Terpenoid cell factory

Abdallah, Ingy Ibrahim Ahmed Fouad

10

Summary, Concluding Remarks and Future Perspectives

Abstract

Terpenoids constitute the largest, most diverse group of natural products with numerous ones garnering commercial and medicinal interest. Most terpenoids are naturally produced in low amounts so their purification suffers from low yields and consumption of large amounts of natural resources. Hence, the extinction of the natural source becomes a real threat and the amount of isolated compound usually cannot cover market demand especially for medicinally important compounds like anticancer paclitaxel and antimalarial artemisinin. Also, synthesis of most terpenoids is problematic and expensive due to the complexity of their structures. Thus, the need for alternate methods of terpenoid production is a pressing issue. Among these methods is the microbial production of terpenoids, which has been researched in the last few decades. The objectives of the thesis by Ingy Abdallah are establishing the Gram-positive bacterium *Bacillus subtilis* as a platform organism for terpenoid production by understanding and improving the biosynthetic routes of terpenoids including the core enzymes involved in the process such as terpene synthases. A *B. subtilis* strain overexpressing the terpenoid biosynthetic pathway was successfully engineered then used for the production of C_{30} carotenoids and the diterpene precursor of paclitaxel, taxadiene. This strain can serve as a cell factory for production of various terpenoids. In addition, the research in this thesis delved into the enzyme family of terpene synthases, which are essential for the formation of terpenoids. The focus was on understanding the structure-function relations of amorphadiene synthase, a key enzyme in artemisinin production, and improving its catalytic activity.

Nature provides a wealth of compounds with commercial and/or medicinal significance. Among these natural products, terpenoids are very famous. Terpenoids are considered one of the largest classes of natural products with vast structural and functional diversity. They are widespread amid various organisms where they are important for several vital functions. Numerous terpenoids are harvested from microbes and plants for countless applications ranging from food to cosmetics to pharmaceuticals or nutraceuticals. For example, carotenoids possessing antioxidant properties, volatile monoterpenes used as flavors and fragrances, anticancer paclitaxel and antimalarial artemisinin.

In spite of the huge diversity in the chemical structures of terpenoids, they are all synthesized by consecutive condensation of two five-carbon precursors, isopentenyl pyrophosphate (IPP) and its isomer dimethylallyl pyrophosphate (DMAPP). These precursors are synthesized via two distinct pathways, the mevalonate (MVA) or the 2-C-methyl-D-erythritol-4-phosphate (MEP) pathways. These building blocks are then fused through different number of repeats to produce the backbones of all terpenoids. The backbones are then cyclized and/or rearranged by terpene synthase enzymes to produce the corresponding terpenoids. Terpene synthases are a class of enzymes that are essential for the production of the desired terpenoids.

Most terpenoids are naturally produced in low concentrations so purification from their natural biological material suffers from low yields and consumption of large amounts of natural resources. This can lead to extinction of the original species and usually the amount of isolated compound cannot cover market demand especially for medicinally important compounds like paclitaxel and artemisinin. Also, organic synthesis of most terpenoids is problematic and costly due to the complexity of their structures. Hence, alternate methods for production of terpenoids have been researched in the last few decades. Among these methods is the microbial production of terpenoids. Microbial hosts such as *Escherichia coli*, *Saccharomyces cerevisiae* and *Bacillus subtilis* were studied for their ability to act as platform organisms for production of terpenoids. An understanding of the biosynthetic routes for production of terpenoids including the core enzymes involved in this process such as terpene synthases is crucial for engineering a host for terpenoid production, in addition, to the knowledge of the inherent metabolic pathways of the selected host. This allows for the transfer of all or parts of the terpenoid biosynthetic pathways to heterologous hosts in order to create a cell factory for production of important terpenoids.

Chapter 1 is an overview of the biosynthesis of terpenoids, their different classes, the terpene synthases enzyme family and some examples of engineering microbial cell factories for terpenoid production.

