Journal of Research in Pharmacy

Research Article

www.jrespharm.com

Heterologous expression of enzymes involved in artemisinin biosynthesis via methylerythritol phosphate pathway from *Artemisia annua* in *Escherichia coli*

Tresna LESTARI [1,4], Catur RIANI [2,3] *, Elfahmi ELFAHMI [1,3], Asep Gana SUGANDA [1]

1 Department of Pharmaceutical Biology, School of Pharmacy, Institut Teknologi Bandung, Bandung, Indonesia.
2 Department of Pharmaceutics, School of Pharmacy, Institut Teknologi Bandung, Bandung, Indonesia.
3 Biosciences and Biotechnology Research Center, Institut Teknologi Bandung, Indonesia
4 Faculty of Pharmacy, Universitas Bakti Tunas Husada, Tasikmalaya, Indonesia.
* Corresponding Author. E-mail: catur@itb.ac.id (C.R.); Tel. +62-22 2504852.

Received: 27 February 2022 / Revised: 10 June 2022 / Accepted: 10 June 2022

ABSTRACT: Artemisinin combination therapies (ACTs) have become the mainstay of treatment for malaria worldwide. This has led to a high demand for artemisinin precursors as starting materials for artemisinin production and their semi-synthetic derivatives. In this study, heterologous expression of enzymes involved in artemisinin biosynthesis was performed in *Escherichia coli* to produce artemisinin precursors, i.e., amorphadiene, artemisinic acid, and dihydroartemisinic acid. These enzymes are farnesyl pyrophosphate synthase (FPS), amorpha-4,11-diene synthase (ADS), cytochrome P450 monooxygenase (CYP71AV1/CYP), artemisinic aldehyde delta-11(13) reductase (DBR2), and aldehyde dehydrogenase 1 (ALDH1). Overexpression of the heterologous 1-deoxy-D-xylulose 5-phosphate (DXP) synthase gene (dxs) from *Bacillus subtilis* and the native isopentenyl diphosphate delta isomerase (IDI) gene (idi) from *E. coli* was also performed to enhance isopentenyl diphosphate (IPP) and dimethylallyl pyrophosphate (DMAPP) via the methylerithritol phosphate (MEP) pathway in *E. coli*. All genes were cloned into three plasmids. Gene expression was performed under isopropyl-β-D-1-thiogalactopyranoside (IPTG) induction and characterized by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and Western blot. The number of IPTG, incubation temperature, length of incubation time, and culture medium were optimized to find the best conditions for protein expression. It is found that all proteins could be expressed under 0.5 mM IPTG induction at an incubation temperature of 20 °C for 24 hours in Luria-Bertani medium. The cloning of seven key enzymes involved in artemisinin biosynthesis in *E. coli* has never been described in other studies. By further optimizing the fermentation process, this research offers a promising future for the production of artemisinin precursors in *E. coli*.

KEYWORDS: heterologous expression; MEP pathway; artemisinin; *Escherichia coli*.

1. INTRODUCTION

Malaria is still considered as a global health problem, particularly in tropical areas like Africa and Southeast Asia. Based on World Health Organization (WHO) data in 2020, a year after Covid-19 pandemic, global malaria cases have increased up to 241 million compared to the number of cases in 2019 which amounted to 227 million cases. Countries in Africa are still the highest contributor to malaria cases in the world, accounting for 95% of the total case. About 228 million cases of malaria occurred in Africa in 2020 and 602 thousand of them caused death. Southeast Asia contributed about 2% of malarial cases worldwide from nine malaria-endemic countries. India is the main contributor with around 83% from 5.0 million of total cases in 2020. As many as 9000 cases of all cases in Southeast Asia cause death [1].

Treatment with artemisinin-based combination therapy (ACTs) to combat malaria is still the first-line therapy recommended by WHO. Artemisinin, a sesquiterpene lactone derived from *Artemisia annua*, is an outstanding antimalarial agent. It is discovered by Youyou Tu in 1972, that brought her to the 2015 Nobel Prize co-receipt in Physiology or Medicine [2]. Artemisinin content from various species of *Artemisia* reported presence at varied level from 0.07% to 0.45%. *A. vachanica* was found to be a novel plant source of artemisinin (0.34%) [3]. Other studies showed that the highest content of artemisinin was found in *A. annua* (up to 1.4%) [4], still a relatively small amount to meet artemisinin needs.

How to cite this article: Lestari T, Riani C, Elfahmi E, Suganda AG. Heterologous expression of enzymes involved in artemisinin biosynthesis via methylerythritol phosphate pathway from Artemisia annua in Escherichia coli. J Res Pharm. 2022; 26(6): 1656-1664.

ISSN: 2630-6344

http://dx.doi.org/10.29228/jrp.256

Artemisinin synthesized in *A. annua*'s trichomes is derived from isopentenyl pyrophosphate (IPP) via mevalonic pathway (MVA) and dimethylallyl pyrophosphate (DMAPP) via non-mevalonic pathway (MEP) [5,6]. Biotechnological approaches such as genetic engineering and metabolic pathway modification of microbes is a promising alternative for producing artemisinin semi-synthetically. *E. coli* is the most common microorganism used for biotechnological manufacture of natural medicinal compounds because of its commercial advantages [7]. Genetic modification of *E. coli* for overexpression of both MVA and MEP pathway increased isoprene yield about 3-fold and 20-fold compared to overexpression of MVA and MEP alone [8]. Previous study reported that introducing heterologous *dxs* from *B. subtilis* doubled isoprenoids production in *E. coli* [9]. When *dxs* was co-expressed with *idi* and *ADS* in *B. subtilis*, 2 mg/L amorphadiene was generated [8]. By combining genes involved in the MVA pathway from *Saccharomyces cerevisiae* to recombinant *E. coli* carrying *ADS* from *A. annua*, amorphadiene production was enhanced to 24 µg/ml [10]. Modification of bioreactor and fermentation condition by restricting glucose level and ammonia has been shown to increase amorphadiene level up to 0.5 g/L and 25 g/L [11-13].

