1 Terpenes: Importance, General Structure, and Biosynthesis

1.1 Term and Significance

The term terpenes originates from turpentine (*lat.* balsamum terebinthinae). Turpentine, the so-called "resin of pine trees", is the viscous pleasantly smelling balsam which flows upon cutting or carving the bark and the new wood of several pine tree species (Pinaceae). Turpentine contains the "resin acids" and some hydrocarbons, which were originally referred to as terpenes. Traditionally, all natural compounds built up from isoprene subunits and for the most part originating from plants are denoted as terpenes [1] (section 1.2).

Conifer wood, balm trees, citrus fruits, coriander, eucalyptus, lavender, lemon grass, lilies, carnation, caraway, peppermint species, roses, rosemary, sage, thyme, violet and many other plants or parts of those (roots, rhizomes, stems, leaves, blossoms, fruits, seed) are well known to smell pleasantly, to taste spicy, or to exhibit specific pharmacological activities. Terpenes predominantly shape these properties. In order to enrich terpenes, the plants are carved, e.g. for the production of incense or myrrh from balm trees; usually, however, terpenes are extracted or steam distilled, e.g. for the recovery of the precious oil of the blossoms of specific fragrant roses. These extracts and steam distillates, known as ethereal or essential oils ("essence absolue") are used to create fine perfumes, to refine the flavor and the aroma of food and drinks and to produce medicines of plant origin (phytopharmaca).

The biological and ecochemical functions of terpenes have not yet been fully investigated. Many plants produce volatile terpenes in order to attract specific insects for pollination or otherwise to expel certain animals using these plants as food. Less volatile but strongly bitter-tasting or toxic terpenes also protect some plants from being eaten by animals (antifeedants). Last, but not least, terpenes play an important role as signal compounds and growth regulators (phytohormones) of plants, as shown by preliminary investigations.

Many insects metabolize terpenes they have received with their plant food to growth hormones and pheromones. Pheromones are luring and signal compounds (sociohormones) that insects and other organisms excrete in order to communicate with others like them, e.g. to warn (alarm pheromones), to mark food resources and their location (trace pheromones), as well of assembly places (aggregation pheromones) and to attract sexual partners for copulation (sexual pheromones). Harmless to the environment, pheromones may replace conventional insecticides to trap harmful and damaging insects such as bark beetles.

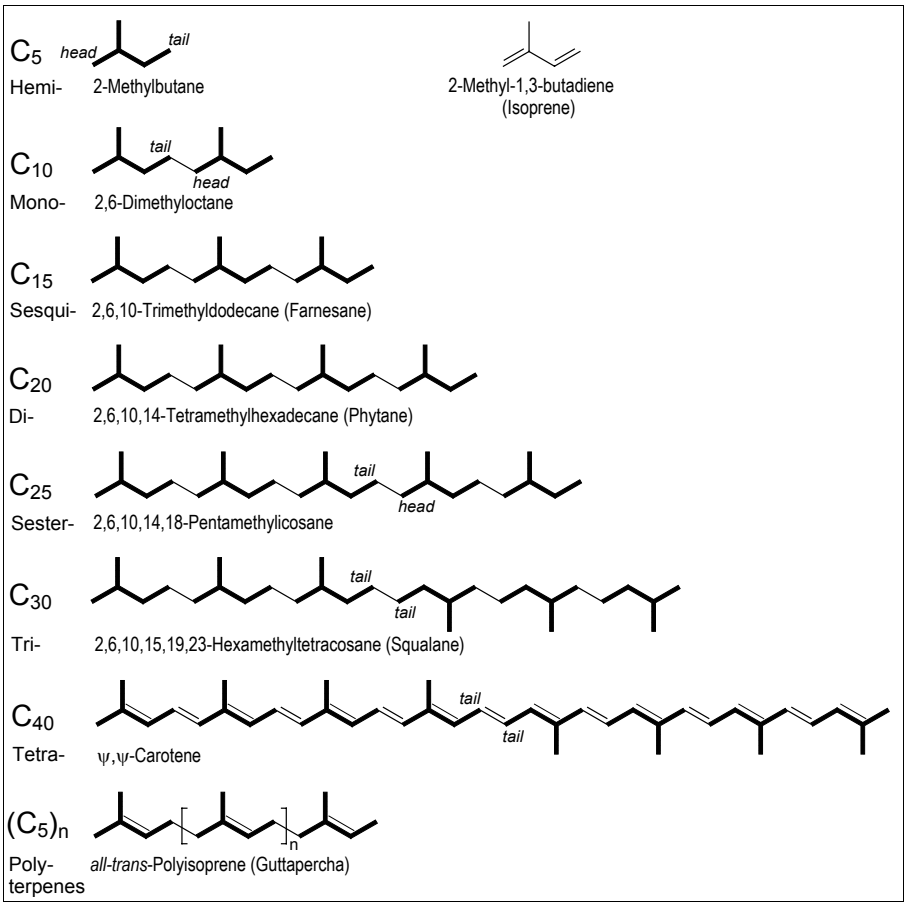
Terpenes. Eberhard Breitmaier.