The research in this thesis aims at metabolic engineering of a microbial host as a cell factory for terpenoid production. *B. subtilis* was chosen as the platform organism due to its successful use for industrial production of proteins, in addition to the fact that it is the highest isoprene (C_5 terpenoid) producer among the well-known microorganism hosts. Moreover, the second part of this thesis focuses on studying the famous sesquiterpene synthase, amorphadiene synthase, which is responsible for production of the precursor of the antimalarial drug artemisinin.

Engineering *Bacillus subtilis* as a cell factory for terpenoid production

B. subtilis has been extensively used as a host for bulk industrial enzyme production. It has high potential to act as a high-grade cell factory. *B. subtilis* possesses an inherent MEP pathway producing high levels of isoprene. It has a wide metabolic potential, and varied substrate range. It has higher growth rate compared to yeast. Moreover, *B. subtilis* is listed by the Food and Drug Administration as generally regarded as safe (GRAS) which is an advantage over *E. coli*. Recently, researchers started focusing on engineering *B. subtilis* to produce valuable terpenoids in a manner similar to the research performed on the yeast *S. cerevisiae* and *E. coli* but taking into account the advantages of *B. subtilis* as a potentially better cell factory. In **Chapter 2**, we review the inherent terpenoid biosynthetic pathways of *B. subtilis* along with the challenges facing the engineering of *B. subtilis* for production of valuable terpenoids, major advances and future directions for using *B. subtilis* as a cell factory.

Our efforts to engineer *B. subtilis* for terpenoid production start in **Chapter 3** where we overexpress different enzymes involved in the MEP pathway to boost the production of the precursors IPP and DMAPP required for biosynthesis of terpenoids. We describe the systematic overexpression of the enzymes from the *B. subtilis* MEP pathway and we use the level of production of orange colored C_{30} carotenoids, which were previously successfully produced in *B. subtilis*, as a read out to evaluate the effect of such modulations on terpenoid production. We assembled the endogenous genes of *B. subtilis* encoding the MEP pathway enzymes one by one in a synthetic operon that is expressed from a plasmid. The operon was constructed to contain up to four genes based on their significance in previous literature reports. The successful expression of the enzymes from a plasmid based system was demonstrated. After that, the *B. subtilis* strains overexpressing different sets of enzymes of the MEP pathway were combined with a C_{30} carotenoid producing plasmid. The orange colored carotenoids, 4,4'-diaponeurosporene and 4,4'-diapolycopene, were extracted from the different strains and quantified by HPLC. It was shown that each consecutive expression of an additional enzyme involved in the MEP pathway resulted in a higher amount of carotenoids detected. The two strains overexpressing four of the MEP pathway

enzymes showed the highest amount of carotenoid production compared to the other strains. The strain overexpressing *ispC*, *ispE*, *ispG*, and *ispA* showed slightly higher amount (10.65 mg/g dcw) compared to the strain overexpressing *dxs*, *ispD*, *ispF*, and *ispH* (9.03 mg/g dcw). Figure 1 shows the upper MEP pathway producing IPP and DMAPP and the lower pathway for conversion of these precursors to the terpenoid backbones, which in turn are transformed by the terpene synthases to the desired terpenoids.

