EXPERIENTIA

Vol. IX - Fasc. 10 Pag. 357–396 15. X. 1953

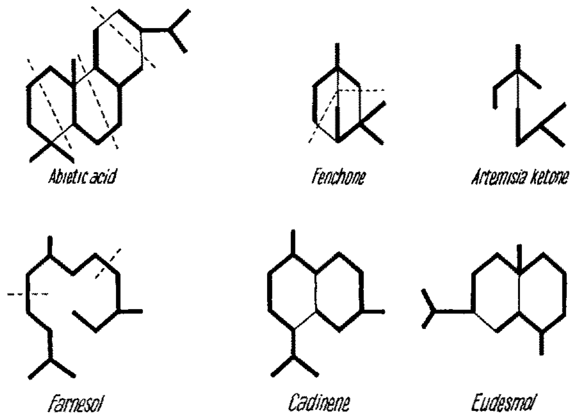
The Isoprene Rule and the Biogenesis of Terpenic Compounds[1]

By L. RUZICKA, Zurich[2]

The systematic study of the higher terpenes was started in Zurich in the early 1920's. At that time work was begun simultaneously on the diterpene abietic acid and on those sesquiterpenes which, like abietic acid, contain a hydroaromatic ring system.

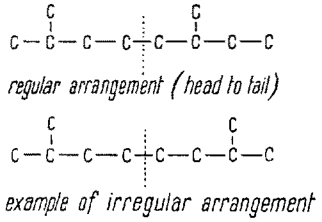
The Carbon Skeleton of the Sesqui- and Polyterpenes[3]

The importance of abietic acid (I)[4] for the development of the chemistry of the polyterpenes lies in the fact that it was soon realised that its carbon skeleton is composed of 4 isoprene units. Likewise the elucidation of the structure of hydroaromatic sesquiterpenes (I) such as cadinene and eudesmol was considerably facilitated by the observation that their carbon skeleton can formally be deduced from farnesol. Farnesol and the sesquiterpenes derived from it contain three isoprene units in regular head-to-tail arrangement (II). In abietic acid, on the other hand, the carbon skeleton is irregular, with three isoprene units arranged head-to-tail as in farnesol, while the fourth is attached the other way round (I).



I.—Regular and irregular terpenic carbon skeletons.

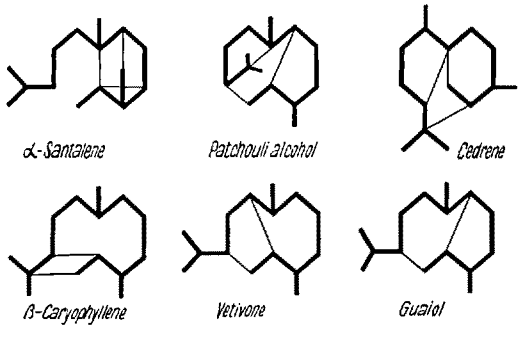
The monoterpenes are also made up of isoprene units, and this represents the common structural feature which links the various compounds of the group, even those in which the carbon skeleton is not derived from p-cymene, as for example fenchone and artemisia ketone (I).



II.—Regular and irregular arrangement of isoprene units.

This led to the formulation of the *isoprene rule*, which states that the carbon skeleton of the terpenes is composed of isoprene units linked in regular or irregular arrangement.

The isoprene rule has successfully been applied in the elucidation of the structure of terpenic compounds. A brief review of the carbon skeletons of the sesqui-, di-, and triterpenes will serve to prove the validity of the isoprene rule, and will also show that *each group has its own special variations of the fundamental rule.*



III.—Sesquiterpene carbon skeletons deduced from farnesol.

In the *sesquiterpene* group almost all compounds have regular carbon skeletons (III). Even the somewhat peculiar looking carbon skeletons of such compounds as α-santalene (SEMMLER), patchouli alcohol

---

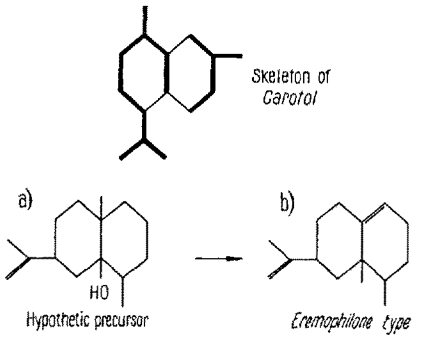
[1] From a lecture held at the XIIIth International Congress of Pure and Applied Chemistry, Stockholm, 29th July, 1953.

[2] Laboratory of Organic Chemistry, E.T.H., Zurich.

[3] In this section only the carbon skeletons are considered (Tables I-VI). The precise formulae are given in the following sections.

[4] *The formulae are grouped in 23 Tables designated by roman numbers (I-XXIII). In the text reference is made to the Table by its roman number and to the formula in the Table by a small latin letter, e.g. "IXb" indicates formula b in Table IX. When the names of compounds are indicated in the Table, only the roman number of the Table is given in the text.*

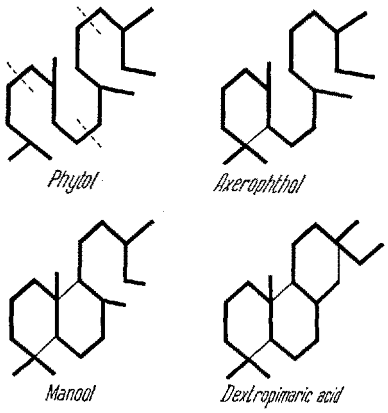
(Treibs), cedrene[1], β-caryophyllene (Šorm)[2], vetivone (Pfau and Plattner), and guaiol (Plattner) can be deduced from farnesol. The wide validity of this structural relationship led to the formulation of the "farnesol rule" which is a special case of the isoprene rule.



IV.—Irregular carbon skeletons of sesquiterpenes.

Of the sesquiterpenes, only carotol (Šorm) does not follow the farnesol rule but is instead built of 3 isoprene units in irregular arrangement (IV).

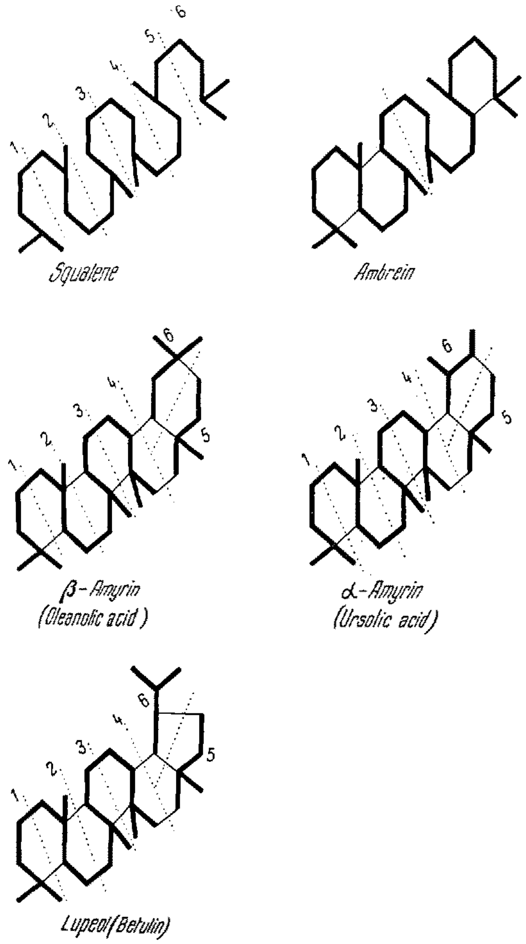
The one exception to the isoprene rule in the sesquiterpene field is eremophilone (IV) (Penfold and Simonsen). However, according to Robert Robinson, the carbon skeleton of eremophilone (IV*b*) may be derived from a eudesmol type precursor (IV*a*) by migration of a methyl group.