ISBN: 3-527-31786-4

1.2 General Structure: The Isoprene Rule

About 30 000 terpenes are known at present in the literature [2-7]. Their basic structure follows a general principle: *2-Methylbutane* residues, less precisely but usually also referred to as *isoprene* units, $(C_5)_n$, build up the carbon skeleton of terpenes; this is the isoprene rule [1] found by RUZICKA and WALLACH (Table 1). Therefore, terpenes are also denoted as *isoprenoids*. In nature, terpenes occur predominantly as hydrocarbons, alcohols and their glycosides, ethers, aldehydes, ketones, carboxylic acids and esters.

Table 1. Parent hydrocarbons of terpenes (isoprenoids).

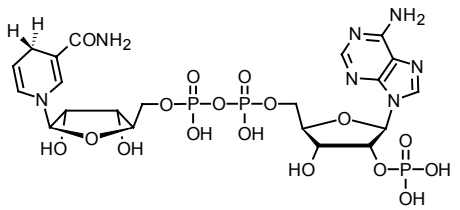


Depending on the number of 2-methylbutane (isoprene) subunits one differentiates between *hemi-* (C_5), *mono-* (C_{10}), *sesqui-* (C_{15}), *di-* (C_{20}), *sester-* (C_{25}), *tri-* (C_{30}), *tetraterpenes* (C_{40}) and *polyterpenes* $(C_5)_n$ with $n > 8$ according to Table 1.

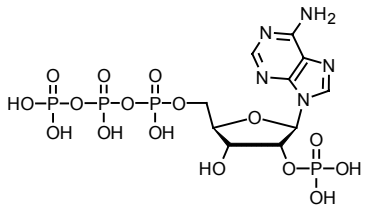
The isopropyl part of 2-methylbutane is defined as the *head*, and the ethyl residue as the *tail* (Table 1). In mono-, sesqui-, di- and sesterterpenes the isoprene units are linked to each other from *head-to-tail*; tri- and tetraterpenes contain one *tail-to-tail* connection in the center.

1.3 Biosynthesis

Acetyl-coenzyme A, also known as *activated acetic acid*, is the biogenetic precursor of terpenes (Figure 1) [9-11]. Similar to the CLAISEN condensation, two equivalents of acetyl-CoA couple to acetoacetyl-CoA, which represents a biological analogue of acetoacetate. Following the pattern of an aldol reaction, acetoacetyl-CoA reacts with another equivalent of acetyl-CoA as a carbon nucleophile to give β-hydroxy-β-methylglutaryl-CoA, followed by an enzymatic reduction with dihydronicotinamide adenine dinucleotide (NADPH + H^+) in the presence of water, affording (*R*)-*mevalonic acid*. Phosphorylation of mevalonic acid by adenosine triphosphate (ATP) *via* the monophosphate provides the diphosphate of mevalonic acid which is decarboxylated and dehydrated to *isopentenylpyrophosphate* (isopentenyldiphosphate, IPP). The latter isomerizes in the presence of an isomerase containing SH groups to *γ,γ-dimethylallylpyrophosphate*. The electrophilic allylic CH_2 group of γ,γ-dimethylallylpyrophosphate and the nucleophilic methylene group of isopentenylpyrophosphate connect to *geranylpyrophosphate* as *monoterpene*. Subsequent reaction of geranyldiphosphate with one equivalent of isopentenyldiphosphate yields *farnesyldiphosphate* as a *sesquiterpene* (Fig. 1).