In this study, the DNA sequence encoding DXS from *B. subtilis* and IDI from *E. coli* were co-expressed to improve IPP/DMAPP production through MEP pathway in *E. coli.* Additionally, native *ADS* from *A. annua* and optimized *FPS, CYP, DBR2* and *ALDH1* genes were co-expressed to promote the synthesis of artemisinin precursor farnesyl pyrophosphate, amorphadiene, artemisinic acid, dihydroartemisinic aldehyde and dihydroartemisininc acid. Up to the best of our knowledge, increasing MEP pathway in *E. coli* combined with overexpression of genes encoding enzymes related to artemisinin biosynthesis from *A. annua* to stimulate the formation of late artemisinin precursors, artemisinic acid and/or dihydroartemisinic acid in *E. coli* has never been reported before.

2. RESULTS

2.1. Plasmids Construction and Characterization

All of the seven genes involved in this study were constructed in three plasmids resulting pETDfac plasmid that carries *FPS, ADS* and *CYP*, pRSFDda plasmid that carries *DBR2* and *ALDH1*, and pCDFDdi plasmid that carries *dxs* and *idi*. All of the plasmids carry lac inducible promoter to control genes expression. Polyhistidine coding sequence was added either at 5'-end or 3'-end of the genes harboring fusion expression with the protein.

The *ADS* and *FPS* were constructed at multiple cloning site (MCS)-1 of pETDUET-1 vector using *Bam*HI/*Not*I restriction sites to insert *ADS* and *Not*I/*Not*I to insert *FPS*. Both genes were regulated under the same promoter but carries each ribosome binding site and translation start/termination sites resulting polycistronic operon structure. Meanwhile, *CYP* was constructed at MCS-2 of pETDUET-1 using *Nde*I/*Xho*I restriction sites as monocistronic operon (Figure 1A). The *DBR2* and *ALDH1* were constructed as monocistronic operon at MCS-1 and MCS-2 of pRSFDUET-1 respectively (Figure 1B) as well as *dxs* and *idi* in pCDFDUET-1 (Figure 1C). The *DBR2* and *dxs* were inserted between *Bam*HI/*Eco*RI restriction sites, while *ALDH1* and *idi* were inserted between *Nde*I/*Xho*I restriction sites in each plasmid.

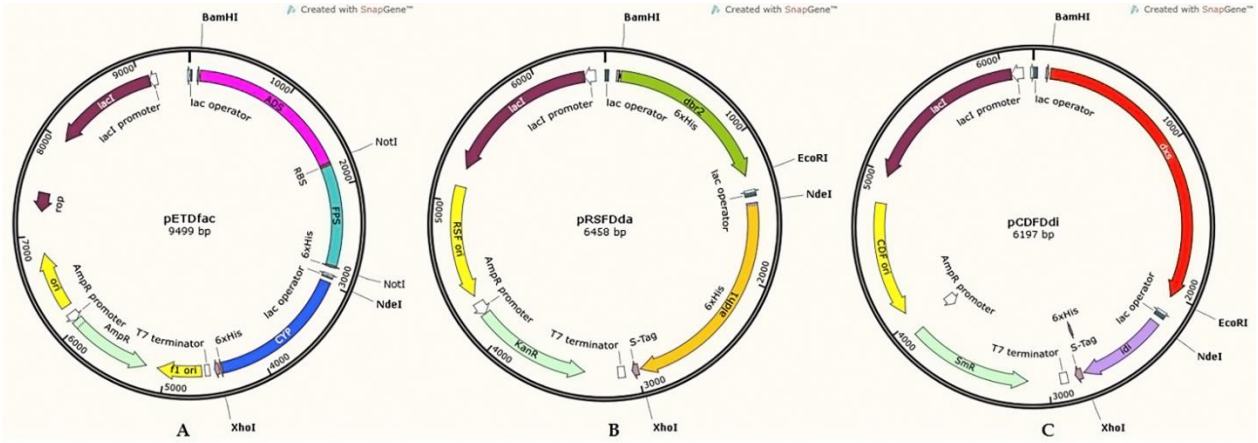

Figure 1. Plasmids map. **A**. pETDfac (9499 bp): pETDUET-1 vector carrying *FPS, ADS, CYP*, **B**. pRSFDda (6458 bp): pRSFDuet-1 vector carrying *DBR2* and *ALDH1*, **C**. pCDFDdi (6197 bp): pCDFDuet-1 vector carrying *dxs* and *idi* (Ilustrated by SnapGene).

Recombinant plasmids characterization resulted that all genes were successfully cloned into plasmids. Migration analysis showed that pETDfac (9499 bp) migrated slower compared to pRSFDda (6458 bp) and pCDFDdi (6197 bp) due to its bigger size. Meanwhile, the smallest plasmid, pCDFDdi migrated faster than the others but only slightly different with pRSFDdi (Figure 2A). Based on single digestion analysis using *Nde*I restriction enzyme, all three recombinant plasmids were seen with DNA bands of the appropriate size (Figure 2B). Consistent with migration analysis, pRSFDda and pCDFDdi were found hard to be differentiated. They both showed bands of similar size that make them difficult to be distinguished on analysis from cells carrying three recombinant plasmids (Figure 2B. Lane 2 and 5). Double digestion analysis showed DNA band of the inserted genes on each plasmid. Digestion of pETDfac using *Bam*HI/*Not*I restriction enzymes cleaved *ADS* (1641 bp) and *FPS* (1072 bp), while digestion using *Nde*I/*Xho*I cleaved *CYP* (1467) from pETDfac (Figure 2C Lane 1 and 2). Digestion of pRSFDda using *Bam*HI/*Eco*RI and *Nde*I/*Xho*I restriction enzymes realeased *DBR2* (1164 bp) and *ALDH1* (1500 bp) from the vectors, respectively (Figure 2C Lane 3 and 4). Characterization of pCDFDdi was carried out using the same restriction enzyme as in pRSFDda. Digestion using *Bam*HI/*Eco*RI resulting *dxs* (1902 bp) and digestion with *Nde*/*Xho*I resulting *idi* (549 bp) (Figure 2C Lane 5 and 6).

