The 6th International Conference on Biochemistry and Molecular Biology

Molecular and functional characterization of sesquiterpene synthase1 from *Piper betle* L.

Kodchakorn Aranyakanon[1], Rattikarn Buaruaeng[1], Thunyared Kaewkhaw[1], Paphitchaya Naree[1], Chonnanit Choopayak[1,*]

[1]Department of Biochemistry, Faculty of Medical Science, Naresuan University, Phitsanulok, Thailand
*E-mail: chonnanitc@nu.ac.th

Abstract

Sesquiterpene volatiles have several biological activities, however, its applications still be limited. To promote sesquiterpene production, biotechnology is the alternative strategy. Previously, we was successfully isolated the full length of *Piper betle sesquiterpene synthase1* (*PbSTS1*) RNA from *Piper betle* L. and cloned into pET28b vector. Nevertheless, its function was not characterized. Here, molecular structure of PbSTS1 was characterized *in silico*. The recombinant protein was produced in *Escherichia coli* and was purified. The enzyme activity was identified using GC-MS and enzyme kinetics was indirectly investigated by pyrophosphate detection kit. *In silico* results indicated that PbSTS1 amino acid was the most similar to germacrene D synthase in *Vitis vinifera* for 44.7%. Three dimension structure analysis showed three catalytic regions; DDXXD, NSE/DTE and RXR motifs that interacted with farnesyl pyrophosphate (FPP). After expression and purification, the only one protein band about 64 kDa was detected. By incubating the purified PbSTS1 with FPP, GC-MS analysis showed only one sesquiterpene product as germacrene D. The Km, Vmax and Kcat values of PbSTS1 to FPP were 32.57 µM, 7.9 µmole/min and 6.4 s^{-1}, respectively. Taken together, we have successfully isolated the functional germacrene D synthase from *P. betle* and capable to synthesize germacrene D *in vitro*. The PbSTS1 had a promising potency to be used for germacrene D production using biotechnology.

Introduction

Piper betle L. is the most cultivation in South and West of Thailand. It often use in ancient religious ceremony and also traditional medicine such as wound healing, dental problems, headaches, arthritis and joint pain[1]. Since, it produces several active compounds such as flavonoids, alkaloids and terpenoids[1]. Terpenoids are a large group of secondary metabolite which greater producing in higher plant especially in *P. betle*. Terpenoids are divided different types base on their structure and function such as monoterpene ($C_{10}H_{16}$), sesquiterpene ($C_{15}H_{24}$) and diterpene ($C_{20}H_{32}$). Sesquiterpene are a class of terpenoids that consist of three isoprene units. Sesquiterpenes are reported with several mechanisms such as anti-fungal activity, anti-inflammatory, anti-oxidant and anticancers[5-7]. Sesquiterpenes are produced via Mevalonate pathway (MVP) in plant cytosol. In MVP pathway, the sesquiterpene products are synthesized by the key enzyme namely sesquiterpene synthase (STS) which catalyzed farnesyl pyrophosphate (FPP) as a substrate to produce sesquiterpenes[2]. Sesquiterpene synthases are localized in the cytoplasm[3]. It contains DDXXD and NSE/DTE motifs which has been implicated in binding with divalent metal cofactors such as Mg^{2+}, together with the other commonly conserved motifs in sesquiterpene synthase consist RXR motif [4]. Sesquiterpene synthase are belonged in enzyme group namely terpene synthase (TPSs). Plant terpene synthase

(TPSs) are grouped into seven subfamily including TPS-a (sesquiterpene), TPS-b (cyclic monoterpene and hemiterpene), TPS-c (copalyl diphosphate and ent-kaurene), TPS-d (gymnosperm specific), TPS-e/f (ent-kaurene and other diterpenes as well as some monoterpene and sesquiterpene) and TPS-g (acyclic monoterpene) [18]. However, the plant producing sesquitepene is unstable source and difficult to synthesized by chemical. Thus, the steady of sesquiterpene production is required. The biotechnology is important process to promote sesquiterpene production. In previous study, microorganism were used as a source for high-level production of sesquiterpene[8]. Thus, the biotechnology could be the alternative attention to increase sesquiterpene. Previously, Chonnanit et al isolated *PbSTS1* RNA from *P. betle* which probably to be sesquiterpene synthase and cloned into pET28b vector, but its function was uncharacterized. In this study, PbSTS1 the deduced amino acid accession number ANA50340 from NCBI database was characterized and predicted structure *in silico*. To demonstrate its functional, the recombinant PbSTS1 protein was expressed in *Escherichia coli* and purified. To confirm its activity, the expected product namely germacrene D was synthesize by enzyme activity assay *in vitro* and then identified by GC-MS.