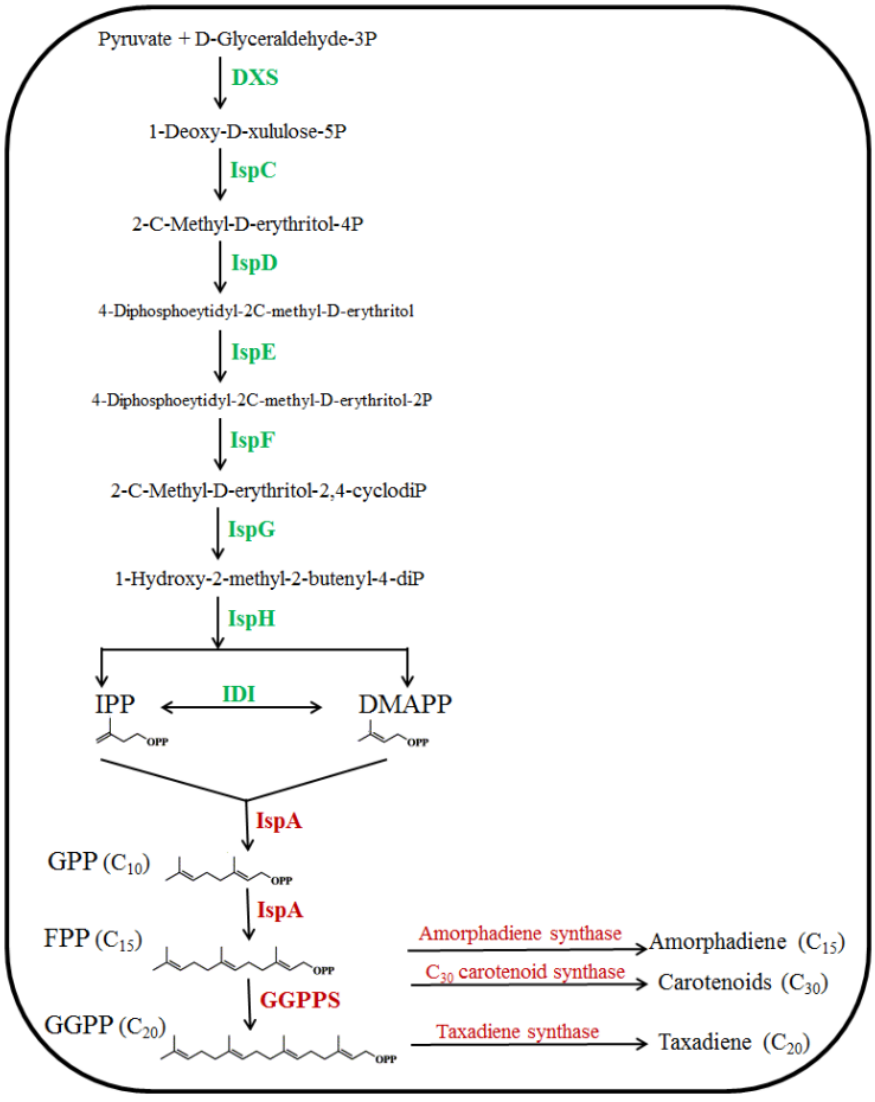

Figure 1. Biosynthesis of terpenoids in *B. subtilis* via the MEP pathway.

As a continuation of the research into the modulation of the MEP pathway of *B. subtilis*, in **Chapter 4**, we compare expression vectors with different origins of replication and promoters to choose the most stable expression system in *B. subtilis*. In addition, all the eight genes of the MEP pathway were combined in a single operon in a plasmid based expression system. The theta-replicating plasmid (pHCMC04G) was proven to be more structurally and segregationally stable compared to the rolling circle replication plasmid (pHB201) especially for large size operons containing three or more genes. Hence, the pHCMC04G vector with its xylose inducible promoter was used as the host of the operon containing the eight genes. RT-qPCR was used to prove that the transcripts are intact and all genes are expressed at the same level irrespective of their position in the operon. Finally, the amount of C_{30} carotenoids produced confirmed that the pHCMC04G plasmid is superior compared to pHB201. The *B. subtilis* strain overexpressing all eight genes of the MEP pathway (p04_SDFHCEGA) showed the highest carotenoid production (21 mg/g dcw) which is nearly twice the amount produced by the strains overexpressing only four genes.