V.—Regular carbon skeletons of diterpenes.

In the *diterpene* group, apart from the carbon skeleton of abietic acid with an irregular sequence of four isoprene units (I), the carbon skeletons of all[3] diterpenes (V) are derived from a regular isoprene tetramer consisting of 4 isoprene units connected head-to-tail. This is best seen in four typical representatives of the diterpenes: the open chain compound phytol (F. G. Fischer), monocyclic axerophthol (Karrer), bicyclic manool (Hosking), and tricyclic dextropimaric acid. This structural relationship may be termed the "phytol rule", in analogy to the farnesol rule of the sesquiterpenes.

The majority of the diterpenes derivable from phytol, as for instance sclareol and agathic acid, possess the same bicyclic structure as manool. All bicyclic and tricyclic diterpenes have an identical carbon skeleton in rings *A* and *B*. It will be shown that this structural pattern is also present in rings *A* and *B* of the triterpenes.



VI.—Symmetrical and irregular carbon skeletons of triterpenes.

The *triterpenes* may be classified on the basis of several structural types. However, a higher analogue of farnesol and of phytol, with a regular head-to-tail arrangement of 6 isoprene units, is not known. A new type is encountered, represented by the aliphatic hydrocarbon squalene (Karrer), which is built of two farnesyl chains symmetrically joined end-to-end (VI).

The symmetrical structure of squalene is reminiscent of the C_{40} carotenoids, which may be considered, on the basis of their carbon skeleton, as typical representatives of the tetraterpenes.

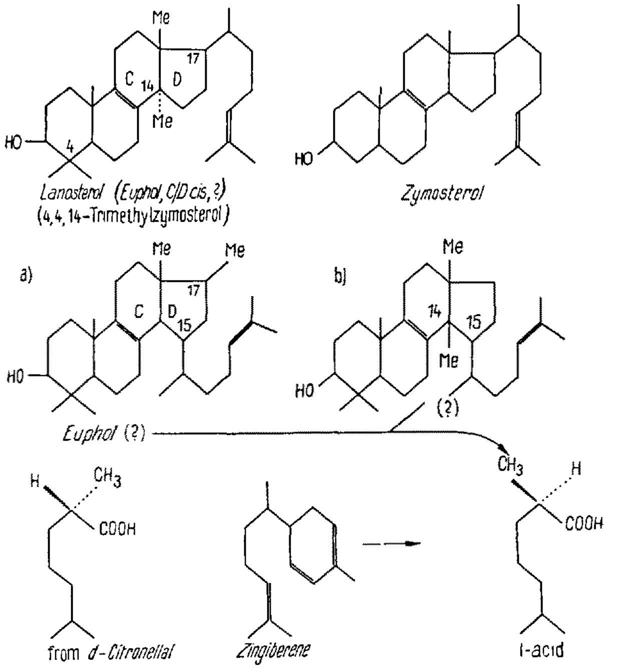
---

[1] The cedrene formula (III, *cf.* also XVIII) was proposed by A. Eschenmoser in his Ph.-D. Thesis, E.T.H., Zurich, 1952. Its correctness was demonstrated experimentally by Plattner *et al. Cf.* the paper by Pl. A. Plattner, XIIIth Internat. Congress of Pure and Applied Chemistry, Stockholm, 29th July, 1953.

[2] Šorm was the first to propose a many-membered ring structure for a terpenic compound. The elucidation of the structural details of the caryophyllenes was completed by Barton, Clemo, Dawson, and Ramage.

[3] α-Camphorene is not considered here, because, though the compound is obtained by dimerisation of myrcene, its structure is not definitely established.

Only one other triterpene belongs to the squalene type, the tricyclic alcohol ambrein (VI). Both squalene and ambrein are of animal origin. Almost all triterpenes of plant origin, on the other hand, have irregular pentacyclic carbon skeletons (VI) which partially deviate from the squalene type. They may be subdivided in three subgroups, the α-amyrin, the β-amyrin, and the lupeol subgroup. These three fundamental carbon skeletons of the pentacyclic triterpenes differ from each other only in the arrangement of one isoprene unit, which is attached at one end of the molecule (isoprene unit 6 in the formulae). The first four isoprene units (1–4 in the formulae) have the same arrangement as the corresponding units in squalene.



VII[1].—Lanosterol and euphol.

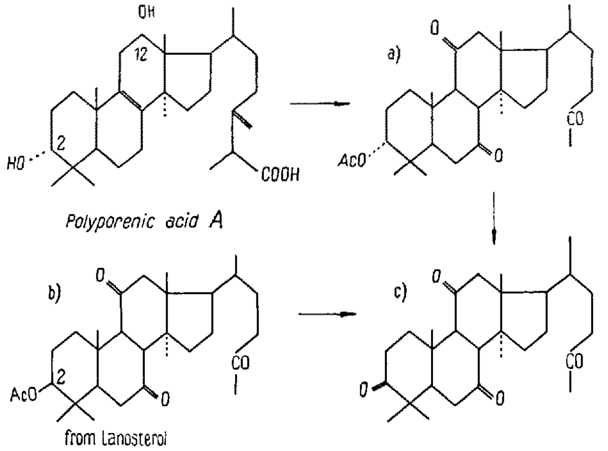
There remains to be mentioned a last group of compounds, which on the basis of their carbon skeleton may be assigned an *intermediate position between the steroids and the triterpenes.* These compounds possess the tetracyclic nucleus typical of the steroids and contain 30 carbon atoms. Two typical representatives of this group are lanosterol and euphol (Jeger).

Lanosterol (VII) was at first assumed to be a triterpene. In the course of degradative studies, however, Jeger found that the carbon skeleton is not built in accordance with the isoprene rule. Lanosterol was therefore defined as the first[2] representative of the group of the C_{30} steroids. The relationship between lanosterol and the steroids is evident at first sight, when one compares lanosterol with zymosterol. Lanosterol may in fact be considered a trimethylzymosterol.

Euphol (VII) and lanosterol differ either in the position or in the configuration of the long side chain and of a methyl group in ring *D*. Formula VII*a* was originally proposed for euphol, but a structure stereoisomeric with lanosterol, *i.e.* having a *cis C/D* junction, is not excluded. For reasons which will be discussed later, formula VII*b* should also be considered[1].

One of the latest results in the degradation of euphol should be mentioned here (VII). The 8–9 double bond of dihydroeuphol shifts easily from rings *B/C* into ring *D*. If the euphol formula differs from VII*a*, it has to be assumed that one or two methyl groups in ring *D* also migrate. In lanosterol, with a probable *trans*-junction of rings *C* and *D*, such a rearrangement does not occur. If in euphol rings *C* and *D* are *cis*-joined, methyl group migration would perhaps be explainable. The isomerisation product obtained from dihydroeuphol yields a diketone on oxidation. Stepwise degradation of this diketone leads to the dextrorotatory 2,6-dimethylheptanoic acid, which has the L configuration as the corresponding acid deducible from the sesquiterpene zingiberene (VII) and is the antipode of the acid obtained from d-citronellal, for which the *D* configuration is established.

Mention should also be made of *polyporenic acid A* (VIII), because of its close relation to lanosterol. With 31 carbon atoms, polyporenic acid *A* is something of a curiosity.



VIII.—Relation between polyporenic acid *A* and lanosterol.

The elucidation of the structure of lanosterol was of decisive assistance in the investigation of polyporenic acid *A*, and made it possible for Halsall, Jones, and Lemin to propose a partial formula. Jeger, Heusser, *et al.* have recently succeeded in establishing a *direct experimental relation between polyporenic acid A and lanosterol.* Polyporenic acid *A* can be converted in several steps to an acetoxytriketone (VIII*a*), which was supposed to be closely related to an acetoxydiketonic acid formerly obtained from lanosterol. However, the acetoxytriketone (VIII*b*) prepared from this acid differed from the degradation product obtained from polyporenic acid *A*. The close relationship between the two com-

---

[1] *Erratum:* In the formulae of the C_9 acids H and CH_3 should be interchanged and l should read d.