Materials and methods

Molecular characterization of PbSTS1 *in silico*

The full length PbSTS1 reported in NCBI database (accession number ANA50340), it was analyzed for comparative homologs sesquiterpene synthase by BLAST algorithm Uniprot database (https://www.uniprot.org/blast/). To identify conserve motifs and catalytic regions, the multiple amino acid sequence alignment of homolog proteins were performed with Clustal Omega program at EMBL (https://www.ebi.ac.uk/Tools/msa/clustalo/). The homolog proteins were related with *P. betle* and grouped in Mesangiospermae. Phylogenetic analysis was generated from neighbor-joining method in MEGA software version 7. To study substrate interaction, the 3D structure of PbSTS1 was predicted in I-TASSER server (https://zhanglab.ccmb.med.umich.edu/I-TASSER/) matching with differential terpene synthase protein's family. PbSTS1 models from I-TASSER were bound with farnesyl pyrophosphate ($C_{15}H_{28}O_7P_2$) molecule and analyzed by using SWISS DOC and Pymol program.

Recombinant PbSTS1 protein production and purification

Previously, *PbSTS1* was cloned into pET28b expression vector. The vector was contained N-terminal poly-histidine (6His)-tagged proteins which required for purification. In this study, the construct *pET28b-PbSTS* was transformed to *Escherichia coli* (BL21star) competent cell and cultured on LB agar with kanamycin (50µg/ml). To produce the recombinant protein, single colony was inoculated in 5 ml LB liquid medium with kanamycin (50µg/ml) and then incubated with vigorous shaking (180 rpm) overnight. We transferred inoculum cell 1% into LB liquid medium with kanamycin (50µg/ml) and incubated until optical density equal 0.4-0.6 at OD_{600}. The 0.5mM isopropyl β-d-1-thiogalactopyranoside (IPTG) was exposed to induce recombinant PbSTS1 production and incubated overnight at 30ºC.

To purify recombinant protein, the induced cells was centrifuged and harvested in assay buffer including 50mM HEPES (pH 7.3), 7.5mM MgCl2, 100mM KCl, 5mM DTT, 10% Glycerol and protease inhibitor. The cells were lysed by Sonicator at 60% of pulse for 2 min on ice. Subsequently, the cell lysate and supernatant were then collected by low speed centrifugation and only cell lysate was kept in 5x loading dye (250 Tris-HCl pH6.8, 10%SDS, 30% Glyerol, 5% β-mercaptoethanol and 0.02% Bromophenol blue). The supernatant was bounded with Ni-sepharose beads by binding buffer including 40mM imidazole and packed into column. Next, the column was eluted with increasing imidazole concentration (500 mM) and collected four fractions. Expression and purity of the recombinant protein was assessed by using a 10%

polyacrylamide gel. Quantitation of recombinant proteins was carried out using Bradford protein assay.

Enzyme kinetics

PbSTS1 kinetic values were obtained by indirect detection using pyrophosphate detection kit P7275 (Sigma-aldrich, USA) according to the instructions of the kit. The principle of detection, sesquiterpene synthase (PbSTS1) catalyze farnesyl pyrophosphate to sequitepene products and free pyrophosphate. The free pyrophosphate convert to DHAP (Dihydroxyacetone phosphate) and it oxidizes by β-NADH (β-Nicotinamide adenine dinucleotid). The reaction (β-NADH molecules) is monitored spectrophotometrically at 340 nm. The enzyme kinetics were presented in K_m, V_{max}, K_{cat} and K_{cat}/K_m values. Control assays using heat-inactivated recombinant PbSTS1 protein. The experiment were done in triplicate.

