A Glimpse into the Biosynthesis of Terpenoids

Abdallah, Ingy I.; Quax, Wim J.

Published in:
International Conference on Natural Resources and Life Sciences (NRLS-2016)

DOI:
10.18502/kls.v3i5.981

Publication date:
2017

Conference Paper

A Glimpse into the Biosynthesis of Terpenoids

Ingy I. Abdallah and Wim J. Quax

Dept. of Chemical and Pharmaceutical Biology, Groningen Research Institute of Pharmacy, University of Groningen, 9713AV Groningen, Groningen, The Netherlands

Abstract

Terpenoids represent the largest class of natural products with a diverse array of structures and functions. Many terpenoids have reported therapeutic properties such as antimicrobial, anti-inflammatory, immunomodulatory and chemotherapeutic properties making them of great interest in the medical field. Also, they are widely used in the flavors and fragrances industries, in addition to being a source of biofuels. Terpenoids suffer from low natural yields and complicated chemical synthesis, hence the need for a more sustainable production method. Metabolic engineering provide an excellent opportunity to construct microbial cell factories producing the desired terpenoids. The biosynthetic mevalonate and non-mevalonate pathways involved in the production of terpenoid precursors are fully characterized so exploring methods to improve their flux would be the first step in creating a successful cell factory. The complexity and diversity of terpenoid structures depends mainly on the action of the terpene synthases responsible for their synthesis. These enzymes are classified into different classes and gaining insight into their catalytic mechanism will be useful in designing approaches to improve terpenoid production. This review focuses on the biosynthesis and biodiversity of terpenoids, understanding the terpene synthase enzyme family involved in their synthesis and the engineering efforts to create microbial cell factories for terpenoid production.

Keywords: amorphadiene; artemisinin; *Bacillus subtilis*; *Escherichia coli*; mevalonate; MEP; terpenoids; terpene synthases; taxol; taxadiene.

Corresponding Author:
Wim J. Quax
w.j.quax@rug.nl

Received: 9 June 2017
Accepted: 15 July 2017
Published: 11 September 2017

Publishing services provided by Knowledge E

Selection and Peer-review under the responsibility of the NRLS Conference Committee.

1. Introduction

Nature is a treasure chest of an infinite number of commercially and/or medicinally significant compounds. Historically, most of new medicines have been derived from natural products (secondary metabolites) where chemical compounds from animals, plants and microbes have been invaluable in treating different human diseases ever since the dawn of medicine. Natural products have the inherent properties of high structural diversity and biochemical specificity making them leading scaffolds for drug discovery in addition to their use in food and fragrance industries [1–4]. It has been reported that 34 % out of new small-molecule medicines approved by the Food and Drug Administration (FDA) in the period of 1981 to 2010 were actually natural products

or derivatives of natural products [5]. Additionally, more than 60 % of chemotherapy drugs and 75 % of drugs for infectious diseases are of natural origin [6].

Terpenoids, with around 64 000 known compounds, are considered the largest and most diverse class of natural products. Terpenoids are secondary metabolites mostly produced by plants and some by bacteria or yeast. They occur in various chemical structures in a usual assortment of linear hydrocarbons or chiral carbocyclic skeletons with different chemical modifications such as hydroxyl, ketone, aldehyde and peroxide groups. Different terpenoidal molecules have been reported to have antimicrobial, antifungal, antiviral, antiparasitic, antihyperglycemic, antiallergenic, anti-inflammatory, antispasmodic, immunomodulatory and chemotherapeutic properties. They can also be used as natural insecticides and protective substances in storing agriculture products. This diverse array of terpenoid structures and functions has incited great interest in their medicinal use and commercial applications as flavors, fragrances and spices. Moreover, terpenoids recently emerge as strong players in the biofuel market. Among the terpenoids with established medical applications are the antimalarial artemisinin and the anticancer taxol [6–9].

This review delves into the world of terpenoids. A brief overview of the importance of terpenoids, their different classes and biosynthesis shedding more light on the key enzymes involved in their synthesis, namely terpene synthases. Additionally, the trend of biosynthesis of terpenoids in engineered microorganisms is discussed.

2. Biosynthesis of terpenoids

Despite the enormous structural differences between terpenoids, they are all derived from the same C_5 skeleton of isoprene. The terpenoidal backbone is synthesized from the two precursors: isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP) through a different number of repeats, rearrangement and cyclization reactions. Two distinct biosynthetic pathways for the formation of these universal precursors have been reported, the classical mevalonate (MVA) pathway and the most recently characterized 2*C*-methyl-D-erythritol-4-phosphate (MEP) pathway, also known as the 1-deoxy-D-xylulose- 5-phosphate (DXP) pathway. The MVA pathway is present in eukaryotes (all mammals, the cytosol and mitochondria of plants, fungi), archaea, and some eubacteria while the non-mevalonate pathway occur in eubacteria, algae, cyanobacteria, and the chloroplasts of plants. The MVA pathway comprises seven enzymatic reactions to convert the precursor acetyl-CoA to IPP and DMAPP (Fig. 1) while the MEP pathway converts the starting materials, pyruvate and glyceraldehyde-3-phosphate, to IPP and DMAPP through eight enzymatic reactions (Fig. 1) [10–12]. The linear prenyl diphosphates such as geranyl pyrophosphate (GPP), farnesyl pyrophosphate (FPP), geranylgeranyl pyrophosphate (GGPP), and farnesyl

geranyl pyrophosphate (FGPP) are synthesized from the two basic building blocks, IPP and DMAPP where a group of enzymes called prenyltransferases repeatedly add the active isoprene unit (IPP) to DMAPP or a prenyl diphosphate in consecutive head-to-tail condensations leading to the production of a range of molecules with fixed lengths and stereochemistry. Geranyl pyrophosphate synthase (GPPS) and farnesyl pyrophosphate synthase (FPPS) catalyze the condensation of IPP and DMAPP to produce GPP (C_{10}) and FPP (C_{15}). Geranylgeranyl pyrophosphate synthase (GGPPS) and farnesyl geranyl pyrophosphate synthase (FGPPS) are responsible for formation of GGPP (C_{20}) and FGPP (C_{25}). The precursors GPP, FPP, GGPP and FGPP, are cyclized and/or rearranged by different terpene synthase enzymes to produce the different classes of terpenoids [6, 13].