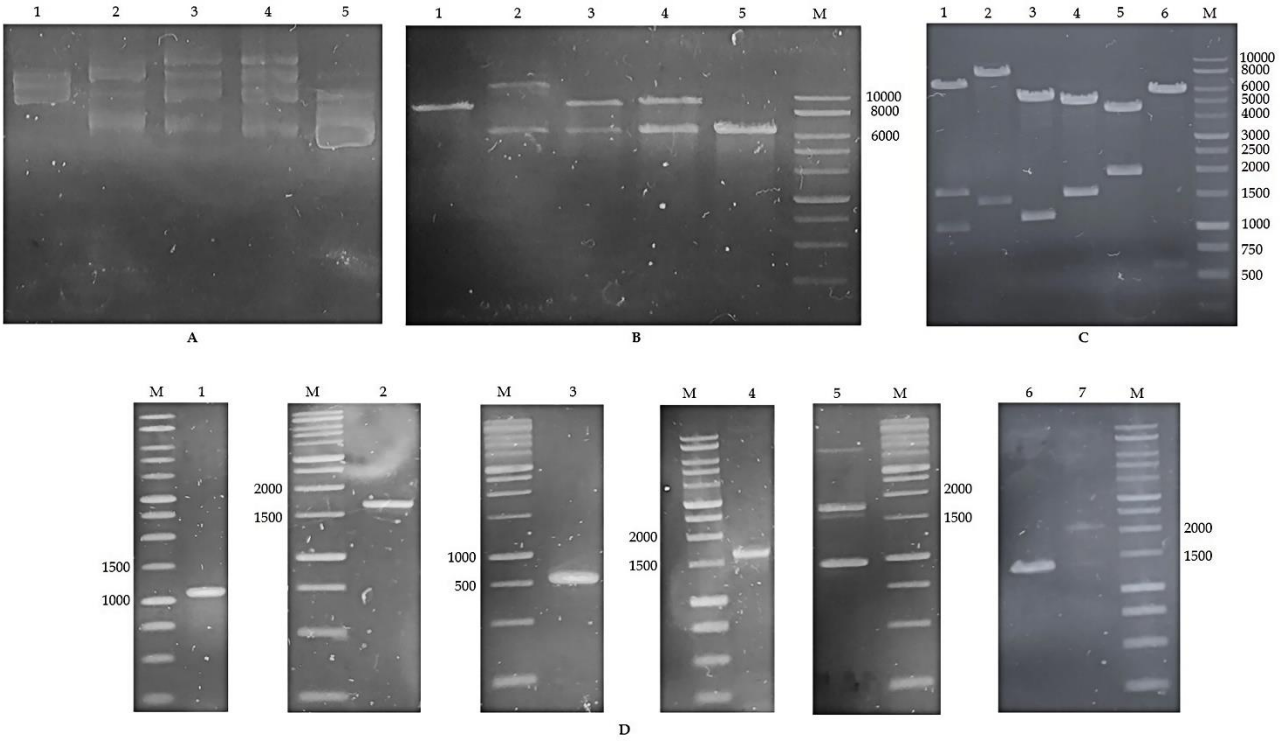

Figure 2. Recombinant plasmid characterizations. **A**. Migration analysis of plasmid: 1. pETDfac (9499 bp); 2. pRSFDda (6458 bp); 3. pETDfac and pRSFDda; 4. pETDfac, pRSFDda and pCDFDdi; 5. pCDFDdi (6197 bp), **B**. Single digestion (*Nde*I) of plasmid: 1. pETDfac (9499 bp); 2. pRSFDda (6458 bp); 3. pETDfac and pRSFDda; 4. pETDfac, pRSFDda and pCDFDdi; 5. pCDFDdi (6197 bp), **C**. Double digestion of plasmid: 1. pETDfac (*Bam*HI/*Not*I) resulting *ADS* (1641 bp) and*FPS* (1076 bp); 2. pETDfac (*Nde*I/*Xho*I) resulting *CYP* (1467 bp); 3. pRSFDda (*Bam*HI/*Eco*RI) resulting *DBR2* (1164 bp); 4. pRSFDda (*Nde*I/*Xho*I) resulting *ALDH1* (1500 bp); 5. pCDFDdi (*Bam*HI/*Eco*RI) resulting *dxs* (1902 bp); 6. pCDFDdi (*Nde*I/*Xho*I) resulting *idi* (549 bp), **D**. PCR analysis of plasmid: 1. *FPS* (1076 bp); 2. *ADS* (1641 bp); 3. *idi* (549 bp); 4. *CYP* (1467 bp); 5. *ALDH1* (1500 bp); 6. *DBR2* (1164 bp); 7. *dxs* (1902 bp); M. DNA ladder 1 kb.

In polymerase chain reaction (PCR) analysis, all the inserted DNA can be amplified. PCR analysis of *FPS* and *ADS* resulted in products that matched the theoretical size of *ADS* and *FPS* because the used primers specifically recognized each terminal of the genes (Figure 2 Lane 1 and 2). In contrast to *FPS* and *ADS*, PCR analysis on *CYP, DBR2, ALDH1, dxs* and *idi* were conducted using universal primers that anneals to each end of MCS of plasmid resulted addition about 200 bp higher than the actual DNA size (Figure 2 Lane 4, 6 and 7). Unspecific binding occurred in the PCR analysis of *ALDH1* (Figure 2 Lane 5), and this still appeared even after optimization of the annealing temperature (data not shown). In general, sequencing

analysis of recombinant plasmids (data not shown) showed that all the genes encoding enzymes involved in artemisinin biosynthesis through MEP pathway have successfully cloned and transformed into *E. coli*.