Enzyme activity

In vitro enzyme activity assay, the crude protein and purified PbSTS1 were used in this study. In case of crude protein, the assay was performed in 500 µl including 496 µl of crude protein and 40 µM of FPP (Sigma, USA). In assay for purified PbSTS1, 40 µg of purified PbSTS1 was mixed with 40 µM of FPP and assay buffer containing 50 mM HEPES, 7.5 mM $MgCl_2$, 100 mM KCl, 5 mM DTT, 10% glycerol and proteinase inhibitor (Roche, Germany). The either assays were overlaid with 250 µl of hexane and incubated at 30 °C for 1 h. After that, the reaction was mixed by vortex and centrifuged at 1,000 g for 30 min at 4 °C. The hexane extract was collect and then subjected to Gas chromatography/Mass spectrometry (GC-MS).

The hexane extract of enzyme activity assay was analyzed by Agilent Technologies GC/MS system. GC Agilent 6890N and MS Agilent 2577A 5973N, equipped with a HP-5MS capillary column (30 m x 0.25 mm x i.d. film thickness 0.25 µm). The oven temperature was held at 45°C for 0.2 min, then programmed to increase to 200°C at a rate of 10°C/min and then increase to 245°C at a rate of 30°C/min and held at 280°C for 5 min. The other parameter: injector temperature at 250°C; ion source temperature at 280°C; carrier gas He at a flow rate of 1 ml/min; splitless, mass range 25-300 m/z. The compound was identified by comparison its mass spectral fragmentation base on the Wiley Mass Spectral library.

Results and discussion

Molecular characterization of PbSTS1

The amino acid sequence of PbSTS1, there were 561 amino acids and encoded a protein molecular weight at 64.7 kDa. The homologs sesquiterpene synthase analysis by BLAST algorithm at Uniprot database showed PbSTS1 the most similarity to germacrene D synthase in *Vitis vinifera* (VIT19) 44.7% of identity. The major product of germacrene D synthase in *Vitis vinifera* was identified as (−)-germacrene D[9]. Therefore, PbSTS1 may serve a similar function as VIT19. The amino acid multiple sequence alignment was compared the protein sequence of PbSTS1 with known sesquiterpene synthase which relatedness taxonomy belonged in piperaceae family such as cadinene synthase in *Piper nigrum* and other were classed in mesangiospermae such as germacrene D synthase in *Vitis vinifera* and β-cubebene synthase in *Magnolia grandiflora*. The result illustrated that the full length of PbSTS1 including start and stop codons which had not been found between the sequence. The conserved motifs DDXXD, NSE/DTE and RXR motifs were presented in PbSTS1 amino acid sequence which is highly conserved in plant sesquiterpene synthase[4]. In addition, the alignment result revealed no N-terminal signal peptide sequences such as the plastid targeting sequence which required for synthesis of monoterpene or signal transmembrane protein[10]. This suggestion was explained its cytosolic localization (**Figure 1**.).