3. Classification of terpenoids

Terpenoids are usually classified according to the number and structural organization of the five carbon isoprene units involved in their synthesis as C_5 hemiterpenoids, C_{10} monoterpenoids, C_{15} sesquiterpenoids, C_{20} diterpenoids, C_{25} sesterterpenoids, C_{30} triterpenoids, C_{40} tetraterpenoids, and $C_{>40}$ polyterpenoids. The properties, significance and examples of the different classes are briefly discussed.

3.1. Hemiterpenoids (C_5)

Hemiterpenoids are the smallest known terpenoids where they are composed of a single five carbon atoms unit. The most famous of which is the volatile hydrocarbon isoprene (Fig. 2). Isoprene is a potential biofuel and a valuable polymer building block in the synthetic chemistry industry. Currently, About 95 % of isoprene production is used to produce cis-1,4-polyisoprene, a synthetic version of natural rubber. The enzyme isoprene synthase is responsible for the conversion of DMAPP to produce isoprene. Many plants possess isoprene synthase but harvesting the volatile isoprene from plants is difficult. Hence, isoprene-producing microorganisms grown in a closed bioreactor offer a more suitable production system for isoprene [14].

3.2. Monoterpenoids (C_{10})

Monoterpenoids are acyclic, monocyclic, or bicyclic C_{10} compounds synthesized from the substrate GPP by monoterpene synthases. Monoterpenoids are components of the essential oils extracted from many plants contributing to the flavor and aroma of these plants. They have high diversity and are widely used in pharmaceutical, cosmetic, agricultural and food industries. A few examples of monoterpenoids (Fig. 2) are the

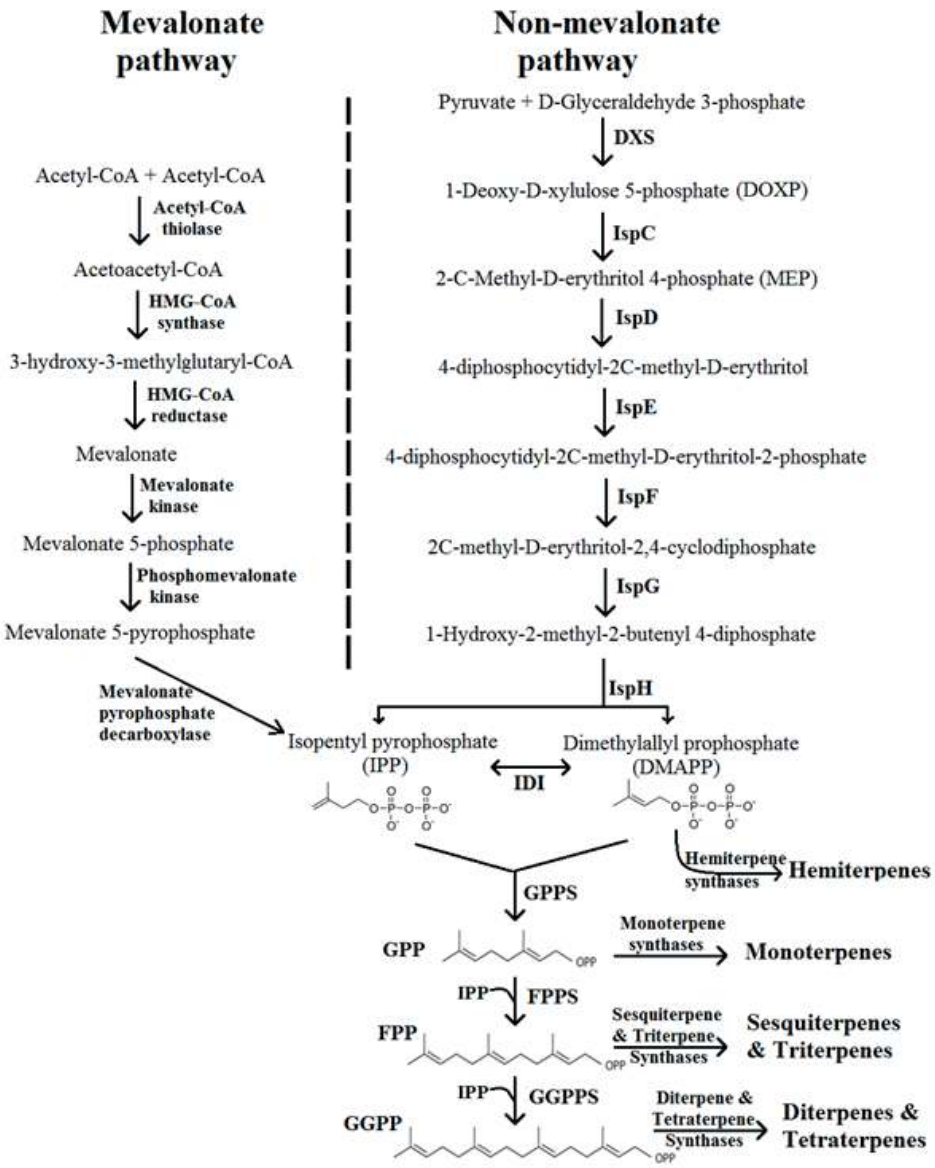

Figure 1: Biosynthetic pathways for terpenoid production.

acyclic myrcene from hops and linalool from lavender, the monocyclic menthol from mint and thymol from thyme, and the bicyclic eucalyptol from eucalyptus and camphor from camphor trees [15, 16].