2.2. Genes Expression and Proteins Characterization

Protein expression analysis was first carried out on *E. coli* Bl21(DE3) carrying three recombinant plasmids induced by 1 mM of IPTG compared to *E. coli* carrying each recombinant plasmid induced by 0.5 mM of IPTG. The results of SDS-PAGE showed that all recombinant proteins (FPS, ADS, CYP, DBR2, ALDH1, DXS and IDI) were expressed. The protein bands obtained corresponding to their theoretical size DXS (71.18 kDa), ADS (65.59 kDa), ALDH1 (54.63 kDa), CYP (54.70 kDa), DBR2 (44.06 kDa), FPS (40.34 kDa) and IDI (21.33 kDa) (Figure 3A. Lane 5, indicated by arrow). The amount of IPTG, time of incubation, temperature and medium culture were further optimized against *E. coli* carrying three recombinant plasmids. Those proteins can be induced by 0.5 mM IPTG with an incubation time of 4 h at 37 °C. Protein expression continued to increase at 6 h and remained expressed until 24 h. However, at the 48 and 72 h of observation, protein expression was seen to decrease. Optimization of temperature of incubation optimization gave better result for soluble protein production at 20 °C compared to 37 °C. While medium optimization gave no significant increment of soluble protein in the use of M9 media compared to Luria-Bertani (LB). All the optimization datas were provided as supplementary material.

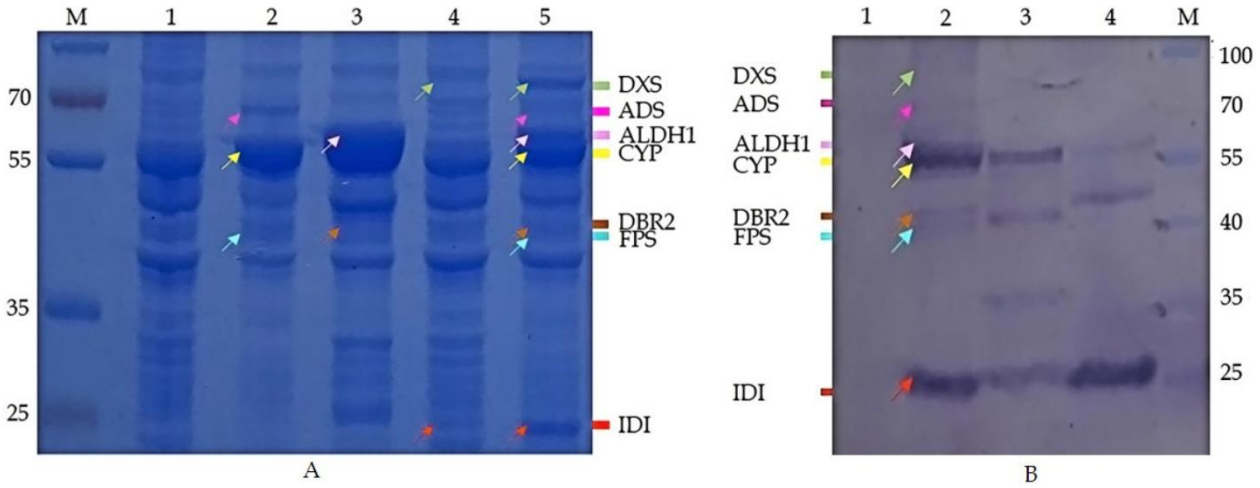

Figure 3. Protein characterizations. **A**. Protein overproduction: 1. *E. coli* BL21(DE3) wild type; 2. *E. coli* BL21(DE3)+pETDfac; 3. *E. coli* BL21(DE3)+pRSFDda; 4. *E. coli* BL21(DE3)+pCDFDdi; 5. *E. coli* BL21(DE3)+3plasmids; **B**. SDS-PAGE analysis of protein from *E. coli* BL21(DE3)+3plasmids; 1. Whole cellular protein from *E. coli* BL21(DE3) wild type; 2. Whole cellular protein from *E. coli* BL21(DE3)+3plasmids; 3. Inclussion bodies of protein from *E. coli* BL21(DE3)+3plasmids, 4. Soluble protein form *E. coli* BL21(DE3)+3plasmids (Arrow with correspond color indicates target protein).

Western blot (WB) analysis was conducted against whole cellular protein in cell lysate, inclusion body protein in debris cell and soluble protein in supernatant. The result reveals that all target protein from pellet cell could be detected in whole cellular protein (Figure 3C Lane 2). IDI was mostly found as soluble protein, followed by CYP and DBR2, which were seen as a small amount of soluble protein (Figure 3C Lane 4). FPS and some of CYP and IDI were observed as inclusion body proteins (Figure 3C Lane 3). Meanwhile, DXS, ADS and ALDH1 could not be identified by WB neither in inclusion body protein nor in soluble protein.

3. DISCUSSION

To date, engineering the MVA pathway from *S. cerevisiae* to *E. coli* successfully produced ADS up to 25 g/L [10,11]. Addition of *dxs* from *Streptomyces avermitilis* and *idi* from *B. subtilis* for the modification of the MEP pathway in *E. coli* succeeded in producing amorphadiene with a yield of 6.1 g/L [14]. *E. coli* has been shown to produce terpenoids through the MEP pathway with higher yields than the MVA pathway [15,16]. In this study an attempt was made to produce further artemisinin precursors, artemisinic acid and dihydroartemisinic acid, which had never been performed in previous studies. The approach was carried out through engineering the MEP pathway in *E. coli* by adding synthetic *dxs* and *idi* to increase the number

of IPP and DMAPP. This study also added *FPS, ADS, CYP*, *DBR2* and *ALDH1* which play a role in the biosynthesis of artemisinin.

Genes construction was conducted in three different vectors selected from different incompatibility groups. This is important to maintain plasmid persistence so the cells can continue to express the expected protein. Since the vectors have two MCS, all genes were designed as monocistronic operons, except *FPS* and *ADS*. Both genes, *FPS* and *ADS*, encode enzymes that function sequentially to produce amorphadiene [6].