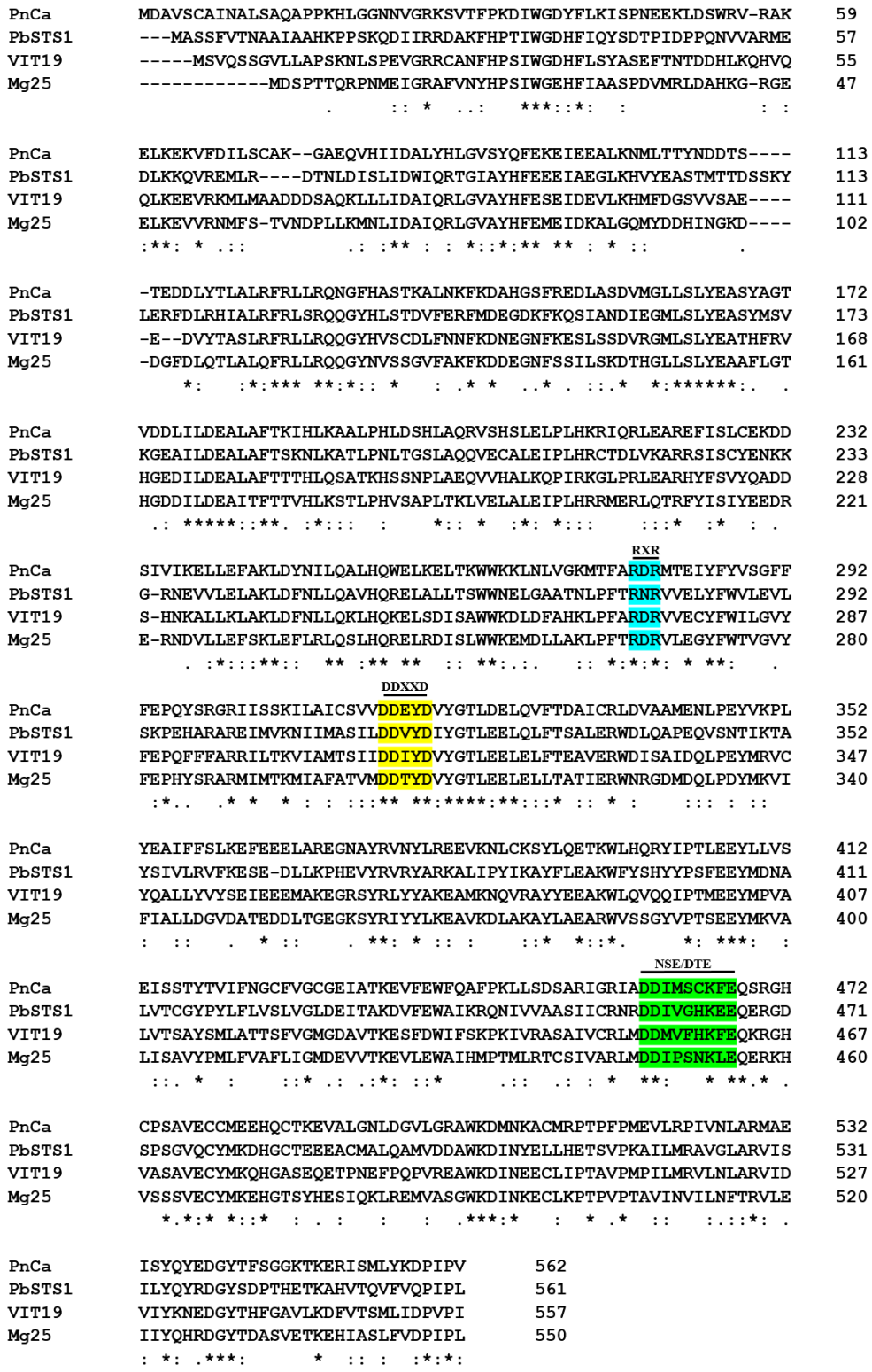

Figure 1. Multiple alignments of PbSTS1 with closely related sesquitepene synthase were performed by Clustal Omega. The RXR, DDXXD and NSE/DTE were presented in blue, yellow and green respectively. PnCa; cadinene synthase (*Piper nigrum*), VIT19; germacrene D synthase (*Vitis vinifera*) and Mg25; β-cubebene Synthase (*Magnolia grandiflora*).

TPS plant family has been divided into seven subfamilies, i.e., TPS-a through to TPS-g, based on sequence relatedness as well as functional assessment[13]. To determine TPSs subfamily of PbSTS1, the amino acid sequence of PbSTS1 was analyzed with known sesquiterpene synthases as well as other terpene synthases of each subfamily. The phylogenetic

tree analysis showed that PbSTS1 was belonged in TPS-a subfamily (**Figure 2**). It was closely with MpFS, SoCa19, F383 and VIT19. TPS-a subfamily is angiosperm specific sesqui-TPSs. The other TPSs subfamily including TPS-b are angiosperm mono-TPSs which consist of R(R)X_8W motif and produced cyclic monoterpene. While, TPS-g is angiosperm mono-TPSs but lacks R(R)X_8W motif which capable to produce acyclic monoterpenes such as geraniol synthase in *Catharanthus roseus*. The TPS-c and e are angiosperm di-TPSs such as copalyl diphosphate synthases and kaurene synthases. TPS-d is specific gymnosperm which is different types of TPSs (mono-, sesqui- as well as di-TPSs) [10].