3.3. Sesquiterpenoids (C_{15})

Sesquiterpenoids are widely distributed in nature and represent the most prevailing class of terpenoids. They are acyclic, monocyclic, bicyclic or tricyclic C_{15} compounds synthesized from the substrate FPP by sesquiterpene synthases. Interestingly, another class of compounds bearing characteristic features as an α-methylene γ-lactone system; α, β-unsaturated carbonyls, or epoxides and chemically distinct

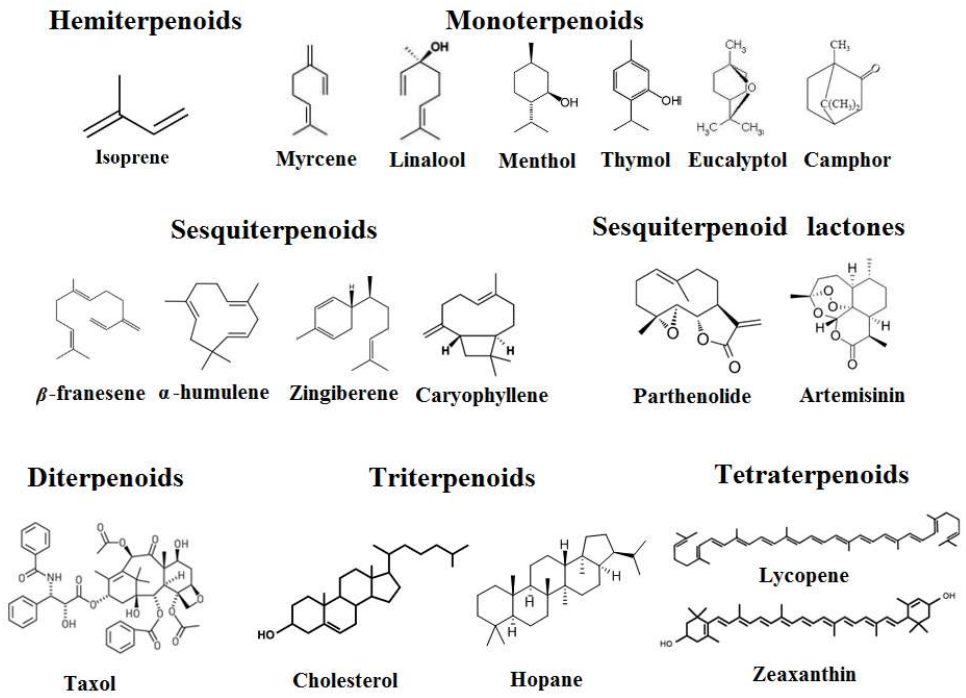

Figure 2: Examples of different classes of terpenoids.

from the sesquiterpenoids are collectively named sesquiterpenoid lactones. Both sesquiterpenoids and sesquiterpenoid lactones demonstrate a wide range of biological functions as antimicrobial, anti-inflammatory and antitumor agents. Some known sesquiterpenoids (Fig. 2) are β-farnesene, α-humulene, zingiberene, and caryophyllene. Among the most famous sesquiterpenoid lactones (Fig. 2) are parthenolide and the antimalarial artemisinin [17, 18].

3.4. Diterpenoids (C_{20})

Diterpenoids are structurally diverse non-volatile C_{20} hydrocarbons derived from the substrate GGPP by the diterpene synthase enzyme family. It has been reported that they mostly originate from plant or fungal sources, but they are also formed by certain insects as well as marine organisms. Chemical synthesis of these compounds is difficult due to their complex structures, and natural extraction is laborious so production in microbial hosts is of great interest. Taxol (Fig. 2) is a well-known diterpenoid that is used in the treatment and management of cancer [6, 19].

3.5. Sesterterpenoids (C_{25})

Sesterterpenoids are rare in nature and are formed from the precursor FGPP. They are generally found in protective waxes of insects and fungi [19].

3.6. Triterpenoids (C_{30})

Triterpenoids are C_{30} hydrocarbons biosynthesized from six isoprene units where they share the acyclic precursor squalene. Based on the numerous possible manners of ring closure in squalene, a large number of triterpenoids having a diversity of skeleton structures can be produced. Squalene itself is a natural antioxidant and is used commercially in cosmetics, nutrition and in vaccines. Triterpenoids may be categorized into two major groups, the steroidal (C_{27}) type with 27 carbon atoms present in the skeleton and the pentacyclic (C_{30}) type. Cholesterol is an example of steroidal triterpenoids and hopane is a pentacyclic triterpenoid (Fig. 2) [19].

3.7. Tetraterpenoids (C_{40})

Tetraterpenoids are C_{40} compounds derived from phytoene formed by two C_{20} GGPPs in a head-to-head condensation reaction. The most famous group of tetraterpenoids is the carotenoid pigments. Carotenoids have important biological functions due to their antioxidant activity, in addition to their commercial use as food colorants. Lycopene and zeaxanthin (Fig. 2) are considered tetraterpenoids [19].

4. Significance of selected terpenoids

4.1. Artemisinin

Artemisinin (Fig. 2), also known as *qinghao su*, is a sesquiterpenoid lactone naturally produced by the plant *Artemisia annua* L. The Nobel Prize was awarded to Youyou Tu in 2015 for her discovery of artemisinin, which she denotes as a gift from traditional Chinese medicine to the world. Artemisinin-based combination therapies (ACTs) are endorsed by The World Health Organization (WHO) as the first-line treatment for *Plasmodium falciparum* malaria. The suggested mechanism of artemisinin is that its endoperoxide moiety interacts with heme, which is ample in malaria parasites, resulting in the generation of carbon-based free radicals which in turn cause death of the *P. falciparum* parasite. Recently, it has been reported that artemisinin has anticancer effect where cancer cells, similar to the malaria parasites, possess high concentration of free iron. Cell death also results from the formation of free radicals by the artemisinin-iron reaction. The benefit of artemisinin as an anticancer agent is not only its potency, but also its selectivity against cancer cells and low toxicity to normal cells. Artemisinin commercial production still largely relies on extraction from its

natural source making ACTs more expensive than other less effective malaria treatments. Hence, research into creating microbial cell factories for sustainable production of artemisinin is of great importance [20, 21, 24].

4.2. Taxol

Taxol (Fig. 2), also known as paclitaxel, is a diterpenoid first isolated from the *Taxus brevifolia* bark. In 1982, it was approved by the FDA as a medicine against different forms of cancer, including various carcinoma's (ovary, breast, lung, head, neck, bladder and cervix), melanoma and AIDS related Kaposi's sarcoma. The activity of taxol is based on inhibition of mitosis where it targets tubulin causing difficulty with the spindle assembly, cell division and also chromosome segregation. Recently, taxol has been reported to be useful in treating neurodegenerative diseases such as Alzheimer's disease. The main shortcoming of taxol seems to be its mass production, which can be resolved by exploring microbial synthesis of paclitaxel since total chemical synthesis proves problematic due to its complex structure [25–27].

