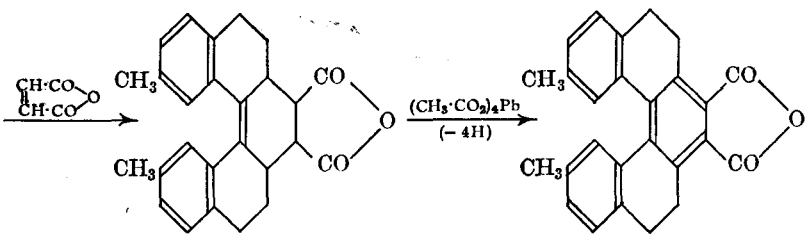

Cl O—CO
CO–O Cl
+ 2
CH3
CH3
AlCl3
C6H5·NO2
CH3
CH3
O–CO
CO–O
CH3
CH3
alkaline
KMnO4
CO2H
HO2C
HO
CO2H
CO2H
OH
CO2H
CO2H
boil in
water
CO2H
CO
O
CO2H
CO2H
O
CO
CO2H
HI–P in
boiling
CH3·CO2H
CO2H
HO2C
CO2H
CO2H
CO2H
CO2H
conc. H2SO4
room temp.
CO2H
HO2C
CO2H
O
CO2H
CO2H
20 %
oleum
CO2H
HO2C
O
O
CO2H
CO2H
H3PO4 + P2O5
340–350°

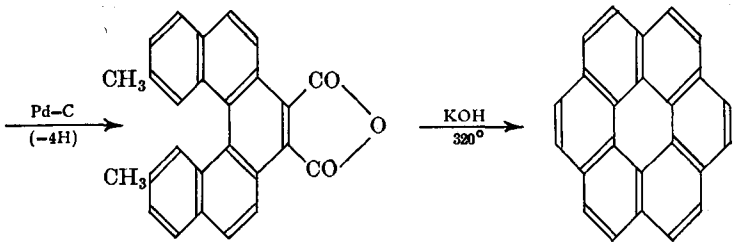




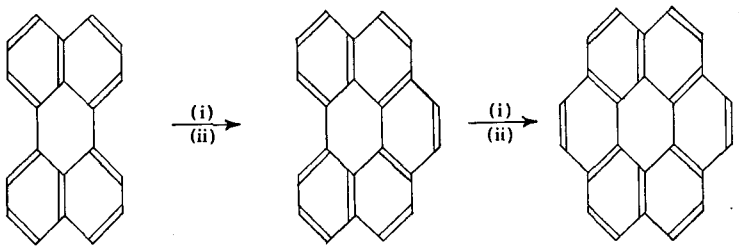
Newman (1940) has also synthesised coronene, starting from 7-methyltetralone, and proceeding as follows:







The simplest and most efficient synthesis of coronene appears to be that of Clar *et al.* (1957). The starting material is perylene (§4c), and this is treated with (i) maleic anhydride and chloranil, and followed by (ii) heating with soda-lime; these processes are then repeated:



READING REFERENCES

Newer Methods of Preparative Organic Chemistry, Interscience Publishers (1948). Dehydrogenation with Sulphur, Selenium and Platinum Metals (pp. 21–59).
Gilman (Ed.), *Advanced Organic Chemistry*, Wiley. Vol. IV (1953), pp. 1232– . Dehydrogenating Agents.
Genie, La Cyclodéshydrogénation Aromatique, *Ind. chim. belg.*, 1953, **18**, 670.
Cook, Polycyclic Aromatic Hydrocarbons, *J.C.S.*, **1950**, 1210.
Traité de Chimie Organique, Masson et Cie., Vol. XVII. Part II (1949).
Encyclopœdia of Organic Chemistry, Elsevier. Vol. 14 (1940). Tetracyclic and Higher-Cyclic Compounds. See also Vol. 14 Supplement (1951).
Cocker, Cross *et al.*, The Elimination of Non-angular Alkyl Groups in Aromatisation Reactions. Part II. *J.C.S.*, **1953**, 2355.
Cook (Ed.), *Progress in Organic Chemistry*, Butterworth. Vol. 2 (1953). Ch. 5. The Relationship of Natural Steroids to Carcinogenic Aromatic Compounds.
Badger, *The Structures and Reactions of Aromatic Compounds*, Cambridge Press (1954).