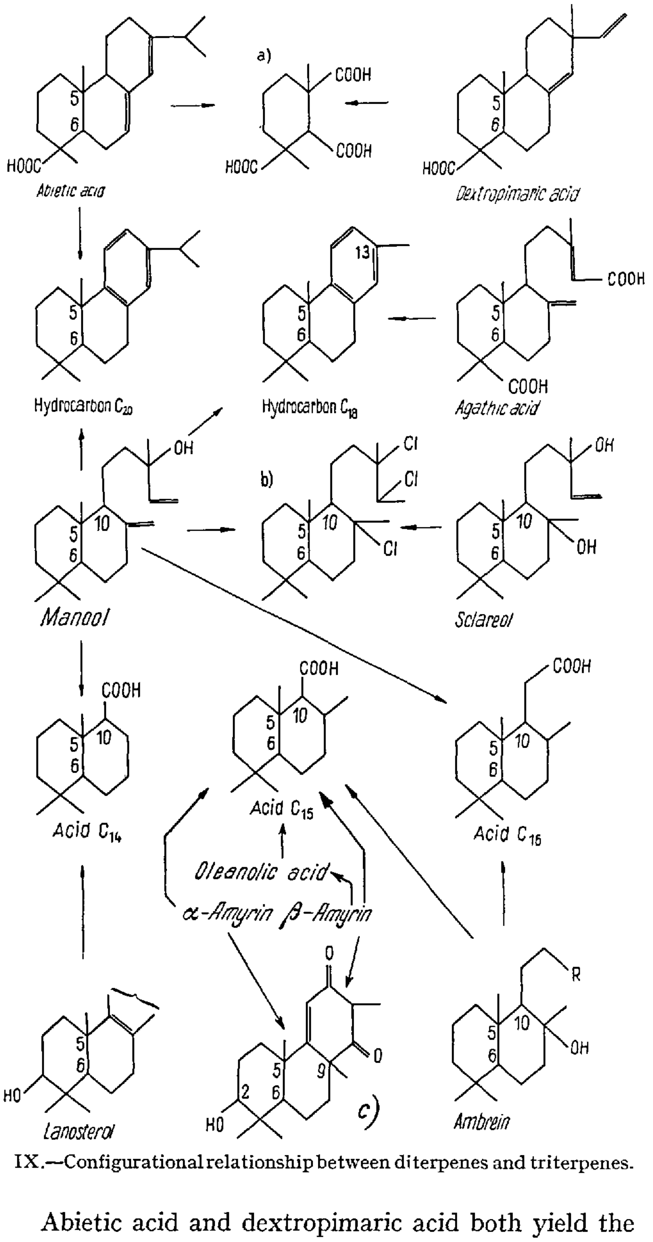
[2] The second representative is cycloartenol (Spring; Barton).

[1] The dehydrogenation of euphadiene with selenium to 1,2,8-trimethylphenanthrene supports the two formulae with the methyl groups in 13 and 14.

pounds was finally established when it was found that both yield the same tetraketone (VIII*c*). Since the hydroxyl group of lanosterol in position 3 has β configuration, it is possible to conclude from this result that the corresponding hydroxyl group of polyporenic acid *A* has α configuration. The configuration of the hydroxyl group in position 12 is according to Halsall *et al.* also α[1].

Configurational Correlations in the Diterpene and Triterpene Series

As has been mentioned before, the first two rings of the diterpenes and of the triterpenes have the same structure. It can be shown that these rings also have identical configuration.



IX.—Configurational relationship between diterpenes and triterpenes.

Abietic acid and dextropimaric acid both yield the same C_{11} tricarboxylic acid on oxidative degradation (IX*a*). Manool (Jeger) and abietic acid (Campbell and Todd) have been correlated through transformation into an identical C_{20} hydrocarbon. The relationship between manool and sclareol is established by the fact that both give an identical trihydrochloride (IX*b*) (Hosking). Finally manool and agathic acid can both be converted into the same C_{18} hydrocarbon (Jeger).

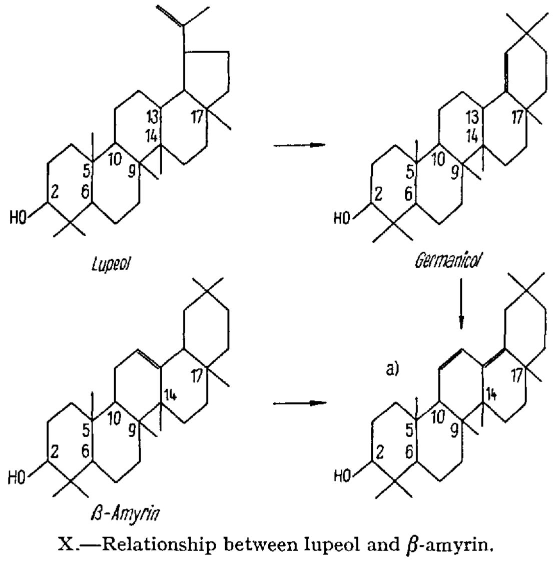
Thus it has been shown that all of these bicyclic and tricyclic diterpenes have the same configuration at the *A*/*B* ring junction. This is of course also true of the other diterpenes which have been experimentally correlated with those just discussed.

Manool has also been correlated with the tricyclic triterpene ambrein (Jeger; Lederer) by degradation of both to an identical C_{16} acid still containing the asymmetric carbons 5, 6, and 10 (IX). Starting from ambrein (Lederer), a C_{15} acid can be prepared, in which the three asymmetric carbons of the C_{16} acid still are present. This C_{15} acid is particularly important, because it has also been obtained both from α-amyrin and from β-amyrin (IX) (Jeger).

Manool, in turn, has been correlated with lanosterol. An identical C_{14} acid can be prepared by degradative reactions both from manool and from lanosterol (IX) (Jeger; Hosking), showing that the two compounds have the same configuration at carbon atoms 5 and 6.

The reactions just discussed do more than merely establish the configurational identity of carbons 5, 6, and 10 in manool, sclareol, α-amyrin, β-amyrin, and ambrein, and of carbons 5 and 6 in lanosterol, abietic acid, and dextropimaric acid. By way of manool they connect lanosterol and the amyrins with the diterpenes, and via the amyrins they embrace all the numerous pentacyclic triterpenes, belonging to the α-amyrin subgroup (*e.g.* ursolic acid) and to the β-amyrin subgroup (*e.g.* oleanolic acid).

Within the triterpene group, α-amyrin and β-amyrin have been degraded (Jeger) to an identical tricylic hydroxydiketone (IX*c*), thus proving that the two



X.—Relationship between lupeol and β-amyrin.

---

[1] T. G. Halsall and E. R. H. Jones in a paper presented at the XIIIth Intern. Congress of Pure and Applied Chemistry, Stockholm, 29th July, 1953, also proposed formula VIII for polyporenic acid *A*.

compounds have the same configuration at carbon atoms 2, 5, 6, and 9. The identity in configuration at position 10 has already been mentioned.

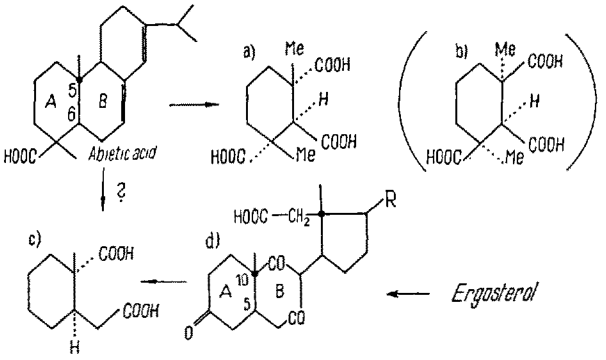
It is mostly through the work of JONES that a third subgroup of the pentacyclic triterpenes, the one of lupeol, has been correlated with β-amyrin (X). JONES obtained germanicol by rearrangement of ring *E* of lupeol. Germanicol itself has been converted (BARTON) into a doubly unsaturated triterpene alcohol (X*a*) formerly obtained from β-amyrin (JEGER).