5. Terpene synthases

Terpene synthases are a family of enzymes responsible for catalyzing the rearrangement and/or cyclization of the precursors GPP, FPP, and GGPP to produce the different classes of terpenoids. The involvement of a terpene synthase is an indispensable requirement for the production of terpenoids. The fascinating structural diversity of terpenoids is based on the orientation of their substrate in the active site of their correlated terpene synthase which then undergo a series of cyclizations and/or rearrangements to produce a certain terpenoid. Terpene synthases are classified into class I and class II terpene synthases based on their substrate activation mechanisms. Class I terpene synthases are characterized by catalyzing the ionization of the allylic diphosphate ester bond in their isoprenyl substrates while class II terpene synthases catalyze protonation-induced cyclization reaction of the substrate, sometimes followed by rearrangement. In addition to the different substrate activation mechanisms, the two different classes of terpene synthases possess unrelated protein folds. A class I terpene synthase uses a tri-nuclear metal cluster liganded by conserved metal ion binding motifs **D**DXX**D** and (**N**,**D**)DXX(**S**,**T**)XXX**E** (bold indicates typical metal ligands) to trigger the ionization of the diphosphate group of their substrate, which initiate catalysis by producing a carbocation. On the other hand, a class II terpene synthase employs general acid catalysis to initiate carbocation generation, using the middle aspartic acid in a DXDD motif to protonate a substrate double bond or oxirane moiety. It is also important to mention that terpene synthases can be further classified into

transoid and cisoid subclasses. The transoid synthases catalyze the ionization of the (*trans*,*trans*)-substrate to generate a transoid intermediate carbocation constituting the backbone of their products while the cisoid synthases perform an initial C-C double bond isomerization producing a (*cis*,*trans*)-intermediate carbocation. This streochemical distinction accounts for the transoid and cisoid distinct product families. Terpene synthases have a wide array of product profile and it is not frequently likely to predict their product profile based on their primary structure alone. Hence, three-dimensional (3D) structures of the enzymes is essential in elucidating structure-function relationships of the amino acid residues in different positions to the catalytic process. Reported structural analyses show a relationship between the catalytic mechanism and the topology of the active site pocket found at the carboxy-terminal domain. In spite of the general structural similarity of terpene synthases, the identification of individual amino acids that are associated with specific mechanistic steps is a difficult task. Promiscuous activity in terpenoid biosynthesis is ingrained in the leniency of the enzyme template that chaperones the conformations of malleable substrate(s) and intermediate(s) through multistep reactions till product formation. Therefore, there is numerous efforts directed towards probing active site residues of different terpene synthases aiming at testing structure-function relationships of protein residues and engineering enzymes with improved catalytic efficiency, product specificity or thermostability [13, 28–33].

5.1. Class I terpene synthases

The class I terpene synthases are characterized by being ionization-initiating enzymes. Microbial Class I terpene synthases are composed of structurally homologous α domains, even in absence of readily obvious sequence homology. Their active site is found within the α domain, which assumes a common α-bundle fold where an aspartate-rich DDXX(XX)D/E motif alongside a less conserved (N,D)DXX(S,T,G)XXXE motif bind essential magnesium co-factors triggering the departure of the substrate pyrophosphate group, and simultaneously initiating the cyclization and rearrangement reactions. Contrary to microbial class I terpene synthases, most plant monoterpene and sesquiterpene synthases assume an αβ assembly wherein the α fold performs its usual function, but the β fold is inactive. In all class I terpene synthases, the formation of a complex of the substrate, metal ions and metal-ion binding motifs prompts conformational changes that sequester the active site from bulk solvent. This points out that the active site pocket does not adopt its product-like contour until after the binding of the substrate. The terpene synthases trigger ionization of the substrate only in this closed enzyme-substrate complex. After ionization, the initially formed allylic carbocation usually undergo cyclization and/or rearrangement. However, sometimes immediate deprotonation is observed corresponding with the

more general designation of this enzyme family as synthases rather than cyclases (though this later nomenclature would better fit the majority of the enzymes in the family). Additionally, after cyclization and/or rearrangement, these enzymes usually deprotonate the final carbocation. Despite that, capture of water by the final carbocation has been detected, with either direct deprotonation to form a hydroxyl group, or even subsequent cyclization before deprotonation, forming a cyclic ether. Finally, class I terpene synthases display a wide array of catalytic promiscuity. Some are fairly specific while others yield a distinctive range of products from the same substrate [28, 30, 34].

5.1.1. Hemiterpene synthases

Isoprene synthase (ISPS) is the only known hemiterpene synthase. ISPS is responsible for the global production of isoprene in nature and biotechnology. ISPS active site contains magnesium ions that interact with the substrate dimethylallyl diphosphate (DMAPP) catalyzing the elimination of inorganic pyrophosphate to yield isoprene. The structure of ISPS reveals a shallower active site cavity compared to other class I terpene synthases, even the monoterpene synthases. This corresponds with its specificity for the smaller substrate DMAPP [28, 35].

5.1.2. Monoterpenesynthases

All monoterpene synthases catalyze the metal-dependent ionization and cyclization of the 10-carbon precursor geranyl pyrophosphate (GPP) to produce different monoterpenes. Monoterpene synthases accomplish outstanding structural and chemical diversity in their product assortments, despite their catalysis of the simplest terpene cyclization cascades where they use the shortest linear isoprenoid substrate [28, 30]. Limonene synthase from *Mentha spicata* L. is an example of monoterpene synthase that quenches the final cyclized carbocation intermediate by deprotonation to form an olefin (limonene) [36]. Cineole synthase from *Salvia fruticosa* Mill. offers an example of the integration of water to form a cyclic ether (cineole) [37]. Bornyl diphosphate synthase from *Salvia officinalis* L. was the first monoterpene synthase to be structurally described and it displays a distinctive example of re-addition of the pyrophosphate anion to the cyclized final carbocation producing bornyl diphosphate [38]. Microbial monoterpene synthases possess only an α-domain (Fig. 3a) while the plant enzymes has both α and β-domains (Fig. 3b).