The *dxs* used in this research was derived from *B. subtilis*, that was reported more efficient to enhance isoprene production in *E. coli* than the native one, while the *idi* gene was originated from *E. coli*. The *dxs* and *idi* genes were reported to regulate the presence of farnesyl diphosphate in *E. coli* via the MEP pathway [17]. DXS has been documented to exert strong flux and concentration control over MEP pathway intermediates [18]. The expression of *dxs* together with *idi* is expected to supply abundant isoprenoids by MEP pathway, and additional enzymes from *A. annua* supposed to promote the next steps conversion to produce other artemisinin intermediate compounds, i.e., farnesyl pyrophosphate, amorphadiene, artemisinic acid and/or dihydroartemisinic acid in *E. coli*.

In this research, *FPS*, *CYP*, *DBR2*, and *ALDH1* nucleotide sequence from *A. annua* were optimized for expression in *E. coli*. This conducted to reduce codon bias problem that might result in either an unexpressed or a truncated recombinant protein. Rare codon in heterologous condition was suggested to force the host to express a protein using random tRNA. The codon optimization was carried out in such a way to avoid errors that can cause translational stalling, premature translation termination, translation frame shift and amino acid mis-incorporation [19-21]. By expressing optimized genes, it is expected to shift the limitation and increase the production of functional enzymes. Beside the *FPS*, *CYP*, *DBR2* and *ALDH1*, the *ADS* gene used in this research was native sequence derived from *A. annua*. Previous work showed that native *ADS* expression in *E. coli* produced amorphadiene synthase wich was active in cyclase test using farnesyl diphosphate (FDP) as substrate [22].

Combination of *FPS, ADS, CYP, DBR2* and *ALDH1* was aimed to produce artemisininc acid and/or dihydroartemisininc acid, a late-stage precursor for artemisinin biosynthesis that was never reported before in *E. coli*. Sanofi created a commercial approach that used modified yeast through MEV pathway modulation developed by Jay D. Keasling to produce artemisinic acid [23,24], which was subsequently converted to dihydroartemisinic acid by diastereoselective reduction using diimide [25]. The production of dihydroartemisinc acid can shorten the semisynthetic route for artemisinin manufacturing. Dihydroartemisinic acid was the starting point for numerous artemisinin synthesis processes [26]. It was able to be converted to artemisinin through photooxidation followed by Hock cleavage reaction and triplet oxygen oxidation subsequently [27].

Protein overproduction was conducted under various IPTG induction. Our finding showed that all proteins could be expressed with no significant differences at the induction of 0.5, 1 and 2 mM IPTG with an incubation time of 4 h (Figure S2.A). In the research process, it was found that DXS from *B. subtilis* was well-expressed in *E. coli* at 0.1 mM of IPTG induction (data not shown) but the level was decreased after co-expression with IDI, although the inducer amount was upgraded ten times (Figure S1). After co-transformation of three plasmids, the number of DXS and IDI increased again to higher level (Figure 3A. Lane 4 and 5). The presence of the recombinant synthase was thought to be able to encourage the production of both enzymes, thus the production of IPP and DMAPP in the cell could also be stimulated.

Time optimization at 0.5 mM IPTG showed that the highest protein production was observed at 6 h and decreased after 24 h. Based on protein band intensity measurement using ImageJ software [28], ALDH1 was accounted for 8%, 10%, 7% at 4, 6 and 24 h, respectively, and reduced significantly to 3% and 2% at 48 and 72 h. Likewise for IDI, the amount produced at 4, 6 and 24 h was 4%, 5% and 4%, then decreased to 1% at 48 h and remain stable until 72 h (Figure S2.B).The depletion of protein number could be due to the reduced amount of IPTG and proteins have been used for compounds biosynthesis. However, the decrease of protein level was not directly proportional to cell viability. The cell density was observed to increase at measurement from 4 to 72 h.

To increase the amount of soluble protein, temperature of incubation was optimized at 20 °C. Based on protein bands intensity on SDS-PAGE analysis, decreasing temperature of incubation to 20 °C could increase the amount of soluble protein compared to 37 °C (Figure S2.C). Low temperature was reported to decrease translation rate and to give sufficient time for protein to fold correctly to avoid the formation of inclusion bodies [29,30]. Further optimization was carried out for medium culture using LB and M9. However, based on analysis using SDS-PAGE and Western Blot, at the same amount of IPTG induction and

incubation time, the soluble protein obtained from LB media was still higher than in M9 media (Figure S2.D).

The only protein that majorly found as soluble protein in WB analysis was IDI. Considering that *idi* is the original gene from *E. coli*, it is suggested that easier for *E. coli* to carry out cloned *idi* translation process and to fold the protein properly, in contrast to other six proteins that were expressed heterologously. FPS was found in WB mostly as inclusion body which was likely to cause it inactive in *E. coli* [31]. Meanwhile, DXS, ADS and ALDH1 could not be detected neither as soluble protein nor inclusion bodies. This could be due to the amount of the proteins (DXS, ADS and ALDH1) in the sample were not sufficient enough to be detected by the WB method used in this study. Finding an appropriate expression condition for such a designed expression system is a big challenge, as well as a proper protein folding. Since optimizing the codons, inducer concentrations, time of induction, culture medium and cultivation temperature were not sufficient enough to make all target protein became soluble protein. Further studies to determine the production of metabolites in the presence of soluble recombinant proteins in *E. coli* (IDI, CYP, DBR2), especially the formation of precursors for artemisinin biosynthesis, is still in progress.

4. CONCLUSION

Two genes related to isoprenoid production through the MEP pathway (*dxs* and *idi*) and five genes related to artemisinin biosynthesis from *A. annua* (*FPS*, *ADS*, *CYP*, *DBR2* and *ALDH1*) have been successfully cloned and expressed in *E. coli*. The best condition to produce soluble and biologically active protein was under induction of 0.5 mM of IPTG for 24 h of incubation at 20 °C.