The triterpenes of the β-amyrin subgroup contain 7 asymmetric carbon atoms in positions 2, 5, 6, 9, 10, 14, and 17, which have the same configuration as the corresponding carbon atoms in lupeol (X). Five of these seven asymmetric centers, carbon atoms 2, 5, 6, 9, and 10, have been shown to have the same configuration as in the α-amyrin subgroup (IX).

Relative and absolute Configuration of Diterpenes and Triterpenes

The first evidence regarding the relative configuration in the diterpene and triterpene series was provided for carbon atoms 5 and 6. The already mentioned C_{11} tricarboxylic acid (IX*a*), which is obtained from abietic acid and from dextropimaric acid, is optically inactive. This means that its asymmetry must be internally compensated. Two configurations (XI*a* and *b*) satisfy this condition. By studying the effect of configuration on the dissociation constants of polycarboxylic acids, BARTON was able to assign configuration XI*a*. As a consequence, *trans*-junction of rings *A* and *B* follows for all di- and triterpenes.

Another proof of the configuration at the *A*/*B* ring junction would be possible by correlation of a suitable degradation product of abietic acid with one derived from rings *A*/*B* of a steroid. Starting from ergosterol the dicarboxylic acid XI*c* has been obtained through the intermediate XI*d*. The two asymmetric carbon atoms of this acid possess the same configuration as carbons 5 and 10 in cholestanol. The preparation of the corresponding dicarboxylic acid from abietic acid is being carried out in Zürich.



XI.—Steric identity at ring A/B junction in diterpenes, triterpenes and steroids.

It is impossible to report in detail all the fundamental work which has led to an almost complete elucidation of the stereochemistry of the triterpenes. Our present day knowledge is mainly due to BARTON, CURTIS *et al.*, JEGER, JONES, KLYNE, MILLS, and PRELOG. The methods used are fundamentally the same as those applied in the steroid field and are based on the following studies:

(1) Studies of the dissociation constants of polycarboxylic acids.

(2) Studies of the relative stability of stereoisomeric 1,2-condensed ring systems, the question being in each case whether the *trans* or the *cis* configuration is more stable.

(3) Studies of the possibility of formation of bridged ring systems, in particular of lactones.

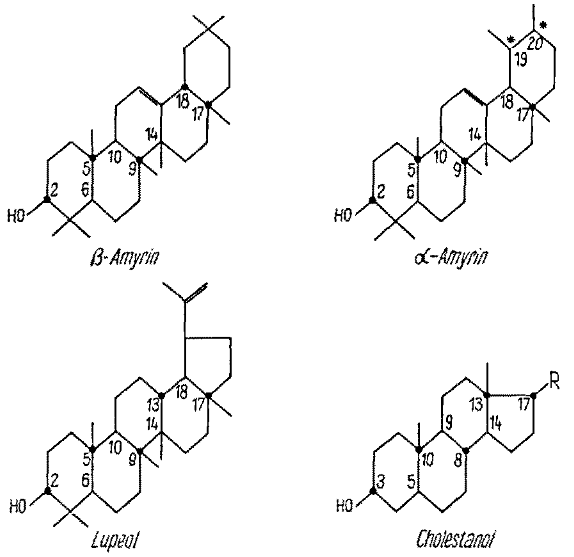
(4) Studies of the steric course of elimination and addition reactions.

(5) Studies of the contribution of individual asymmetric carbon atoms to the molecular rotation.

(6) Studies of the molecular structure by X-ray diffraction.

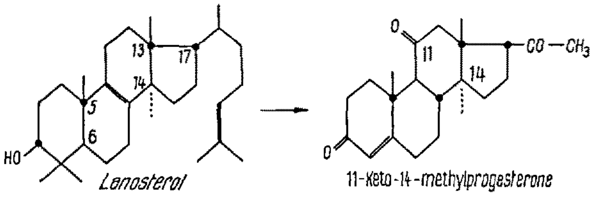
(7) Studies of the absolute configuration by PRELOG's method using asymmetric synthesis.

Of these methods, two provide information about the absolute stereochemical configuration: the method of MILLS, based on the contribution of the single asymmetric carbon atoms to the total optical activity, and the method of PRELOG, based on asymmetric synthesis. Combination of the results obtained by these two methods with those achieved by the other methods mentioned above leads to the assignement of the absolute configuration of the asymmetric carbons in the triterpenes.



XII.—Configurational relationship between triterpenes and steroids.

The resulting information regarding the absolute configuration of the triterpenes is presented in the formulae of α-amyrin, β-amyrin, and lupeol (XII). The configurational formula of cholestanol is also shown for comparison purposes. In the formulae of the three triterpenes the carbon atoms which have β configuration are marked by heavy dots, according to Linstead's notation. All other asymmetric carbons in α-amyrin, β-amyrin, and lupeol have α-configuration, as do the corresponding atoms in cholestanol. The carbons 19 and 20 in α-amyrin are of unknown configuration and are marked by asterisks in the formula. In the case of lupeol (and of β-amyranol) there are 7 asymmetric carbons (2, 5, 6, 9, 10, 13, and 14) having *the same configuration as the corresponding carbons in cholestanol*, while in the case of α- and β-amyrin there are 6 (carbons 2, 5, 6, 9, 10, and 14).



XIII.—Hormone from lanosterol.

Lanosterol has been shown to have the same absolute configuration at the *A*/*B* ring-junction as the triterpenes and cholestanol. Mention should also be made of a tentative deduction of the configuration of carbon atoms 13, 14, and 17 from physiological properties (Jeger). A compound prepared starting from lanosterol, 11-keto-14-methylprogesterone (XIII), has been found to have the same qualitative and quantitative activity in the Corner-Allen test as 11-keto-progesterone. It is a well known fact that the hormonal activity of progesterone is highly dependent on configuration, and it may be deduced from the activity of 11-keto-14-methylprogesterone that the 4 asymmetric carbons of lanosterol, which are present in 11-keto-14-methylprogesterone (XIII), have probably the same configuration as the corresponding 4 atoms in progesterone, and consequently in cholesterol.

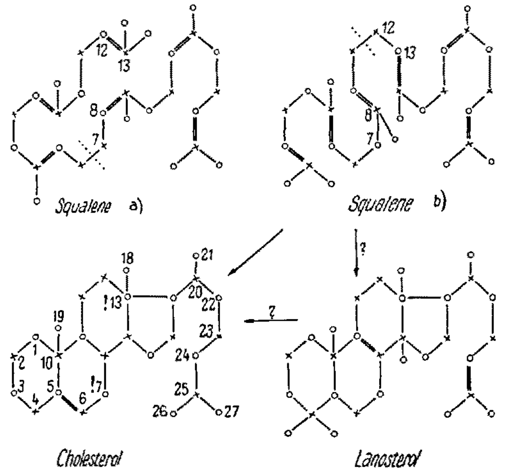
The same conclusion has been reached by Curtis *et al.* by X-ray diffraction study of lanosterol. Curtis' results lead him to assign to the lanosterol molecule a flat structure, analogous to the one established by Bernal 20 years ago for cholesterol.

Biogenesis of Steroids and Terpenic Compounds

(together with A. Eschenmoser[1] and H. Heusser)

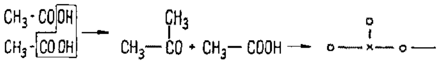
Once the structure and the configuration of a natural compound has been established, the question arises how the compound is synthesized in nature.