5.1.3. Sesquiterpene synthases

Sesquiterpene synthases are responsible for catalyzing the conversion of farnesyl pyrophosphate (FPP) into more than 300 known monocyclic, bicyclic, and tricyclic products. In general, there is low amino acid sequence identity amid sesquiterpene synthases from bacteria, fungi, and plants. However, these enzymes assume the distinctive class I terpene synthases fold where there structural homology comprises not just the α-helical domain but also the signature metal ion binding motifs within, linked to binding the metal co-factors essential for catalysis. Pentalenene synthase from *Streptomyces* UC5319, and epi-isozizaene from *Streptomyces coelicolor* A3(2), trichodiene synthase from *Fusarium sporotrichioides* Sherb, and aristolochene synthase from *Penicillium roqueforti* Thom. are examples of bacterial and fungal sesquiterpene synthases, respectively. Their reported crystal structures depict the characteristic α-domain of this class (Fig. 3a). On the other hand, plant sesquiterpene synthases such as epi-aristolochene synthase from *Nicotiana tabacum* L., δ-cadinene synthase from *Gossypium arboreum* L., and *Artemisia annua* L. α-bisabolol synthase contain an N-terminal domain, in addition to the α-domain (Fig. 3b). This extra domain (β-domain) resembles class II terpene synthase fold and is catalytically silent but plays a part in capping the active site of the C-terminal domain [28, 30, 39].

5.1.4. Diterpene synthases

Diterpene synthases catalyze the cyclization of the linear C_{20} geranylgeranyl pyrophosphate (GGPP) to produce a range of cyclic and polycyclic diterpenes. Among the very few characterized diterpene synthases are taxadiene synthase from *Taxus brevifolia* Nutt. tree and abietadiene synthase from *Abies grandis* (Douglas ex D. Don) Lindl. These plant diterpene synthases contain three domains where in addition to the usual plant terpene synthase β and α domains, they possess a γ domain (fig. 3c) [28–30].

5.2. Class II terpene synthases

The class II terpene synthases are characterized by being protonation-initiating enzymes. This class is composed of Class II diterpene synthases and triterpene synthases which can be squalene-hopene or oxido-squalene synthases. Bacterial diterpene synthases and all triterpene synthases comprise β and γ domains (Fig. 3d) while plant class II diterpene synthases consist of α,β and γ domains (Fig. 3e). Their active site is located between β/γ domains, both of which display an α-barrel fold where a DXDD motif in the β domain offers the proton donor that activates initial carbocation formation. After the initial carbocation production, these enzymes often catalyze

stereochemically complex cyclization reactions producing from one to five rings, followed with subsequent rearrangement. Similar to the class I terpene synthases, enzymes of this class do not essentially directly deprotonate the final carbocation but sometimes water is captured tailed by deprotonation to form a hydroxylated product. Also they exhibit a wide range of catalytic promiscuity [28, 30].

5.3. Selected terpene synthases

5.3.1. Amorpha-4,11-diene synthase

Amorphadiene synthase (ADS) is a class I cisoid sesquiterpene synthase. It is a key enzyme in the biosynthesis of the antimalarial drug artemisinin in the plant *A. annua* where it catalyzes the first rate limiting step of converting the substrate FPP to amorpha-4,11-diene which is the precursor of artemisinin. There is no crystal structure reported for ADS, however, a 3D homology model representing the conformation of this enzymes has been recently published (Fig. 4a). The model was constructed using another sesquiterpene synthase from *A. annua* as a template, namely, α-bisabolol synthase (BOS). Both ADS and BOS share high sequence identity which made BOS the ideal template for homology modelling of ADS. The created model of ADS showed the characteristic metal-ion binding motifs of class I terpene synthases chelating three magnesium ions in the active site. In addition, the substrate FPP was docked in the active site and its correct orientation was confirmed. Since ADS belongs to the cisoid family, its multistep mechanism begins with isomerization of the C2-C3 double bond of FPP to produce nerolidyl diphosphate (NPP) which is ionized into a 2,3-cis-farnesyl cation. This cation will initially undertake 1,6-cyclization to give bisabolyl cation followed by 1,10-ring closure to produce the major product amorpha-4,11-diene. Probing of different amino-acid residues in the active site of ADS helped in providing more insight into its catalytic mechanism. Moreover, efforts of engineering ADS to improve catalytic efficiency and alter product profile have yielded interesting results [24, 40].

5.3.2. Taxadiene synthase

Taxadiene synthase (TXS), a class I diterpene synthase, catalyzes the first step in biosynthesis of taxol in the bark of *T. brevifolia* by metal-dependent cyclization of GGPP to produce taxa-4(5),11(12)-diene which is the precursor of taxol. The full-length of the enzyme is 862-residue (98 kD) but a terminal transit sequence of around 80 amino acid residues is cleaved off after maturation in plastids. Hence, the reported crystal structure of TXS is that of a truncated variant, lacking the transit sequence, complexed