Dihydro **n**icotinamide **a**denine **d**inucleotide **p**hosphate (NADPH + H^+)



Adenosine **t**ri **p**hosphate (ATP)

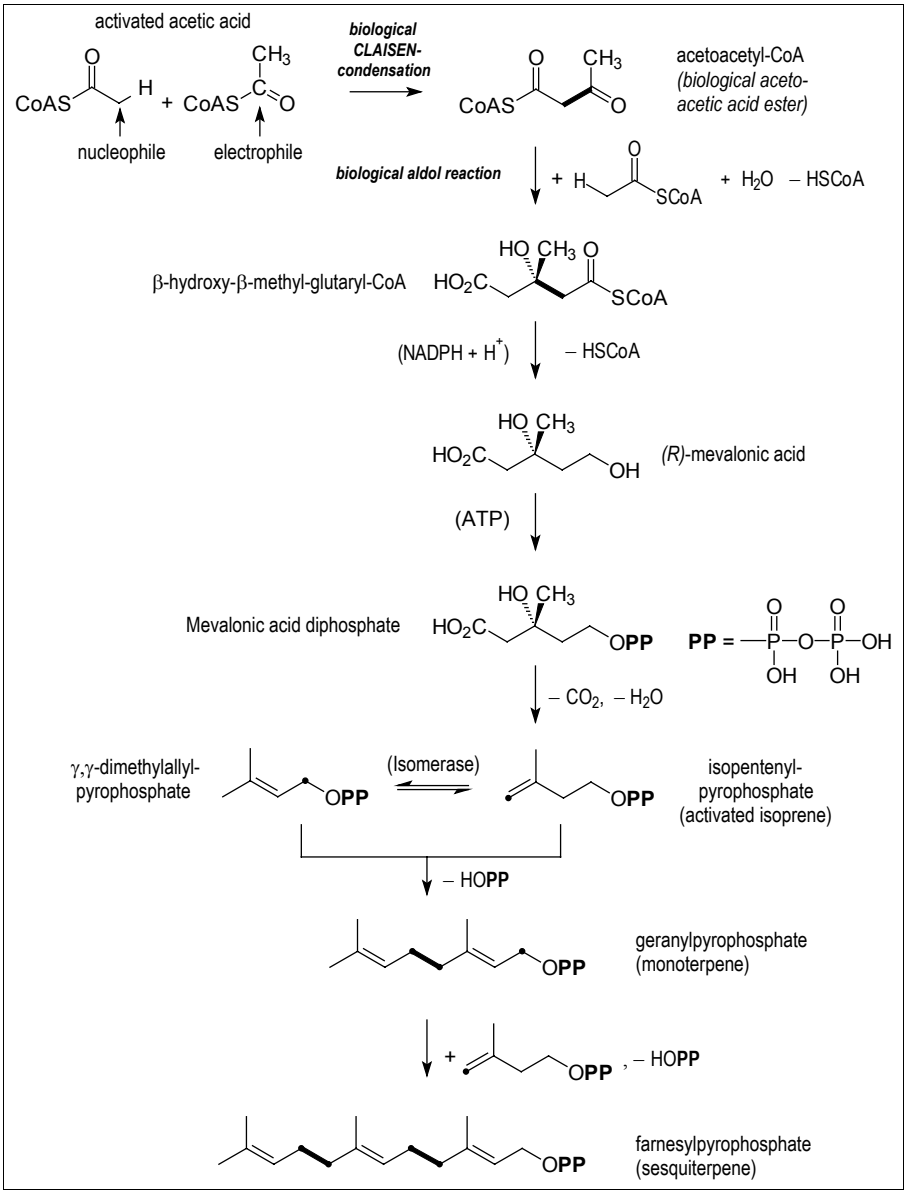
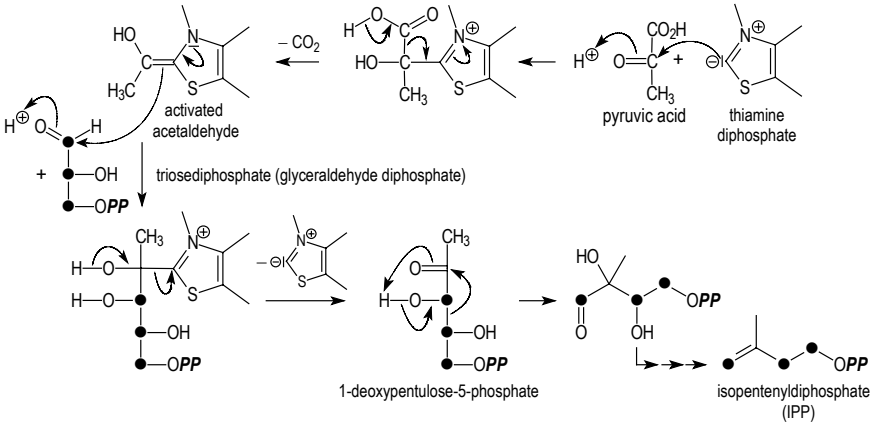