The first definite answer to such a question in regard to the carbon skeleton of the steroids and the triterpenes was given by Bloch through a biological synthesis of cholesterol from acetic acid labelled with C^{14}. Bloch has also provided evidence indicating that in all probability squalene is an intermediate in this synthesis. The two ways by which squalene may be transformed into cholesterol, are shown in schemes XIV*a* and XIV*b*. Scheme *a* represents the cyclisation of squalene as proposed by Robinson. Scheme *b* is the mechanism proposed by Woodward and Bloch.



XIV.—Biogenesis of cholesterol and lanosterol from squalene.
° = Carbon from acetate methyl,
× = Carbon from acetate carboxyl.

In both schemes the carbon atoms originating from the methyl group of acetic acid are indicated by a small circle, and those originating from the carboxyl group of acetic acid by a cross, in accordance with the arrangement of the isoprene units in squalene (VI) and with the formation of the isoprene skeleton from acetic acid (XV).



$$\begin{matrix}CH_3\cdot CO\boxed{OH}\\ CH_3\cdot\boxed{CO\,OH}\end{matrix} \longrightarrow CH_3-\overset{CH_3}{\overset{|}{C}}O + CH_3-COOH \longrightarrow \overset{\circ}{\overset{|}{\circ}}-\times-\circ-\times$$

XV.—Distribution of acetate carbons in the biogenesis of the isoprene skeleton.

Comparison of the two schemes XIV*a* and *b* shows that the cholesterol carbons at positions 7, 8, 12, and 13 are derived from different acetate carbons, depending on whether scheme *a* or scheme *b* is followed. An experimental decision between these two cyclisation schemes was reached by appropriate degradation of the biosynthetically labelled cholesterol. By degradation of cholesterol obtained from methyl labelled acetate, Bloch has been able to prove that carbon atoms 7[1] and 13 are radioactive[2]. It follows that scheme *b* represents the cyclisation mechanism, if squalene really is the intermediate in the biosynthesis of cholesterol.

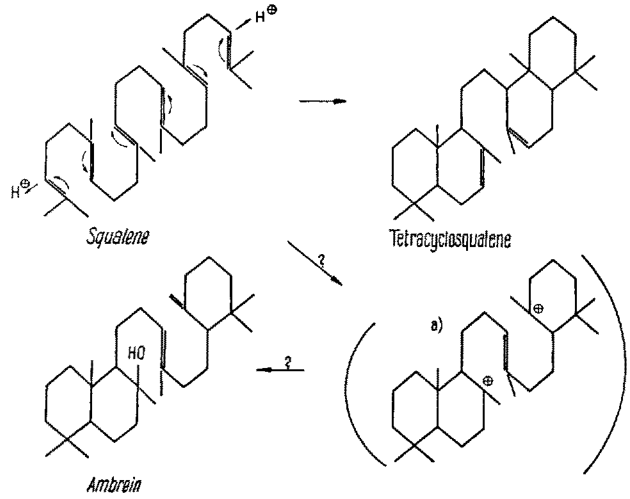
It had been shown previously that carbons 20–27 of the side chain (Würsch, Huang, and Bloch) and 1–6,

---

[1] *Cf.* A. Eschenmoser, Ph.-D. Thesis, E.T.H., Zurich, 1952, p. 16–38.

[1] Private communication.

[2] A. Mondon, Angew. Chemie *65*, 333 (1953), has recently published a paper in which he proposes a hypothetical iso-squalene as the biological precursor of cholesterol. In this hypothesis the rearrangement required in the deduction of the cholesterol formula from squalene is avoided. However, Mondon's proposal is in disagreement with Bloch's experimental results, as it does not account for the presence of a carbon atom derived from an acetate methyl group in position 13 of cholesterol.



XVI.—Acid-catalysed cyclisation of squalene.

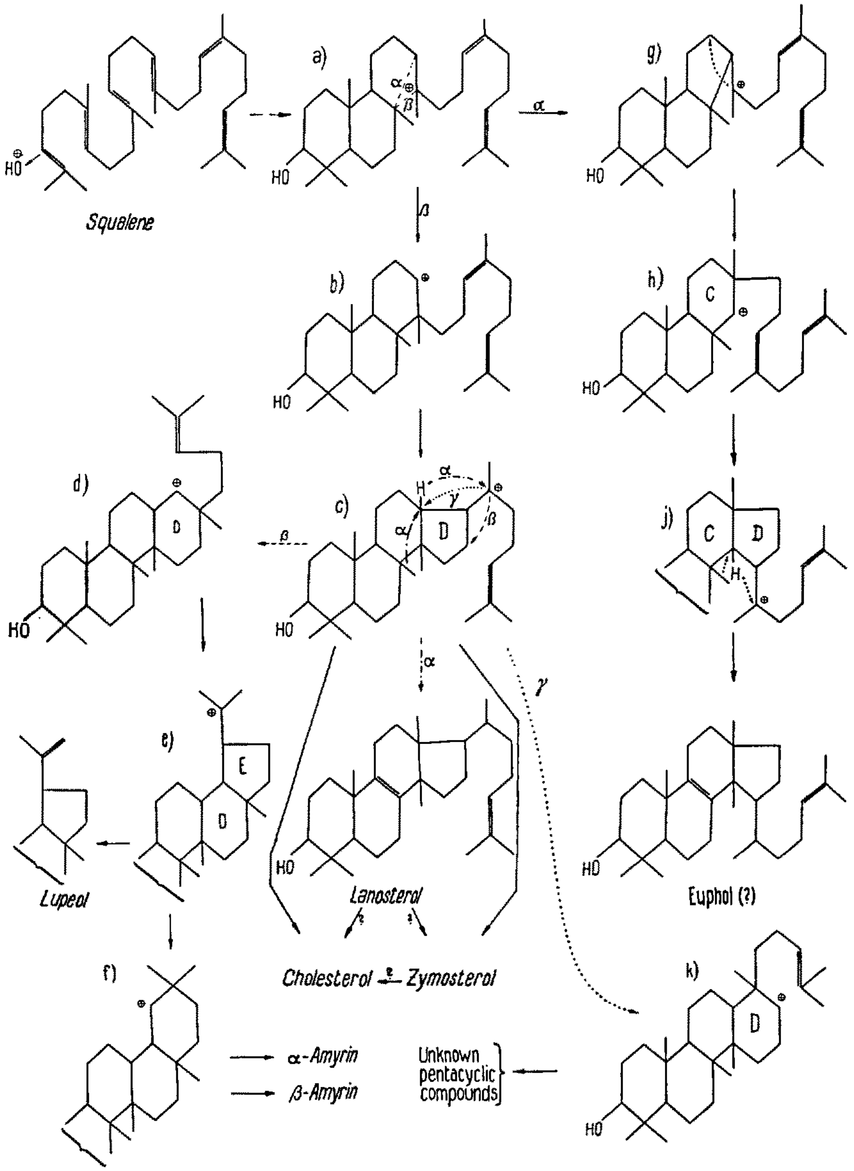
10, and 19 of rings *A*/*B* of cholesterol (Hunter, Popjak, and Cornforth) are arranged according to the alternation principle required by both cyclisation schemes (XIV).

Though squalene most probably is the biological precursor of cholesterol, it should be emphasised that the only product obtained so far by acid-catalysed cyclisation of squalene is tetracyclosqualene[1] (XVI). It can be assumed that also ambrein is formed by cyclisation of squalene. In the formation of tetracyclosqualene, as well as in that of ambrein, cyclisation starts simultaneously at both ends of the squalene molecule (XVI).