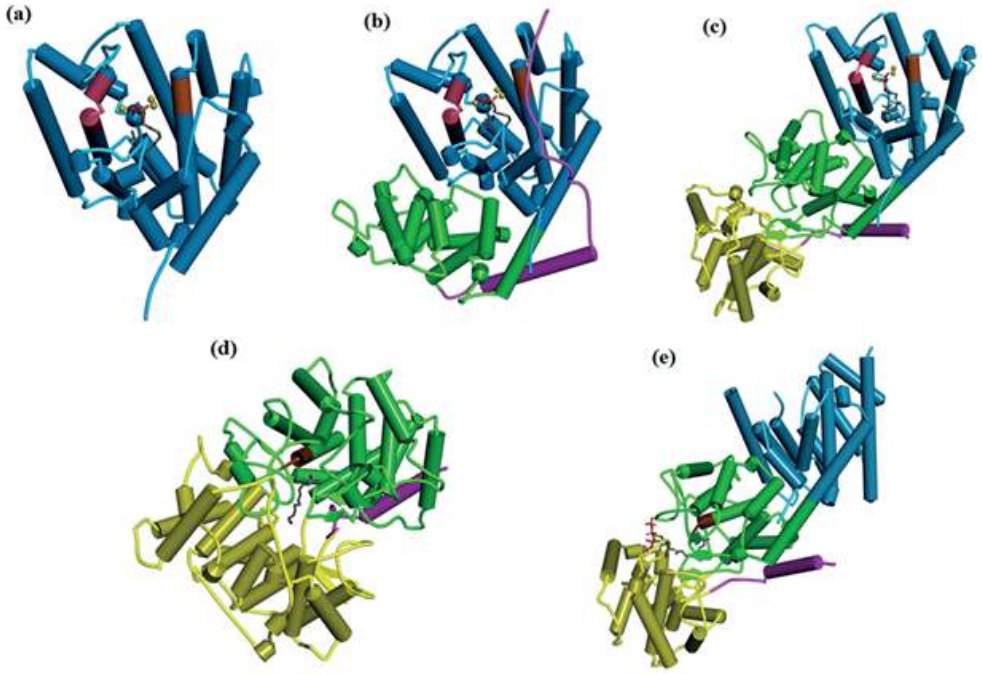

Figure 3: Schematic representation of the general structure of different terpene synthases. (a) Microbial class I mono- and sesquiterpene synthases; (b) Plant class I mono- and sesquiterpene synthases; (c) Plant class I diterpene synthases; (d) Class II triterpene and bacterial diterpene synthases; (e) Plant Class II diterpene synthases. Note that the α domain is in blue, β domain in green, γ domain in yellow and N-terminal in purple. The metal ion binding motifs DDXX(XX)D/E and (N,D)DXX(S,T,G)XXXE are in orange and pink, respectively. The DXDD motif is in brown. The three yellow balls represent magnesium ions and the red side chain is the pyrophosphate group of the substrate.

with its substrate (GGPP) (Fig. 4b). This enzyme have a tri-domain structure where it possess not only the typical plant terpene synthase β and α domains, but also a γ domain, that is inserted between the first and second helices of the β domain so its final structure contains both class I and class II folds. The enzyme C-terminal contains conserved metal-binding motifs with three magnesium metal clusters to bind and activate the substrate but the N-terminal and insertion domain lack the characteristic DXDD motif indicating that the enzyme functions as a class I terpene synthase [28, 29].

6. Engineering microbial cell factories for terpenoid production

The need for sustainable production of terpenoids, being a very famous class of natural products, is massive. The problem of low natural yield of terpenoids and expensive or difficult chemical synthesis can be overcome by engineering microbial cells to act as biofactories for the sustainable production of terpenoids. This approach would require transfer of biosynthetic pathways from the native source of terpenoids to these microbes with all its challenges. These microbial factories provide the benefits of the use of cheap carbon sources, ability to increase production yield by genetic

manipulation, and environmentally friendly chemistry. Since all terpenoids originate from the same C_5 precursors IPP and DMAPP produced by MVA or MEP pathway, engineering a platform strain producing large amounts of these precursors is beneficial for manufacturing different types of terpenoids where the terpene synthase responsible for production of a terpenoid of interest can be directly introduced into the platform strain. In the last few decades, biosynthesis of terpenoids in microorganisms has focused mostly on carotenoids along with precursors for important drugs such as artemisinin and taxol [41, 42]. *Escherichia coli* is one of the most widely used platform organisms. Numerous reports exploiting its inherent MEP pathway by overexpression for production of terpenoids were successful. Also, efforts were made to introduce the heterologous MVA pathway in *E. coli*. Numerous terpenoids including amorphadiene and taxadiene were effectively produced in *E. coli*. One of the drawbacks of *E.coli* is the possibility of contamination of the final product by endotoxins which make it until now not designated as a Generally Regarded As Safe (GRAS) organism by the FDA [43, 44]. Another organism that has been widely researched for terpenoid production is the yeast *Saccharomyces cerevisiae* Meyen ex E.C. Hansen. *S. cerevisiae* can tolerate low pH and increased osmotic pressure compared to bacteria making it highly favored in industry. This yeast possess an endogenous MVA pathway, however most of the FPP produced by the pathway is consumed for production of sterols. Hence, researchers focused on increasing the pool of the GPP, FPP, and GGPP precursors for terpenoid production. This can be achieved by the suppression of competing pathways that drain these precursors along with upregulation of the MVA pathway and expression of the desired terpene synthases. The major disadvantage of *S. cerevisiae* is its slow growth rate so it would take more time to produce the same terpenoid compared to *E. coli* [9, 45]. In the recent years, interest in using *Bacillus subtilis* (Ehrenberg 1835) Corn 1872 as a cell factory for terpenoid production has grown. *B. subtilis* is a Gram-positive bacteria that contains an inherent MEP pathway capable of isoprene production in amounts higher than most eubacteria counting *E. coli*. It has a high growth rate, extensive substrate range and is considered a GRAS organism by the FDA. Hence, *B. subtilis* emerges as a strong candidate for terpenoid production by enhancing the MEP pathway flux. Overexpression of the MEP pathway genes, *dxs* and *idi*, increased the production of amorphadiene in *B. subtilis*. Also, Expression of heterologous *CrtM* and *CrtN* genes in *B. subtilis* successfully allowed the production of C_{30}carotenoids. The production of these carotenoids was further enhanced by overexpression of different MEP pathway genes, in addition to, allowing the systematic analysis of the functionality of the different MEP pathway enzymes [8, 46–48]. Furthermore, Photosynthetic microorganisms as cyanobacteria offer an additional advantage in production of terpenoids over both plants and other microbial systems. Similar to plants, they have the ability to directly use CO_2 as their carbon source and light as their source of energy. They