Figure 1. Scheme of the biogenesis of mono- and sesquiterpenes.

However, failing incoporations of ^{13}C-labeled acetate and successful ones of ^{13}C-labeled glycerol as well as pyruvate in hopanes and ubiquinones showed isopentenyldiphosphate (IPP) to originate not only from the acetate mevalonate pathway, but also from *activated acetaldehyde* (C_2, by reaction of pyruvate and thiamine diphosphate) and glyceraldehyde-3-phosphate (C_3) [12]. In this way, *1-deoxy-pentulose-5-phosphate* is generated as the first unbranched C_5 precursor of IPP.



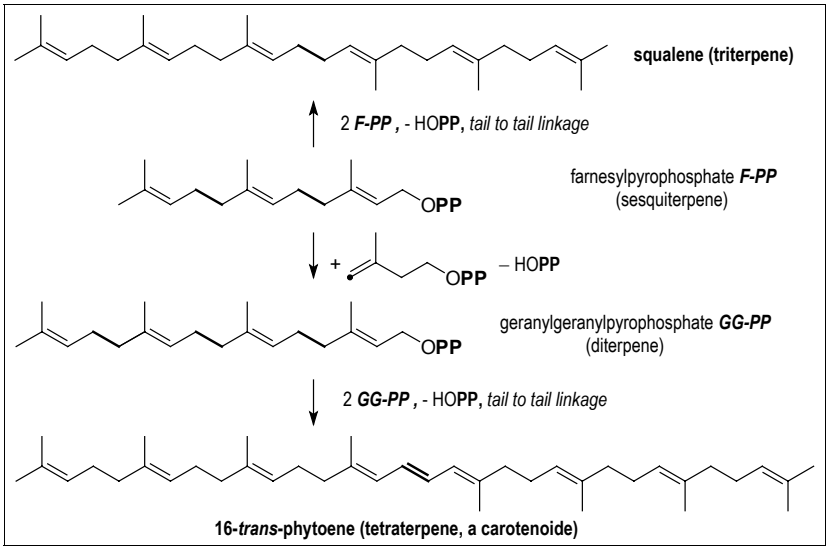
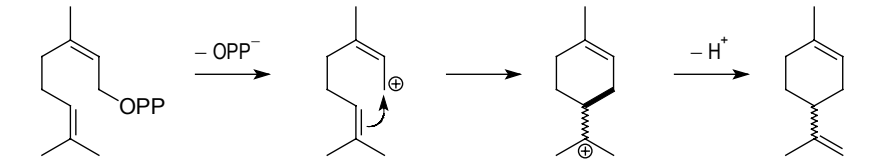


Figure 2. Scheme of the biogenesis of di-, tri- and tetraterpenes.