An essential characteristic of the biological cyclisation of squalene to the steroids is the attack of a cation (e.g. OH^+) at one end only of squalene (XVII)[2], whereupon the cyclisation proceeds synchronously to completion[3]. In the case of *lanosterol*, *zymosterol*, and *cholesterol*, the cyclisation yields (by way of the intermediates XVII *a* and *b*) the cation XVII *c*, which is stabilised (reaction α in XVII *c*) through a 1–3 hydrogen shift[1] (from position 13 to position 20) and a double 1–2 methyl shift[2] (positions 8, 14, and 13).

The cation XVII *c* could also be stabilised by Wagner-Meerwein rearrangement, and this by two different routes (routes β and γ in XVII *c*) each leading to different pentacylic systems. Both rearrangements result in the enlargement of ring *D* to a 6-membered ring (XVII *d* and *k*).

The cation XVII *d* leads to the formation of a 5-membered *E*-ring (XVII *e*), and hence to *lupeol*. A further Wagner-Meerwein rearrangement of XVII *e* leads by way of the intermediate XVII *f* to *α-amyrin* and *β-amyrin*. This route from lupeol to β-amyrin via the intermediates XVII *e* and XVII *f* is nothing but a formulation of the well-known transformation of lupeol into compounds of the β-amyrin type, which, as we have seen, has been experimentally achieved (X).



XVII.—Ionic mechanisms in the biogenesis of steroids and triterpenes from squalene.

---

[1] The position of the two double bonds in tetracyclosqualene, which has been proved by G. Büchi (private communication), explains why further cyclisation of tetracyclosqualene has so far been impossible.

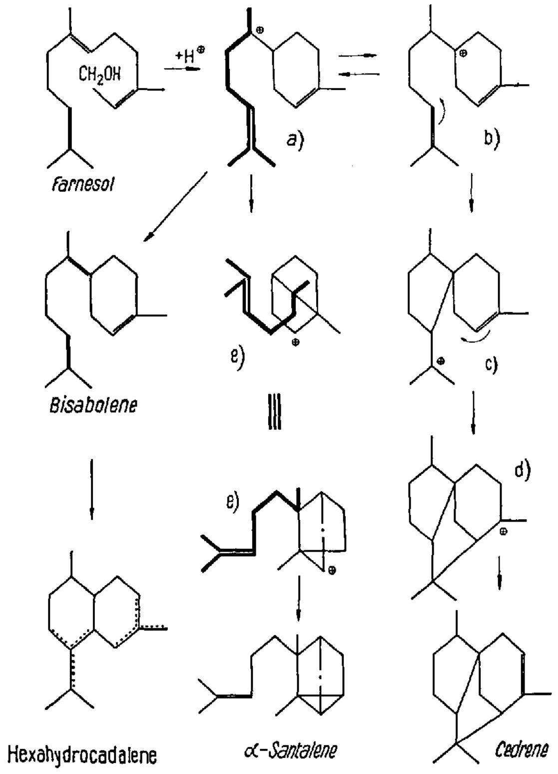
[2] In Table XVII, for the sake of simplicity, the configuration of the methyl groups is not indicated.

[3] It is not known at what stage the methyl groups in positions 4 and 14 may be eliminated.

[1] Or a double 1—2 shift.

[2] Or a 1—3 shift.

The cation XVII*k*, on the other hand, represents the precursor of pentacyclic compounds of structural types, which so far have not been observed in nature. It is worth noting that ring *E* in these hypothetical compounds would have structures similar to those of ring *E* in lupeol, α-amyrin, and β-amyrin.



XVIII.—Ionic mechanisms in the biogenesis of sesquiterpenes (6-membered ring intermediate).

A last variation of the squalene cyclisation should be mentioned. The limiting structure XVII*g* of the π-complex XVII*a* could undergo a Wagner-Meerwein rearrangement (from 14 to 12) to an intermediate XVII*h*, with subsequent formation of the five-membered *D*-ring (XVII*j*). A 1–3 shift of a hydrogen atom (from 14 to 20) and a 1–2 shift of a methyl group (from 8 to 14) would lead to one of the formulae proposed for *euphol* (VII*b*). On the basis of this mechanism, formula VII*b* appears to be more likely than the one with the methyl group in 17 (VII*a*) and should therefore also be taken into consideration.

In an E.T.H. thesis[1] published in 1925, one may find the following statements:

"The *hypothesis* may be formulated that the *steroids* and the *triterpenes have, at least partially, a common origin.* For instance, one only has to insert a methyl group in each of the rings *A*, *B*, and *C* of cholesterol, in order to obtain a carbon skeleton corresponding to a triterpene. It would not be very profitable to discuss *the question whether steroids may originate from such triterpenes by elimination of three methyl groups.*[1]"

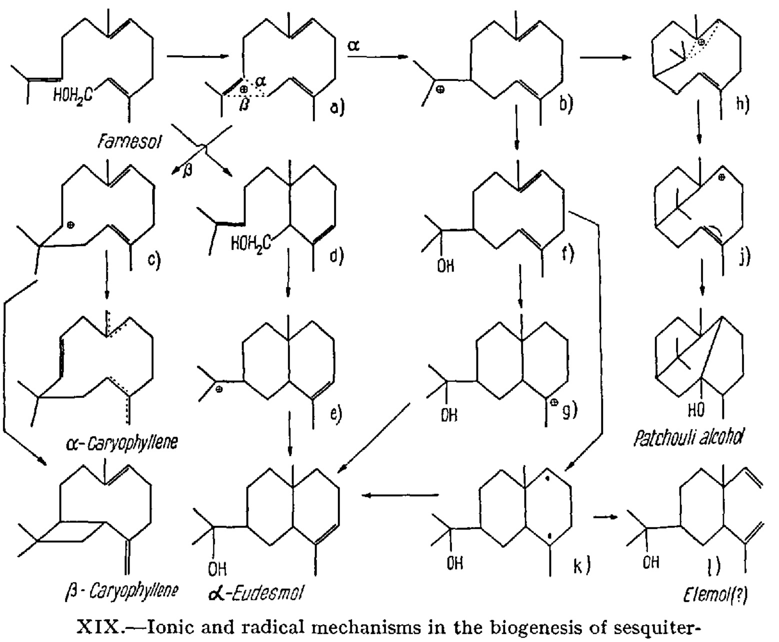
28 years later, in the light of Bloch's experimental results and of the constitutional and configurational relations of steroids and triterpenes, the common origin of steroids and triterpenes is not only a matter of profitable discussion, but achieves the status of a working hypothesis for future chemical and biochemical work.

It should be emphasized that all of the biogenetic schemes discussed above are based on generally accepted reaction mechanisms. *Analogous biogenetic schemes*[2] *can also be proposed for the biosynthesis of sesquiterpenes, diterpenes, and monoterpenes.*

The biogenesis of *sesquiterpenes* possessing a carbon skeleton derived from farnesol may be assumed to follow a route starting from farnesol, farnesene or nerolidol.

One may distinguish two types of biogenetic cyclisation of farnesol. One is characterised by the formation of 6-membered ring intermediates and the other of 10- or 11-membered ring intermediates.

The *acid-catalysed cyclisation* of farnesol or nerolidol, which was carried out experimentally in Zurich in



XIX.—Ionic and radical mechanisms in the biogenesis of sesquiterpenes (10- and 11-membered ring intermediates).

---

[1] E. A. Rudolph, Ph.-D. Thesis, E.T.H., Zürich, 1925, p. 98, 99.

[1] The cholesterol formula which had been proposed by Wieland and Windaus when this Thesis was written was far from being established and no triterpene formula at all had as yet been published. *The above biogenetic hypothesis was advanced on principle, because it appeared more likely than the one deriving the steroids from fatty acids* (Wieland).