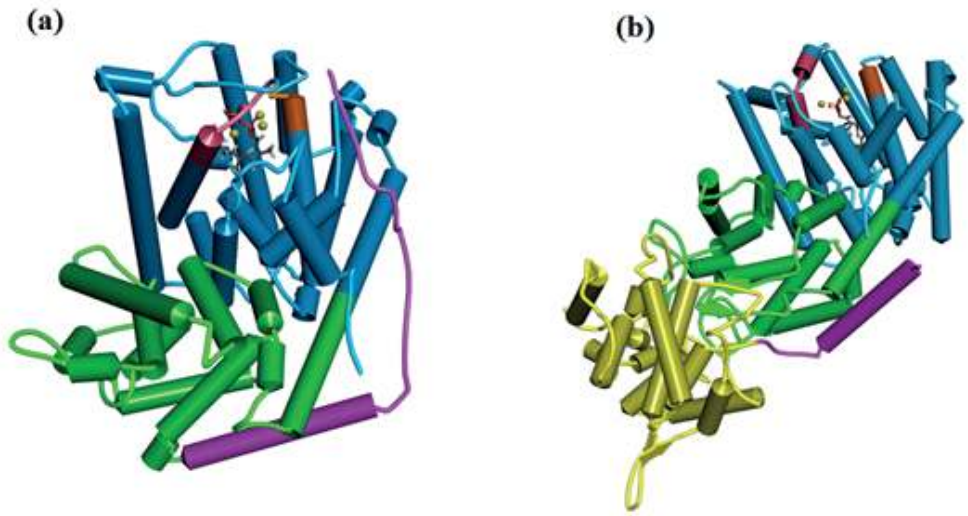

Figure 4: (a) Reported 3D model of amorpha-4,11-diene synthase; (b) Reported crystal structure of taxadiene synthase.

can even perform that more efficiently with faster growth rates and improved solar energy conversion than plants. Simultaneously, certain strains of cyanobacteria have the same upsides as other microbial systems where they can grow to high densities in photobioreactors, can be genetically modified, and provide simpler extraction and purification processes for the target terpenoid than plant systems. Also, they provide better likelihood of functional expression of plant enzymes and metabolic pathways compared to other microbial systems [14, 49].

7. Future perspective

In the medicinal and commercial market, terpenoids will always be valuable compounds of vast interest. The biosynthetic pathways involved in terpenoid production are fully described, however, more insight into the catalytic mechanism of the enzymes involved in these pathways, especially terpene synthase, is of grave importance. The characterization of different terpene synthases and exploring the structure-function relationships of their amino acid residues with regard to their interaction with the substrate will be the basis of manipulating these enzymes. Protein engineering of terpene synthases will provide the chance to improve the enzymes stability, catalytic efficiency and product specificity aiming at more sustainable production of their respective terpenoids. In spite of the progress made in understanding microbial metabolic regulation and creating suitable genetic tools, there are several challenges still facing the construction of microbial cell factories for the commercial production of terpenoids. These challenges can be summarized into the precursor supply problem, pathway optimization, microbial tolerance, and efficient product extraction. The future research should focus on further optimization of flux through MEP or MVA pathways

to provide high supply of precursors and engineering terpene synthase enzymes to increase the production of desired terpenoids. Also, efforts should be made to improve microbial tolerance to high levels of terpenoid production and to develop suitable extraction methods of terpenoids, especially volatile ones, during production.

References

[1] Koehn FE, Carter GT. The evolving role of natural products in drug discovery. Nature Reviews Drug Discovery 2005;4:206–220.

[2] Lahlou M. The success of natural products in drug discovery. Pharmacology & Pharmacy 2013;4:17–31.

[3] Dias DA, Urban S, Roessner U. A Historical overview of natural products in drug discovery. Metabolites 2012;2:303–36.

[4] Harvey AL, Edrada-Ebel R, Quinn RJ. The re-emergence of natural products for drug discovery in the genomics era. Nature Reviews Drug Discovery 2015;14:111–29.

[5] Newman DJ, Cragg GM. Natural products as sources of new drugs over the 30 years from 1981 to 2010. Journal of Natural Products 2012;75:311–35.

[6] Wang G, Tang W, Bidigare RR. Terpenoids as therapeutic drugs and pharmaceutical agents. Natural products: Drug discovery and therapeutic medicine. In: Zhang L, Demain AL (Eds.). Totowa, NJ: Humana Press; 2005. p. 197–227.

[7] Thoppil RJ, Bishayee A. Terpenoids as potential chemopreventive and therapeutic agents in liver cancer. World Journal Hepatology 2011;3:228–249.

[8] Guan Z, Xue D, Abdallah II, Dijkshoorn L, Setroikromo R, Guiyuan L, et al. Metabolic engineering of *Bacillus subtilis* for terpenoid production. Applied Microbiology and Biotechnology 2015;99:9395–9406.

[9] Ajikumar PK, Tyo K, Carlsen S, Mucha O, Phon TH, Stephanopoulos G. Terpenoids: Opportunities for biosynthesis of natural product drugs using engineered microorganisms. Molecular Pharmaceutics 2008;5:167–190.

[10] Lange BM, Rujan T, Martin W, Croteau R. Isoprenoid biosynthesis: the evolution of two ancient and distinct pathways across genomes. Proceedings of the National Academy of Sciences of USA 2000;97:13172–13177.

[11] Dewick PM. The mevalonate and deoxyxylulose phosphate pathways: terpenoids and steroids. In: Medicinal natural products. John Wiley & Sons, Ltd; 2001. p. 167–289.

[12] Eisenreich W, Schwarz M, Cartayrade A, Arigoni D, Zenk MH, Bacher A. The deoxyxylulose phosphate pathway of terpenoid biosynthesis in plants and microorganisms. Chemistry and Biology 1998;5:R221–R233.

[13] Tholl D. Terpene synthases and the regulation, diversity and biological roles of terpene metabolism. Current Opinion in Plant Biology 2006;9:297–304.

[14] Pattanaik B, Lindberg P. Terpenoids and their biosynthesis in cyanobacteria. Life 2015;5:269–293.

[15] Loza-Tavera H. Monoterpenes in essential oils. Biosynthesis and properties. Advances in Experimental Medicine and Biology 1999;464:49–62.

[16] Banthorpe DV, Charlwood BV, Francis MJO. Biosynthesis of monoterpenes. Chemical Review 1972;72:115–155.

[17] Cordell GA. Biosynthesis of sesquiterpenes. Chemical Review 1976;76:425–460.

[18] Chadwick M, Trewin H, Gawthrop F, Wagstaff C. Sesquiterpenoids lactones: Benefits to plants and people. International Journal of Molecular Sciences 2013;14:12780–12805.

[19] Bhat SV, Sivakumar M, Nagasampagi BA. Chemistry of natural products. Berlin: Narosa; 2005.