Geranylgeranylpyrophosphate as a diterpene (C_{20}) emerges from the attachment of isopentenylpyrophosphate with its nucleophilic head to farnesylpyrophosphate with its electrophilic tail (Fig. 2). The formation of sesterterpenes (C_{25}) involves an additional head-to-tail linkage of isopentenylpyrophosphate (C_5) with geranylgeranylpyrophosphate (C_{20}). A tail-to-tail connection of two equivalents of farnesylpyrophosphate leads to squalene as a triterpene (C_{30}, Fig. 2). Similarly, tetraterpenes such as the carotenoid 16-*trans*-phytoene originate from tail-to-tail dimerization of geranylgeranylpyrophosphate (Fig. 2).

The biogeneses of cyclic and polycyclic terpenes [9,10] are usually assumed to involve *intermediate carbenium ions*, but evidence for this *in vivo* was given only in some specific cases. In the simple case of monocyclic monoterpenes such as limonene the allylic cation remaining after separation of the pyrophosphate anion cyclizes to a cyclohexyl cation which is deprotonated to (*R*)- or (*S*)-limonene.



The non-classical version of the intermediate carbenium ion (also referred to as a carbonium ion) resulting upon dissociation of the pyrophosphate anion from farnesylpyrophosphate explains the cyclization to several cyclic carbenium ions [8], as demonstrated for some sesquiterpenes (Fig. 3). Additional diversity arises from *1,2-hydride* and *1,2-alkyl shifts* (WAGNER-MEERWEIN rearrangements) and *sigmatropic reactions* (COPE rearrangements) on the one hand, and on the other hand from the formation of diastereomers and enantiomers provided that the cyclizations generate new asymmetric carbon atoms (Fig. 3) [8-10].

Thus, the non-classical carbenium ion arising from dissociation of the diphosphate anion from farnesylpyrophosphate permits formation of the monoyclic sesquiterpenes humulatriene and germacratriene after deprotonation (Fig.3). A COPE rearrangement of germacratriene leads to elematriene. Protonation of germacratriene following MARKOWNIKOW orientation initially provides the higher alkylated and therefore more stable carbenium ion which undergoes 1,2-hydride shifts resulting in bicyclic carbenium ions with an eudesmane or guaiane skeleton. Subsequent deprotonations yield diastereomeric eudesmadienes and guajadienes. Finally, eudesmanes may rearrange to eremophilanes involving 1,2-methyl shifts (Fig. 3).

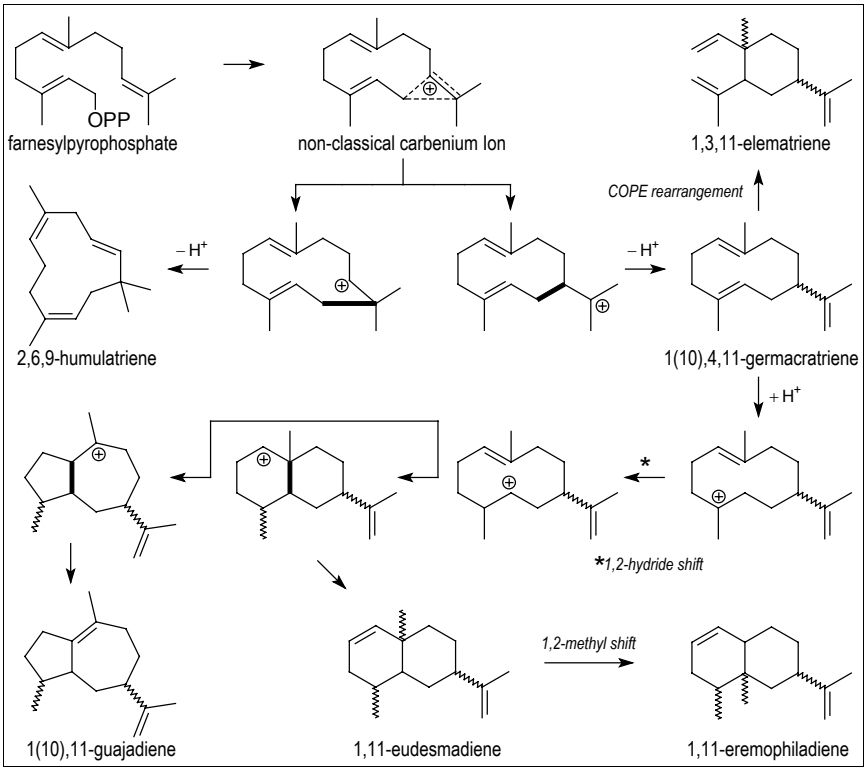
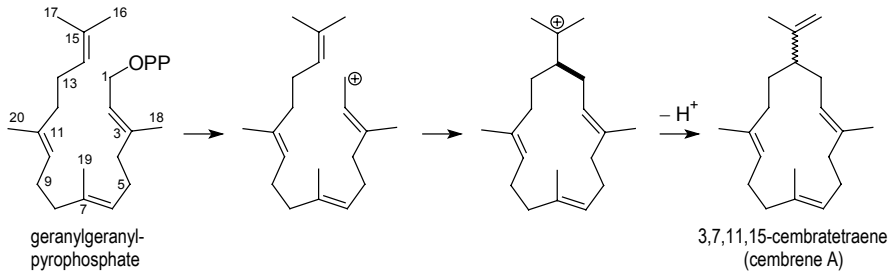