5. MATERIALS AND METHODS

5.1 Materials

LB medium (1% NaCl, 1% tripton, 0.5% yeast extract, 1.5% bacto agar), TSS (LB medium containing 10% PEG 6000, 5% DMSO and 50 mM $MgCl_2$), TAE 1x {diluted from TAE 50x (2 M Tris, 1 M glacial/acetic acid, 50 mM EDTA pH 8.0), DNA 1 kb ladder (Fermentas™), agarose (Top vision MDBio), Paq5000 DNA polymerase 2x master mix (Agilent Technologies), primers purchased from IDTDNA (Tabel), plasmid isolation kit (Qiagen), IPTG, reagents for SDS-PAGE {stacking buffer (0.5 M Tris-HCl pH 6.8), separating buffer (1.5 M Tris-HCl pH 8.8), 10% SDS, 40% acrylamide, 10% APS, TEMED}, SDS-PAGE running buffer (25 mM Tris, 192 mM glycine, 0.1% SDS), polyvinylidene fluoride (PVDF) membrane, blocking buffer (5% skim milk in TBS-T 1x), Western Blot transfer buffer (25 mM Tris, 192 mM glycine, 10% methanol), Tris-Buffered Saline-Tween (TBS-T) 1X {diluted from TBS 10x (200 mM Tris and 1500 mM NaCl) with 0.1% Tween 20}, reagents for protein visualization consisting of primary antibody 6x-His Tag rabbit polyclonal antibody (Invitrogen), secondary antibody alkaline phosphatase conjugated goat antirabbit IgG (Sigma), NBT (Bio Basic)/BCIP (Sigma) solution as substrate for alkaline phosphatase.

5.2 Methods

5.2.1 Plasmids construction

The gene fragments cloned in this research were *FPS*, *ADS*, *CYP*, *DBR2*, *ALDH1*, *dxs* and *idi*. The native sequence of the DNA can be accessed on NCBI using Acc. Nr. KJ609177.1, DQ241826.1, HQ315834.1, EU704257.1, FJ809784.1, NC_000964 and NC_000913, respectively. The DNA sequence of *ADS*, *dxs* and *idi* were native as in the original organisms. Changes were only made to some sequences to avoid the restriction sites used when inserting the DNA into plasmids. *FPS*, *CYP*, *DBR2* and *ALDH1* were optimized to the preferred codon usage of *E. coli* (http://www.kazusa.or.jp/codon/). The optimized *FPS* and *CYP* were synthetically made in pJexpress401 plasmid by DNA 2.0. *Nde*I and *Xho*I restriction sites were added at 5′ and 3′ ends of the sequence during DNA synthesis. The optimized *DBR2* and *ALDH1* were also synthetically produced and cloned into pRSFDUET-1 by Genscript resulting pRSFDda. Similarly, *dxs* and *idi* were synthetically produced and constructed into pCDFDUET-1 commercially by Genscript resulting pCDFDdi recombinant plasmid. *Bam*HI and *Eco*RI restriction sites were added at 5′ and 3′ ends of *DBR2* and *dxs*, while *Nde*I and *Xho*I were added at 5′ and 3′ ends of *ALDH1* and *idi*.

To produce pETDfac, the first gene cloned into pETDUET-1 was *CYP*. The *CYP* gene fragment was transferred from pJexpress_cyp to pETDUET-1 by restriction and ligation resulting pETDc recombinant plasmid. The next gene inserted was *ADS* that was isolated from pET15b-ads (kindly given by Prof. Wim J Quax from Groningen University) by PCR method. The *Bam*HI and *Not*I restriction sites were added to the sequence during the PCR process. The PCR product was firstly cloned into pGEMT-easy and then was transferred to pETDc by the same strategy as on pETDc, resulting pETDac recombinant plasmid. Finally,

FPS gene was isolated from pJexpress_fps by PCR to produce *FPS* gene carrying *Not*I restriction site at both 5′ and 3′ ends. PCR product and pETDac were digested using *Not*I and then ligated to produce pETDfac.

5.2.2 Plasmid transformation

Plasmids pETDfac and pRSFDda were transformed into *E. coli* TOP10, while pCDFDdi was transformed into *E. coli* DH5α for maintaining the plasmids. All the plasmids were further transformed into *E. coli* BL21(DE3) for protein overproduction. Plasmid transformation was conducted using heat shock method [32]. As much as 100 ng DNA were added to 200 µL competent cells treated with transformation and stock solution (TSS) containing $MgCl_2$. The mixture was then incubated at 42 °C for 90 s and quickly moved to the ice for 2 min. After the incubation, 800 µL LB medium was added into the mixture and then the mixture was incubated further using a shaker incubator at 37 °C, 150 rpm for 1 h. Before plating, cell culture was centrifuged (3000 *g*, 5 min). The obtained pellet cells were then re-suspended using 100 µL LB medium and then plated onto LB agar containing ampicillin 100 µg/ml for pETDfac, kanamycin 30 µg/ml for pRSFDda and streptomycin 50 µg/ml. Plasmid pETDfac, pRSFDda and pCDFDdi were also co-transformed into *E. coli* BL21(DE3) competent cells using the same method and plated on LB medium containing the three antibiotics for selection. Recombinant plasmids obtained from transformant cells were then analyzed.

5.2.3 Plasmids characterization

Plasmids isolated from transformant cells were characterized by migration, restriction and PCR analysis. Migration analysis performed in 8% agarose gel electrophoresis by running the plasmids compared to another plasmid. Confirmation of plasmids size was carried out by a single digestion using *Xho*I incubated at 37 °C. To investigate the genes inserted, double digestion using appropriate restriction enzymes and PCR analysis was conducted. PCR process was carried out using specific primers for each gene or universal primer for DUET plasmid provided by Novagen as shown in the Table.