In **Chapter 5**, we make use of the best *B. subtilis* strain generated from the previous two chapters (p04_SDFHCEGA) to aim at the production of taxa-4,11-diene which is the precursor of the anticancer paclitaxel. In the effort towards engineering *B. subtilis* as a host for the production of paclitaxel, we expressed the plant-derived taxadiene synthase (TXS) enzyme, which is responsible for the conversion of the precursor geranylgeranyl pyrophosphate (GGPP) to taxa-4,11-diene. This is the first committed intermediate in paclitaxel biosynthesis. The *txs* gene was integrated into the genome of *B. subtilis* and the TXS enzyme was successfully expressed for the first time in *B. subtilis*. Additionally, genes encoding biosynthetic enzymes from the MEP pathway [*dxs*, *ispD*, *ispF*, *ispH*, *ispC*, *ispE*, *ispG*] and early terpenoid biosynthesis [*ispA* encoding IspA that performs the function of geranyl pyrophosphate synthase (GPPS) and farnesyl pyrophosphate synthase (FPPS), *crtE* encoding geranylgeranyl pyrophosphate synthase (GGPPS)] were cloned into *B. subtilis* and their effect on taxadiene production was evaluated. The over expression of the MEP pathway enzymes along with IspA and GGPPS caused 83-fold increase in the amount of taxadiene produced compared to the *B. subtilis* strain only expressing TXS and relying on the innate pathway of *B. subtilis*. The total amount of taxadiene produced by that strain was 17.2 mg/L which is higher than the amounts reported for *E. coli* and *S. cerevisiae* on shake flask fermentation level.

Study of terpene synthases with emphasis on amorphadiene synthase

The participation of a terpene synthase is an indispensable prerequisite for the synthesis of terpenoids. Terpene synthases constitute a family of enzymes tasked with catalyzing the cyclization and/or rearrangement of the precursors geranyl pyrophosphate (GPP), farnesyl pyrophosphate (FPP), and geranylgeranyl pyrophosphate (GGPP) to create the different classes of terpenoids. The mechanism of terpene synthases is complex where the diverse array of terpenoid structures depend on the orientation of the flexible substrate in their large active site. Terpene synthases are divided into different classes to produce the numerous terpenoids. We focus on the class of sesquiterpene synthases, which are responsible for converting the substrate FPP to more than 300 sesquiterpenes (C_{15}). They are classified into transoid sesquiterpene synthases that catalyze the ionization of the (*trans*, *trans*)-substrate and cisoid ones which execute an initial double bond isomerization to create a (*cis*, *trans*)-intermediate carbocation. All sesquiterpene synthases possess a tri-nuclear metal cluster liganded by conserved metal ion binding motifs DDXXD and (N,D)DXX(S,T)XXXE that prompt the ionization of the pyrophosphate group of the substrate to initiate the catalytic mechanism.

Among the most prominent sesquiterpene synthases is amorphadiene synthase (ADS). ADS catalyzes the cyclization of FPP to amorpha-4,11-diene, which is considered the first and rate limiting step in the synthesis of the antimalarial drug artemisinin. Artemisinin-based combination therapies (ACTs) are recommended as the first-line treatment of malaria by The World Health Organization (WHO). Also, artemisinin has recently been gaining attention for its possible chemotherapeutic effect in fighting cancers. Understanding the biosynthesis of artemisinin will help optimize its production and increase its availability at lower price. Hence, we focused on studying the enzyme amorphadiene synthase to gain better insight into the structure-function relations within the catalytic mechanism and to aim at improving its activity.

Amorphadiene synthase is a cisoid sesquiterpene synthase. In spite of the numerous research into this enzyme, obtaining an X-ray crystal structure has not been successful. Thus, the structural basis of its catalytic mechanism is still unknown, as the molecular conformation of the enzyme has not been characterized. In **Chapter 6**, we created a reliable three-dimensional (3D) homology model representing the conformation of the ADS enzyme using Discovery Studio software. The model was produced using the crystal structure of α-bisabolol synthase, an enzyme with high sequence identity with ADS, as a template. Three magnesium ions representing the tri-nuclear metal cluster were positioned in the active site and confirmed by their chelation with the metal ion binding motifs. Subsequently, the substrate FPP was docked into the active site and the

different conformations produced were evaluated for their folding and interactions leading to the selection of one conformation as the best fit for the substrate in the active site . The 3D model was validated by assessing its stereochemical quality in addition to probing some of the active site residues through site-directed mutagenesis. The generated variants confirmed the validity of the ADS model.