[20] Tu Y. The discovery of artemisinin (qinghaosu) and gifts from Chinese medicine. Nature Medicine 2011;17:1217–1220.

[21] Lai HC, Singh NP, Sasaki T. Development of artemisinin compounds for cancer treatment. Investigational New Drugs 2013;31:230–246.

[22] Das AK. Anticancer effect of antimalarial artemisinin compounds. Annals of Medical and Health Sciences Research 2015;5:93–102.

[23] Lai H, Sasaki T, Singh NP. Targeted treatment of cancer with artemisinin and artemisinin-tagged iron-carrying compounds. Expert Opinion on Therapeutic Targets 2005;9:995–1007.

[24] Abdallah II, Czepnik M, van Merkerk R, Quax WJ. Insights into the three-dimensional structure of amorpha-4,11-diene synthase and probing of plasticity residues. Journal of Natural Product 2016;79:2455–2463.

[25] Priyadarshini K, Keerthi Aparajitha U. Paclitaxel against cancer: A short review Medical Chemistry 2012;2:139–141.

[26] Boghigian BA, Salas D, Ajikumar PK, Stephanopoulos G, Pfeifer BA. Analysis of heterologous taxadiene production in K- and B-derived Escherichia coli. Applied Microbiology and Biotechnology 2012;93:1651–1661.

[27] Hezari M, Croteau R. Taxol biosynthesis: An update. Planta Medica 1997;63:291–295.

[28] Gao Y, Honzatko RB, Peters RJ. Terpenoid synthase structures: A so far incomplete view of complex catalysis. Natural Product Reports 2012;29:1153–1175.

[29] Köksal M, Jin Y, Coates RM, Croteau R, Christianson DW. Taxadiene synthase structure and evolution of modular architecture in terpene biosynthesis. Nature 2011;469:116–120.

[30] Christianson DW. Structural biology and chemistry of the terpenoid cyclases. Chemical Reviews 2006;106:3412–3442.

[31] Noel JP, Dellas N, Faraldos JA, Zhao M, Hess BA, Smentek L, et al. Structural elucidation of cisoid and transoid cyclization pathways of a sesquiterpene synthase using 2-fluorofarnesyl diphosphates. ACS Chemical Biology 2010;5:377–392.

[32] Bloom JD, Meyer MM, Meinhold P, Otey CR, MacMillan D, Arnold FH. Evolving strategies for enzyme engineering. Current Opinion in Structural Biology 2005;15:447–452.

[33] Diaz JE, Lin CS, Kunishiro K, Feld BK, Avrantinis SK, Bronson J, Greaves J, Saven JG, Weiss GA. Computational design and selections for an engineered, thermostable terpene synthase. Protein Science 2011;20:1597–1606.

[34] Christianson DW. Unearthing the roots of the terpenome. Current Opinion in Chemical Biology 2008;12:141–150.

[35] Köksal M, Zimmer I, Schnitzler J-P, Christianson DW. Structure of isoprene synthase illuminates the chemical mechanism of teragram atmospheric carbon emission. Journal of Molecular Biology 2010;402:363–373.

[36] Hyatt DC, Youn B, Zhao Y, Santhamma B, Coates RM, Croteau RB, et al. Structure of limonene synthase, a simple model for terpenoid cyclase catalysis. Proceedings of National Academy of Sciences USA 2007;104:5360–5365.

[37] Kampranis SC, Ioannidis D, Purvis A, Mahrez W, Ninga E, Katerelos NA, et al. Rational conversion of substrate and product specificity in a Salvia monoterpene synthase: structural insights into the evolution of terpene synthase function. Plant Cell 2007;19:1994–2005.

[38] Whittington DA, Wise ML, Urbansky M, Coates RM, Croteau RB, Christianson DW. Bornyl diphosphate synthase: structure and strategy for carbocation manipulation by a terpenoid cyclase. Proceedings of National Academy of Sciences USA 2002;99:15375–15380.

[39] Miller DJ, Allemann RK. Sesquiterpene synthases: Passive catalysts or active players? Natural Product Reports 2012;29:60–71.

[40] Li JX, Fang X, Zhao Q, Ruan JX, Yang CQ, Wang LJ, et al. Rational engineering of plasticity residues of sesquiterpene synthases from *Artemisia annua*: Product specificity and catalytic efficiency. Biochemical Journal 2013;451:417–426.

[41] Klein-Marcuschamer D, Ajikumar PK, Stephanopoulos G. Engineering microbial cell factories for biosynthesis of isoprenoid molecules: beyond lycopene. Trends in Biotechnology 2007;25:417–424.

[42] Chang MCY, Keasling JD. Production of isoprenoid pharmaceuticals by engineered microbes. Nature Chemical Biology 2006;2:674–681.

[43] Chen X, Zhou L, Tian K, Kumar A, Singh S, Prior BA, Wang Z. Metabolic engineering of *Escherichia coli*: A sustainable industrial platform for bio-based chemical production. Biotechnology Advances 2013;31:1200–1223.

[44] Martin VJJ, Pitera DJ, Withers ST, Newman JD, Keasling JD. Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. Nature Biotechnology 2003;21:796–802.

[45] Kampranis SC, Makris AM. Developing a yeast cell factory for the production of terponoids. Computational and Structural Biotechnology Journal 2012;3:1–7.

[46] Xue D, Abdallah II, de Haan IEM, Sibbald MJJB, Quax WJ. Enhanced C30 carotenoid production in *Bacillus subtilis* by systematic overexpression of MEP pathway genes. Applied Microbiology and Biotechnology 2015;99:5907–5915.

[47] Zhou K, Zou R, Zhang C, Stephanopoulos G, Too HP. Optimization of amorphadiene synthesis in *Bacillus subtilis* via transcriptional, translational, and media modulation. Biotechnology and Bioengineering 2013;110:2556–2561.

[48] Yoshida K, Ueda S, Maeda I. Carotenoid production in *Bacillus subtilis* achieved by metabolic engineering. Biotechnology Letters 2009;31:1789–1793.

[49] Ducat DC, Way JC, Silver PA. Engineering cyanobacteria to generate high-value products. Trends Biotechnology 2011;29:95–103.