Table. Primers used for PCR analysis

Genes	Primers Pairs	Primers Sequences (5′→ 3′)
FPS	For_fps_NotI1Duet	TTTGTTGCGGCCGCTCTAGAA
	Rev_fps_AflIIDuet	CTGCGGCTTAAGTTAGTGGTGATGGTG
ADS	ADS_Duet_F_BamHI	GAGGACGGATCCGATGTCACTTACAG
	ADS_Duet_R_EcoRI	CCGGACGCCATGGTCATATACTCATAG
CYP	DUETUP2	TTGTACACGGCCGCATAATC
	T7terminator	GCTAGTTATTGCTCAGCGG
DBR2	T7promoter	TAATACGACTCACTATAGGG
	DUETDOWN1	GATTATGCGGCCGTGTACA
ALDH1	DUETUP2	TTGTACACGGCCGCATAATC
	T7terminator	GCTAGTTATTGCTCAGCGG
dxs	T7promoter	TAATACGACTCACTATAGGG
	DUETDOWN1	GATTATGCGGCCGTGTACA
idi	DUETUP2	TTGTACACGGCCGCATAATC
	T7terminator	GCTAGTTATTGCTCAGCGG

5.2.4 Protein overproduction

Single colonies of recombinant cells were grown in media containing the appropriate antibiotics and inoculated overnight at 37 °C in a shaking incubator at a speed of 150 rpm. As much as 5% of overnight cultures were then subcultured in fresh medium and incubated until OD_{600} reading reached 0.6-0.7. Furthermore, the cultures were induced with IPTG and incubated further at specified time and condition. Optimization of the overproduction process was carried out on IPTG concentrations (0.1, 0.5 and 1.0 mM), incubation time (4, 6, 24, 48 and 72 h), incubation temperature (37 and 20 °C) and culture media (LB and M9). The cells were harvested by centrifugation at 3750 rpm for 20 min at 4 °C.

5.2.5 Protein characterization

SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS-PAGE analysis was conducted using 12% of acrylamide gel. To determine whether the proteins were expressed under the condition used, pellet from cell culture to yield OD_{600} 2.0 was re-suspended in

DTT:loading dye (1:9) and heated on boiled water for 10 min. The mixture was further centrifuged at 10.000 rpm for 10 min. As much of 5 µL of supernatant was loaded and run for SDS-PAGE analysis.

To determine whether the target proteins (FPS, ADS, CYP, DBR2, ALDH1, DXS, IDI) were expressed in soluble or inclusion bodies form, the cells harvested from 200 mL culture were resuspend in LEW buffer pH 8 (1:5) containing 1 mM phenylmethylsulfonyl fluoride (PMSF) and lysed by sonication at 70% amplitude for 20 min (15 s on, 15 s off) on ice. The lysate and cell debris were separated by centrifugation at 10.000 rpm for 10 min at 4 °C. The lysate contained crude protein in soluble form while cell debris contained protein in inclusion bodies form. As much of 40 µl of crude protein was added with 10 µl of loading dye containing bromophenol blue. Meanwhile, the cell debris obtained from 1.5 mL lysis mixture was weighed and re-suspended in loading dye containing bromophenol blue resulting 50% mixture concentration. Both crude protein and cell debris mixture were heated in boiled water for 10 min. The mixtures were centrifuged at 10.000 rpm for 10 min. Hereafter, 10 µl of each crude protein and cell debris supernatant were loaded and separated using SDS-PAGE.

Western Blot

The protein from polyacrylamide gel was transferred into methanol-activated PVDF membrane. The proteins were blocked by incubating the membrane in 5% skim milk in TBS-T 1x solution. To visualize the protein, the 6x-His Tag rabbit polyclonal antibody was used as the primary antibody (1:1000), and the secondary antibody was goat antirabbit IgG alkaline phosphatase conjugated (1:10000). The NBT/BCIP solution was used as substrate for alkaline phosphatase.

Acknowledgements: This work was conducted under grant research from Ministry of Research, Technology and Higher Education, Republic of Indonesia, and Lembaga Penelitian dan Pengabdian kepada Masayarakat, Institut Teknologi Bandung.

Author contributions: Concept – T.L., C.R., E.E.; Design – T.L., C.R., E.E.; Supervision – C.R., E.E., A.G.S.; Resources – E.E., A.G.S.; Materials – E.E, A.G.S.; Data Collection and/or Processing – T.L.; Analysis and/or Interpretation – T.L., C.R.; Literature Search – T.L.; Writing – T.L., C.R., E.E, A.G.S.; Critical Reviews – C.R., E.E., A.G.S.

Conflict of interest statement: The authors declared no conflict of interest in the manuscript.

REFERENCES

[1] WHO World malaria report 2021. https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2021 (accsessed on 15 December 2021).

[2] Tu Y. Artemisinin-A gift from traditional chinese medicine to the world (Nobel Lecture). Angew Chem Int Ed Engl. 2016; 55(35): 10210-10226. [CrossRef]

[3] Numonov S, Farukh S, Aminjon S, Parviz S, Sunbula A, Ramazon S, Maidina H. Assessment of artemisinin contents in selected Artemisia species from Tajikistan (Central Asia). Medicines (Basel). 2019; 6(1): 23. [CrossRef]

[4] Czechowski T, Larson TR, Catania TM, Harvey D, Wei C, Essome M, Brown GD, Graham IA. Detailed phytochemical analysis of high and low artemisinin-producing chemotypes of Artemisia annua. Front Plant Sci. 2018; 9: 641. [CrossRef]

[5] Nguyen KT, Arsenault PR, Weathers P J. Trichomes + roots + ROS = artemisinin: regulating artemisinin biosynthesis in Artemisia annua L. In Vitro Cell Dev Biol Plant. 2011; 47(3): 329–338. [CrossRef]

[6] Wen W, Yu R. Artemisinin biosynthesis and its regulatory enzymes: Progress and perspective. Phcog Rev. 2011; 5(10): 189–194. [CrossRef]

[7] Pateraki I, Heskes AM, Hamberger B. Cytochromes P450 for terpene functionalisation and metabolic engineering. Adv Biochem Eng Biotechnol. 2015; 148: 107-139. [CrossRef]

[8] Yang C, Gao X, Jiang Y, Sun B, Gao F, Yang S. Synergy between methylerythritol phosphate pathway and mevalonate pathway for isoprene production in Escherichia coli. Metab Eng. 2016; 37: 79-91. [CrossRef]