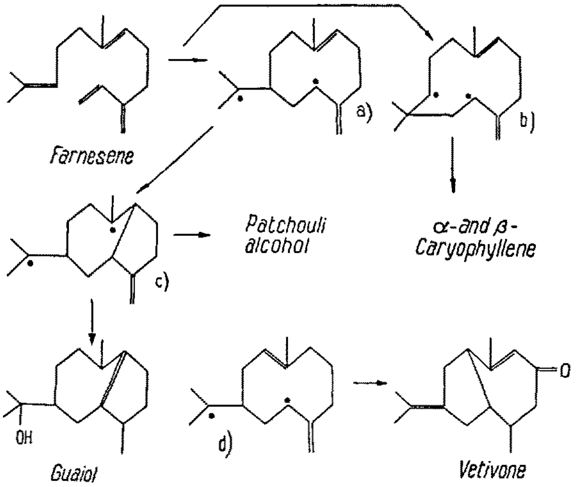
[2] For the sake of greater clarity, the hypothetical intermediates in the acid-catalysed cyclisations are formulated as classical carbenium ions.

1925, is an example of the first type. By way of the intermediate *6-membered ring* cation XVIII*a*, the reaction leads to *bisabolene* and further to a mixture of hexahydrocadalenes. The intermediate XVIII*a* could undergo a different cyclisation leading to the cation XVIII*e*, in analogy to the schematic transition from XXII*a* to the bornyl cation (XXII*b*). The cation XVIII*e* would be stabilised to α- or to *β-santalene.* The cation XVIII*b*, also derived by cyclisation of farnesol, leads to *cedrene* by way of the intermediates *c* and *d*.

The second type of farnesol cyclisation starts from the π-complex XIX*a* of the farnesyl cation, and may lead to a cation with a *10-membered ring* (*b*) or to a cation with an *11-membered ring* (*c*). The intermediate *c* yields *α-caryophyllene* (humulene) directly, or else, on formation of a 4-membered ring, *β-* or *γ-caryophyllene.*

The 10-membered ring intermediate XIX*b*, which is derived from the π-complex of the farnesyl cation, leads to *patchouli alcohol* by way of the π-complex XIX*h* and the bicyclic intermediate XIX*j*.

The biogenesis of *eudesmol* may be formulated along two different ionic routes. The one leads to α- or β-eudesmol by way of the already mentioned 10-membered ring intermediate XIX*b* and the intermediates *f* and *g*. The other route starts with a cyclogeraniol type cyclisation of farnesol to an intermediate XIX*d*, followed by an α-terpineol type cyclisation (XIX*e*) to α-eudesmol.



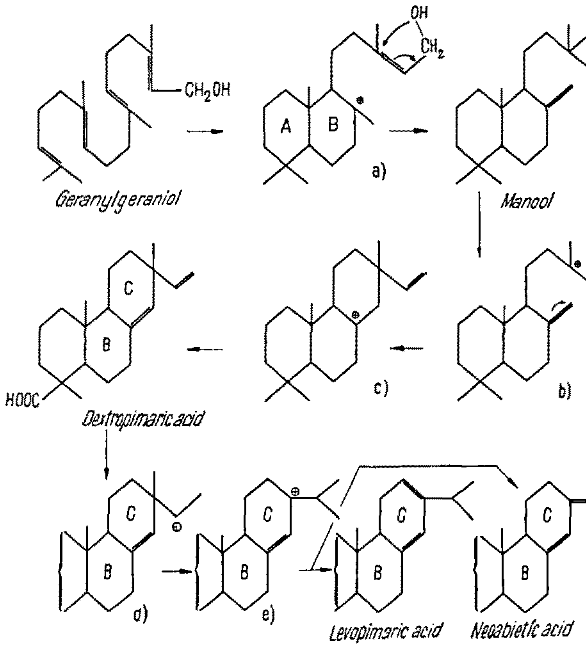
XX.—Radical mechanisms in the biogenesis of sesquiterpenes (10- and 11-membered ring intermediates).

So far only ionic cyclisations have been discussed. *Radical mechanisms* may, however, also be involved (XX). In fact, in certain cases the radical mechanism may be more appropriate than the ionic one, as for instance in the cyclisation of farnesene to guaiol, to vetivone, and also to the caryophyllenes. In the case of guaiol the intermediate would be the monocyclic biradical XX*a* which can cyclise to the bicyclic biradical *c* of the guaiol type. From here stabilisation may take place either directly to *guaiol* or by further cyclisation to patchouli alcohol (formula in Table XIX). The intermediate on the way from farnesene to the caryophyllenes (formulae in Table XIX) is the biradical XX*b*. A shift of the ring double bond of the biradical XX*a* would lead to a new 10-membered ring biradical (XX*d*), which may be stabilised by cyclisation, yielding *vetivone.*

Formula XIX*l*, recently advanced for *elemol* by Šorm[1], on the basis of the isolation of methyldiisopropylbenzene on dehydrogenation, may be derived from the intermediate XIX*f* by cyclisation to the biradical XIX*k*, which could then undergo ring opening. This ring opening to XIX*l* corresponds closely to the one that occurs in the experimental conversion of α-pinene into ocimene (or *allo*-ocimene) and of β-pinene into myrcene by way of the biradicals XXIII*a* and XXIII*b*.

Another stabilisation route of the biradical XIX*k* would lead to α-eudesmol. This stabilisation is analogous to the conversion of the biradicals XXIII*a* and XXIII*b* into limonene, which has actually been obtained by pyrolysis of α- and β-pinene.

Farnesol thus occupies a key position not only because it possesses the carbon skeleton traceable in sesquiterpenes, but also because it is a possible biological precursor of the cyclic sesquiterpenes and, in the form of difarnesyl (*i.e.* squalene), also of the triterpenes and of the steroids.

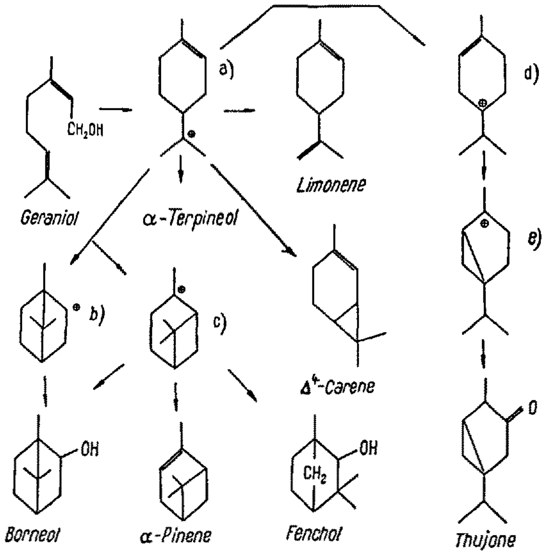


XXI.—Ionic mechanisms in the biogenesis of diterpenes.

---

[1] F. Šorm in a paper presented at the XIIIth Internat. Congress of Pure and Applied Chemistry, Stockholm, 29th July, 1953.

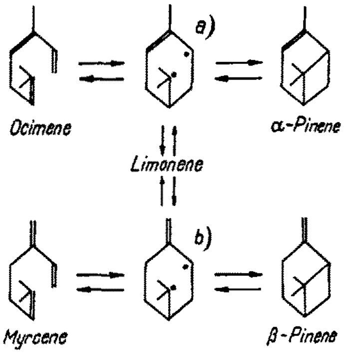
The *diterpenes* can easily be derived from geranylgeraniol[1], which so far has not been found in nature. Double cyclisation at one end of the molecule, accompanied by an allylic rearrangement at the other end (XXI*a*), would lead to *sclareol* and *manool*. Manool could in turn yield *dextropimaric acid* by cyclisation of the intermediate XXI*b* to XXI*c*. From dextropimaric acid, finally, one may derive *neoabietic acid* and *levopimaric acid*, by way of a WAGNER-MEERWEIN rearrangement of the intermediate XXI*d* to XXI*e*. As is well known, neoabietic acid and levopimaric acid readily yield *abietic acid*.



XXII.—Ionic mechanisms in the biogenesis of monoterpenes.