The generated ADS model was used in **Chapter 7** to select active site residues for creating a mutability landscape of ADS. A mutant library of 257 variants encompassing the sixteen residues: R262, R440, Q518, H392, H448, L515, K449, V396, F525, Y519, W271, T296, T399, G400, G439 and D523 was produced. This mutant library was screened for catalytic activity and product profile to determine the residues involved in the mechanism of ADS and their respective roles. This screening proved that aromatic residues (W271, Y519 and F525) are essential for stabilization of the reactive carbocations in the active site. The basic histidines, 392 and 448, were excluded as the active site catalytic base while W271 proved to be a possible candidate for this function. The role of T399 in regioselective deprotonation was corroborated. Moreover, R262 creates a region of high positive charge along with the magnesium ions in the active site to neutralize the negative charge of the released pyrophosphate preventing it from interfering with the reactive carbocations. R440 and K449 were excluded from this function. Also, the landscape helped select variants with improved catalytic activity compared to wild type ADS. Several variants having the same product profile as wild type ADS showed improved catalytic activity. The H448A variant displayed high improvement with almost four times increase in catalytic efficiency compared to wild type ADS. The double mutant T399S/H448A showed turnover rate (k_{cat}) five folds higher than the wild type. Figure 2 summarizes the research in chapters 6 and 7 starting with creating the ADS model then generating the mutant library followed by screening the mutability landscape for product profile and catalytic activity.

Finally, **Chapter 8** focused on evaluating the impact of mutating a single active site residue on the promiscuity of amorphadiene synthase. The major product of ADS is amorpha-4,11-diene along with several minor products such as β-farnesene, γ-humulene, α-bisabolol, amorpha-4,7-diene and amorpha-4-en-11(7)-ol. Using the ADS mutability landscape for product profile generated in Chapter 7, we identified variants that increase the production of one or more of these minor products at expense of the major amorpha-4,11-diene. Six plasticity residues were probed, namely V396, T399, G400, H448, L515 and D523. Mutation of residue L515 reduced product specificity unless conserved as isoleucine, valine or proline. Residue G400 must be conserved to maintain activity and product specificity. Replacing V396 with acidic residues impaired the regioselective deprotonation as well as substituting T399 with amino acids other

than serine. Also, the two variants, H448P and D523A, greatly diminished the product specificity of ADS by producing amorpha-4-en-7(11)-ol as the major product.