Figure 3. Biogenesis of some mono- and bicyclic sesquiterpenes from farnesylpyrophosphate.

A simliar cyclization generates the 14-membered skeleton of cembrane from which other polycyclic diterpenes are derived. 3,7,11,15-Cembratetraene, better known as cembrene A, emerges directly from geranylgeranylpyrophosphate (Fig. 2) involving the 1,14-cyclization of the resulting allylic cation [9,10].



The biogenesis of pimarane, the parent compound of many polycyclic diterpenes, is assumed to arise from *iso*-geranylgeranylpyrophosphate [9,10]. After dissociation of the pyrophosphate anion, the remaining acyclic allylic cation undergoes a 1,3-sigmatropic hydrogen shift and thereby cyclizes to a monocyclic carbenium ion which, itself, isomerizes to the ionic precursor of the pimarane skeleton.

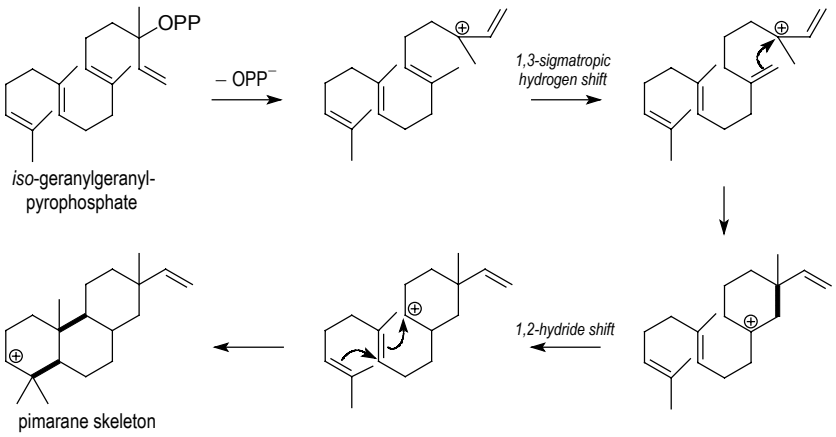
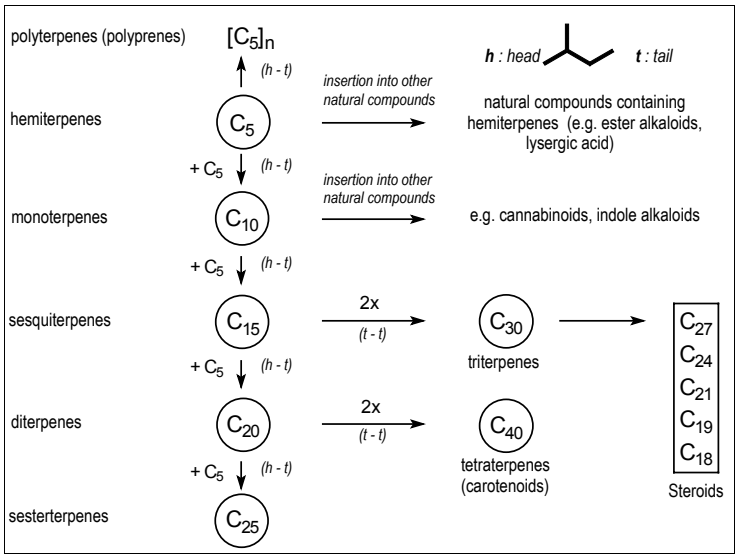