[9] Zhao Y, Yang J, Qin B, Li Y. Biosynthesis of isoprene in Escherichia coli via methylerythritol phosphate (MEP) pathway. Appl Microbiol Biotechnol. 2011; 90: 1915–1922. [CrossRef]

[10] Zhou K, Zou R, Zhang C, Stephanopoulos G, Too H. Optimization of amorphadiene synthesis in Bacillus subtilis via transcriptional, translational and media modulation. Biotechnol Bioeng. 2013; 110(9): 2556–2561. [CrossRef]

[11] Martin VJJ, Pitera DJ, Withers ST, Newman JD, Keasling JD. Engineering a mevalonate pathway in Escherichia coli for production of terpenoids. Nat Biotechnol. 2003; 21(7): 796–802. [CrossRef]

[12] Newman JD, Marshall J, Chang M, Nowroozi F, Paradise E, Pitera D, Newman KL, Keasling JD. High-Level Production of amorpha-4 , 11- diene in a two-phase partitioning bioreactor of metabolically engineered Escherichia coli. Biotechnol Bioeng. 2006; 95(4): 684-691. [CrossRef]

[13] Tsuruta H, Paddon CJ, Eng D, Lenihan JR, Horning T, Anthony LC, Regentin R, Keasling JD, Renninger NS, Newman JD. High-level production of amorpha-4, 11-diene, a precursor of the antimalarial agent artemisinin, in Escherichia coli. PLoS One. 2009; 4(2): e4489. [CrossRef]

[14] Wang, J, Apptec W, Xiong Z. Engineering MEP Pathway in Eschericia coli for amorphadiene production and optimizing the bioprocess through glucose feeding control. Chin J Biotech. 2014; 30(1): 64-75. [CrossRef]

[15] Chung BK, Klement M, Lakshmanan M, Mohanty B, Lee D. Genome-scale in silico modeling and analysis for designing synthetic terpenoid-producing microbial cell factories. Chem Eng Sci. 2012; 103(September): 1–9. [CrossRef]

[16] Wang Q, Quan S, Xiao H. Towards efficient terpenoid biosynthesis : manipulating IPP and DMAPP supply. Bioresour Bioprocess. 2019; 6(6): 1-13. [CrossRef]

[17] Yuan LZ, Rouvière PE, LaRossa RA, Suh W. Chromosomal promoter replacement of the isoprenoid pathway for enhancing carotenoid production in *E. coli*. Metab Eng. 2006; 8(1): 79–90. [CrossRef]

[18] Volke DC, Rohwer J, Fischer R, Jennewein S. Investigation of the methylerythritol 4-phosphate pathway for microbial terpenoid production through metabolic control analysis. Microb Cell Fact. 2019; 18(1): 192. [CrossRef]

[19] Kane JF. Effects of rare codon clusters on high-level expression of heterologous proteins in Escherichia coli. Curr Opin Biotechnol. 1995; 6(5): 494–500. [CrossRef]

[20] Kurland C, Gallantt J. Errors of heterologous protein expression. Curr Opin Biotechnol. 1996; 7(5): 489-493. [CrossRef]

[21] Gustafsson C, Minshull J, Govindarajan S, Ness J, Villalobos A, Welch M, Drive OB, Suite A, Park M. Engineering genes for predictable protein expression. Protein Expr Purif. 2012; 83(1): 37–46. [CrossRef]

[22] Wallaart TE, Bouwmeester HJ, Hille J, Poppinga L, Maijers NCA. Amorpha-4,11-diene synthase: Cloning and functional expression of a key enzyme in the biosynthetic pathway of the novel antimalarial drug artemisinin. Planta. 2001; 212(3): 460–465. [CrossRef]

[23] Paddon CJ, Westfall PJ, Pitera DJ, Benjamin K, Fisher K, McPhee D, Leavell MD, Tai A, Mai A, Eng D, Polichuk DR, Teoh KH, Reed DW, Treynor T, Lenihan J, Jiang H, Fleck M, Bajad S, Dang G, … Newman JD. High-level semi-synthetic production of the potent antimalarial artemisinin. Nature. 2013; 496(7446): 528–532. [CrossRef]

[24] Peplow M. Synthetic malaria drug meets market resistance. Nature. 2016; 530: 389–390. [CrossRef]

[25] Feth MP, Rossen K, Burgard A. A Pilot Plant PAT Approach for the Diastereoselective Diimide Reduction of Artemisinic Acid. Org Process Res Dev. 2013; *17*(2): 282–293. [CrossRef]

[26] Amara Z, Bellamy JFB, Horvath R, Miller SJ, Beeby A, Burgard A, Rossen K, Poliakoff M, George MW. Applying green chemistry to the photochemical route to artemisinin. Nat Chem. 2015; 7: 489-495. [CrossRef]

[27] Gilmore K, Daniel K, Ju WL, Zoltán, McQuade DT, Seidel-Morgenstern A, Seeberger PH. Continuous synthesis of artemisinin-derived medicines. Chem Commun (Camb). 2014; 50(84): 12652-12655. [CrossRef]

[28] Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods. 2012; 9(7): 671-675 [CrossRef]

[29] Schein CH. Production of soluble recombinant proteins in bacteria. Nat Biotechno. 1989; 7: 1141–1149. [CrossRef]

[30] Feller G, Bussy OLE. Expression of psychrophilic genes in mesophilic hosts : Assessment of the folding state of a recombinant alpha-amylase. Appl Environ Microbiol. 1998; 64(3): 1163–1165. [CrossRef]

[31] Vallejo LF, Rinas U. Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins. Microb Cell Fact. 2004; 3(1): 11. [CrossRef]

[32] Chung CT, Niemela SL, Miller RH. One-step preparation of competent Escherichia coli : Transformation and storage of bacterial cells in the same solution. Proc Natl Acad Sci USA. 1989; 86(7): 2172–2175. [CrossRef]