The cyclic *monoterpenes* can also be derived from an open chain precursor such as geraniol[2] by way of the common intermediate XXII*a*. In Table XXII the well-known formation of *limonene* from geraniol is formulated. The intermediate *a* can cyclise to XXII*b* or *c* and lead either to α- and *β-pinene*, or to *borneol*. From pinene, by way of XXII*c* bornyl and *fenchyl* derivatives can be obtained. By an intramolecular electrophylic substitution the intermediate XXII*a* can also lead to a 3-membered ring and consequently to *Δ⁴-carene*. *Thujone* and *sabinene* may be derived from the cation XXII*d*, an intermediate in the formation of α- and γ-terpinene from limonene, by way of the bicyclic cation *e*.

For the biogenesis of the monoterpenes it is also possible to propose radical mechanisms, as for example for the formation of limonene and α-pinene from ocimene, and of limonene and β-pinene from myrcene, by way of the intermediate biradicals XXIII*a* and *b*. The reactions in Table XXIII are nothing but the reversal of the experimentally achieved thermal ring openings of α-pinene to yield limonene and *allo*-ocimene (which is formed by isomerisation of the primary product ocimene), and of β-pinene to yield limonene and myrcene.



XXIII.—Radical mechanisms in the biogenesis of monoterpenes.

Isoprene rule

Only few of the cyclisations mentioned above have been carried out in the laboratory. These reactions can follow various routes, each leading to different natural compounds and, so far, most of the routes are, on steric grounds, not accessible *in vitro*. The course of a cyclisation is not only dependent on the reaction mechanism, but also to a high degree on the *constellation*[1] of the precursor and of the intermediate. It is an important function of the enzymes to bring about the required constellation and thus cause the cyclisation to follow one specific course.

An interesting example of the importance of the constellation in the biogenesis of terpenic compounds is given in the sesquiterpene group, where a great variety of carbon skeletons is encountered. The one carbon skeleton which has never been observed, is the bicyclofarnesol skeleton, so typical of all cyclic diterpenes and triterpenes. This appears to indicate that the biogenesis of steroids, diterpenes, and triterpenes differs in some fundamental detail from that of monoterpenes and sesquiterpenes.

Whether the hypothetical schemes which have been discussed have a real significance in the laboratory of nature, remains of course to be proved by experiment. BLOCH's merit is to have provided a firmer basis for such an experimental approach to the biogenesis of terpenic compounds as well as of the steroids.

Moreover, the *deduction of structures of natural terpenic compounds by accepted reaction mechanisms from the hypothesized simple precursors squalene, geranylgeraniol, farnesol, geraniol ("biogenetic isoprene rule")* not only serves to outline possible biogenetic routes,

---

[1] Or from analogous compounds which can yield suitable ionic or radical intermediates, *e.g.* geranyllinalool, geranylmyrcene.

[2] Or from analogous compounds, *e.g.* linalool, myrcene.

[1] The word "*constellation*" (F. EBEL in K. FREUDENBERG, Stereochemie [F. Deuticke, Leipzig and Wien, 1933], p. 825) has the same meaning as "*conformation*" (*cf.* D. H. R. BARTON, Soc. *1953*, 1027).

but also represents a new helpful tool in the structural elucidation of terpenic compounds. This new tool limits the number of carbon skeletons which can be proposed on the basis of the original isoprene rule[1] ("*empirical isoprene rule*") alone. It is obvious that any formula deduced in accordance with the "biogenetic" as well as with the "empirical" isoprene rule requires experimental proof.

The "empirical isoprene rule" was deduced formally from the structure of the natural terpenic compounds. From the "biogenetic isoprene rule", however, it would follow that the carbon skeleton of the biological end product is not necessarily identical to the carbon skeleton of the precursor. In other words the validity of the "empirical isoprene rule" depends on the mechanisms of formation of the natural compounds, and the failure of a terpene to obey this rule does not necessarily disprove its origin from isoprene units.

Nomenclature

The terms terpene, diterpene, triterpene are retained here in preference to the terms terpenoid, diterpenoid, triterpenoid, which are occasionally used in the modern literature. These terms have unnecessarily been introduced in analogy to the expression "steroid", which denotes a group of compounds having an irregularly varying number of carbon atoms. In the terpenic field the expression terpenoid should be reserved by analogy for compounds in which the number of carbon atoms varies irregularly[2], in contrast to the terpenes proper, where the number of carbon atoms is always a multiple of five.

The terms monoterpene, sesquiterpene, diterpene, triterpene are unambiguous and need not be replaced by nebulous synonyms ending in "oid". The term terpene, on the other hand, originally designated the monoterpenes alone. In order to avoid confusion the term terpene should be used only to designate the whole class of the terpenic compounds, and for the C_{10} group the term monoterpene should be used exclusively.

Zusammenfassung

Es wird eine Übersicht über die Kohlenstoffgerüste der Sesquiterpene, Diterpene und Triterpene gegeben. Die Isoprenregel zeigt in jeder dieser Gruppen besondere Eigenheiten. Auch die C_{30}-Steroide und ein Steroid mit 31 Kohlenstoffatomen, die mit den Triterpenen verwandt sind, werden besprochen. Die Zusammenhänge zwischen Diterpenen und Triterpenen in bezug auf ihre Konstitution und Konfiguration werden diskutiert, wobei auf die weitgehende Übereinstimmung ihrer Konfiguration mit der Konfiguration der Steroide hingewiesen wird. Die kürzlich durch die Experimentalarbeiten von BLOCH gewonnenen Kenntnisse über die biologische Entstehung des Cholesterins aus Essigsäure, unter sehr wahrscheinlicher Zwischenbildung des Triterpens Squalen, geben Anlass zur Erörterung hypothetischer Wege für die Biogenese der pentazyklischen Triterpene aus Squalen, der zyklischen Sesquiterpene aus Farnesol, der zyklischen Diterpene aus Geranyl-geraniol und schliesslich auch der zyklischen Monoterpene aus Geraniol, unter Beachtung der elektronischen Zyklisierungsmechanismen der organischen Chemie. Schliesslich wird die «biogenetische» Isoprenregel definiert und ihre Bedeutung diskutiert.

---

[1] *Cf.* page 357.

[2] *E.g. santene* (C_9), *irone* (C_{14}), *lupulone* (C_{21}).

Brèves communications - Kurze Mitteilungen
Brevi comunicazioni - Brief Reports

The Configuration of Digoxigenin

Although crystalline Digoxin (the tridigitoxoside of digoxigenin) was isolated in 1930[1] and has been in clinical use in the United Kingdom since 1933[2] the accepted proof of its structure is vitiated by a discrepancy between the optical rotation of the derived methyl 3:12-dihydroxyetianate, $[\alpha]_D + 39°$ (methanol)[3] and of authentic methyl $3\alpha:12\beta$-dihydroxyetianate, $[\alpha]_D + 52°$ (methanol)[1], with which it gave no melting point depression[2]. This matter was discussed with Professor REICHSTEIN and Dr. D. A. H. TAYLOR in 1952 and in the knowledge that the former was planning to prepare the four isomeric 3:12-dihydroxyetianates and the latter was repeating the degradation of digoxigenin, we studied partial acetylation of digoxigenin and the anhydrodigoxigenins.

---

[1] S. SMITH, J. Chem. Soc. *1930*, 508.

[2] Brit. Med. J. *295*, 364 (1933).

[3] M. STEIGER and T. REICHSTEIN, Helv. chim. Acta *21*, 828 (1938).

[1] V. WENNER and T. REICHSTEIN, Helv. chim. Acta *27*, 965 (1944).

[2] M. STEIGER and T. REICHSTEIN, Helv. chim. Acta *21*, 828 (1938). – H. L. MASON and W. M. HOEHN, J. Amer. Chem. Soc. *60*, 2824 (1938); *61*, 1614 (1939).