Table 2. Isoprenoids



2,3-Epoxysqualene has been shown by isotope labeling to be the biogenetic precursor of tetracyclic triterpenes with perhydrocyclopenta[*a*]phenanthrene as the basic skeleton (also referred to as *gonane* or *sterane*). *Steroids* [13] are derived from these tetracyclic triterpenes. These include cholestanes (C_{27}), pregnanes (C_{21}), androstanes (C_{19}) with *trans* fusion of the rings *A* and *B* (5α), estranes (C_{18}) with a benzenoid ring *A* (estra-1,3,5-triene; Fig. 4) [9,10] as well as cholic acid and its derivatives (C_{24}) with *cis* fusion of the rings *A* and *B* (5β). The biogenetic origins of tetracyclic triterpenes and steroids are summarized in Table 2.

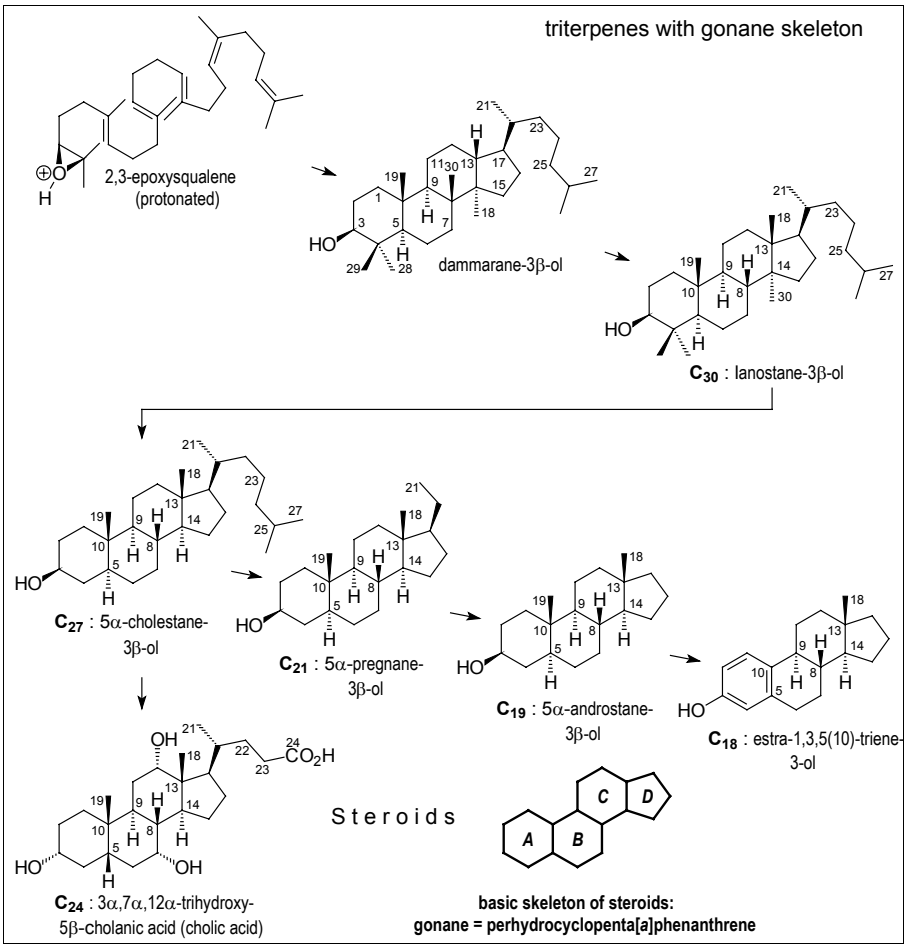


Figure 4. Biogenetic origin of steroids.