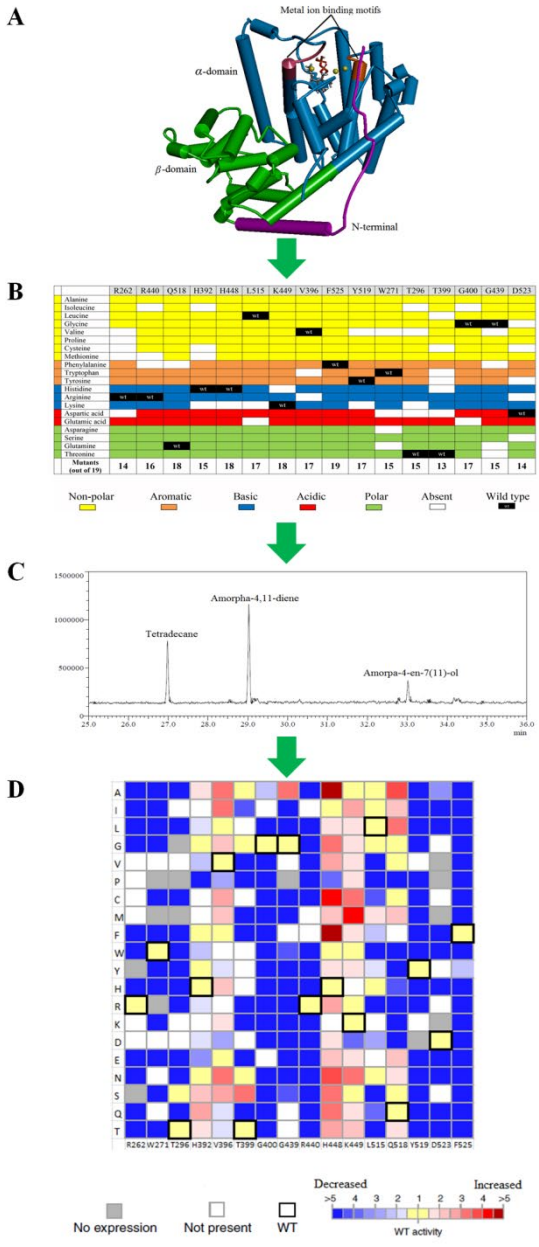

Figure 2. Summary of the research in chapters 6 and 7.

(A) The generated ADS 3D model. **(B)** The ADS mutant library consisting of 257 variants. **(C)** GC chromatogram showing the product profile of wild type ADS. **(D)** Mutability landscape of ADS for catalytic activity.

Conclusions and future perspectives

The overall aim of this thesis is to create a sustainable terpenoid cell factory. To achieve this goal, we focused on two main areas. The first part of the thesis researched the engineering of *B. subtilis* as a platform organism for terpenoid production. We started by modulating the inherent biosynthetic MEP pathway in *B. subtilis* required for production of the terpenoid precursors and we evaluated its effect on production of C_{30} carotenoids. We successfully generated a strain stably overexpressing all eight genes of the MEP pathway from a plasmid based system in *B. subtilis*. That strain produced a high amount of C_{30} carotenoids. Moreover, the gene for the enzyme taxadiene synthase was incorporated in the genome of the *B. subtilis* strain overexpressing the full MEP pathway leading to the production of high levels of taxa-4,11-diene, the precursor of paclitaxel, superior to *E. coli* and *S. cerevisiae*. The second part of our work delved into the terpene synthase enzymes essential for synthesis of terpenoids. We focused on studying the enzyme amorphadiene synthase which is important due to its role in the synthesis of artemisinin. We successfully created a 3D model of the enzyme and used it to build a mutability landscape that helped understand the role of some active site residues in the mechanism of ADS and identify variants with improved activity compared to the wild type.

In the future, the superior *B. subtilis* created in this research can be used as a cell factory for production of any number of terpenoids where the terpene synthase of the desired terpenoid can be introduced into its genome similar to our work with taxadiene synthase. In addition, further improvement to the production levels can be achieved by optimizing the growth medium and/or incorporating the mevalonate biosynthetic pathway to further boost the production of terpenoid precursors. Scaling up to large scale industrial fermenters along with further tuning of growth conditions should be performed to optimize the *B. subtilis* strain producing taxadiene. This cell factory can be used for industrial production of paclitaxel by combining it with the remaining enzymes in their biosynthetic pathway and optional chemical conversions. Based on the advantages of *B. subtilis* compared to other cell factories, this can provide a path for a more efficient approach for production of semisynthetic paclitaxel or artemisinin. Moreover, the ADS variant with improved catalytic activity can be cloned into any microbial host to achieve better production of amorpha-4,11-diene to what is known now using wild type ADS. This is the stepping stone for improving production of artemisinin.


4
5
3
2
6
1
7
8
9
Hoofdstukken

4
6
3
7
1
2
5
9
8
Samenvatting