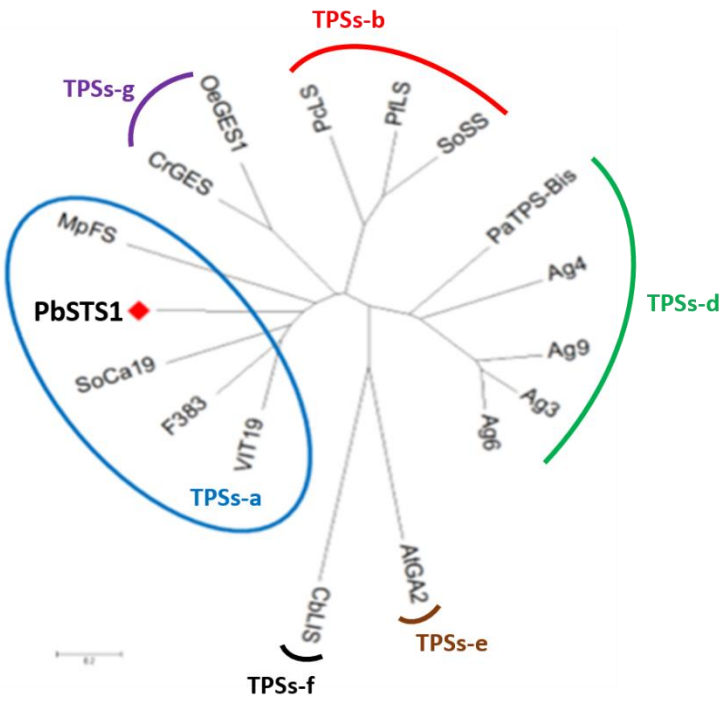

Figure 2. Phylogenetic tree analysis of PbSTS1 and other proteins in TPSs subfamily a-g. TPSs-a: VIT19; germacrene D synthase [*Vitis vinifera*], MpFS; β-farnesene synthase [*Mentha piperita*], SoCa19; germacrene D synthase [*Solidago canadensis*], F383; germacrene D synthase [*Gossypium arboretum*]. TPSs-b: PcLS; limonene synthase [*Perilla citriodora*], PfLS; linalool synthase [*Perilla frutescens*], SoSS; sabinene synthase [*Salvia officinalis*]. TPSs-d: Ag3; pinene synthase [*Abies grandis*], Ag4; δ-selinene synthase, Ag9; terpinolene synthase, PaTPS-Bis; α-bisabolene synthase [*Picea abies*]. TPSs-e: AtGA2; kaurene synthase [*Arabidopsis thaliana*]. TPSs-f: CbLIS; linalool synthase [*Clarkia breweri*]. TPSs-g: CrGES; geraniol synthase [*Catharanthus roseus*] and OeGES1; geraniol synthase [*Olea europaea*].

A three-dimension structure of PbSTS1 was predicted by I-TASSER and cadinene synthase of *Gossypium arboretum* was used as a model. The TM-Score and C-Score equal 0.883 and -0.29 respectively which the most suitable for synthesis of PbSTS1 three-dimension structure. The PbSTS1 three-dimension structure was a globular protein. In the structure, PbSTS1 was found α-helix and β-pleated sheet which contained N-terminus (red) and C-terminus (green) at the 1-139 and 240-560 of sequence. The RXR, DDXXD and NSE/DTE were all expressed in C-terminus and represented at amino acid 275-277, 314-318 and 457-465 respectively. Likewise, there was reported sesquiterpene synthase from *Polygonum minus* which consisted DDXXD motif at 314-318 and NSE/DTE at 465-473 of its sequence[12] which similar to conserved motifs of PbSTS1. In mention above, we knew conserve motifs including RXR, DDXXD and NSE/DTE of sesquitepene synthase chelate divalent metal ions and required these motifs for cyclization of farnesyl pyrophosphate (FPP) to generate varies sesquiterpene products[13]. To predict possibility of interaction with the specific substrate,

PbSTS1 model was observed interaction with FPP substrate. The result showed Arg278 of RXR, Asp315 of DDXXD and Asp 459 of NSE/DTE interacting with farnesyl pyrophosphate. The H atom number 21 of Arg278 interacted with first O atom of FPP, Asp315 generated hydrogen bound between H atom number 5 with fourth O of FPP and O atom of Asp495 interacted with H number 27 of FPP, the distance between molecules were 2.6, 3.3 and 10.5 ºA respectively. In the catalysis of sesquitepene synthase enzyme, these motifs have been reported flank the entrance of active site via involved a trinuclear magnesium cluster. Two magnesium bind with DDXXD motif and other one magnesium ion interacts to NSE/DTE motif. In addition, the Mg^{2+} cluster attaches to the pyrophosphate (PPi) group of FPP and induces the position of hydrophobic substrate binding in pocket enzyme[14]. Therefore, PbSTS1 was possible to interact with FPP and generates sesquiterpene.

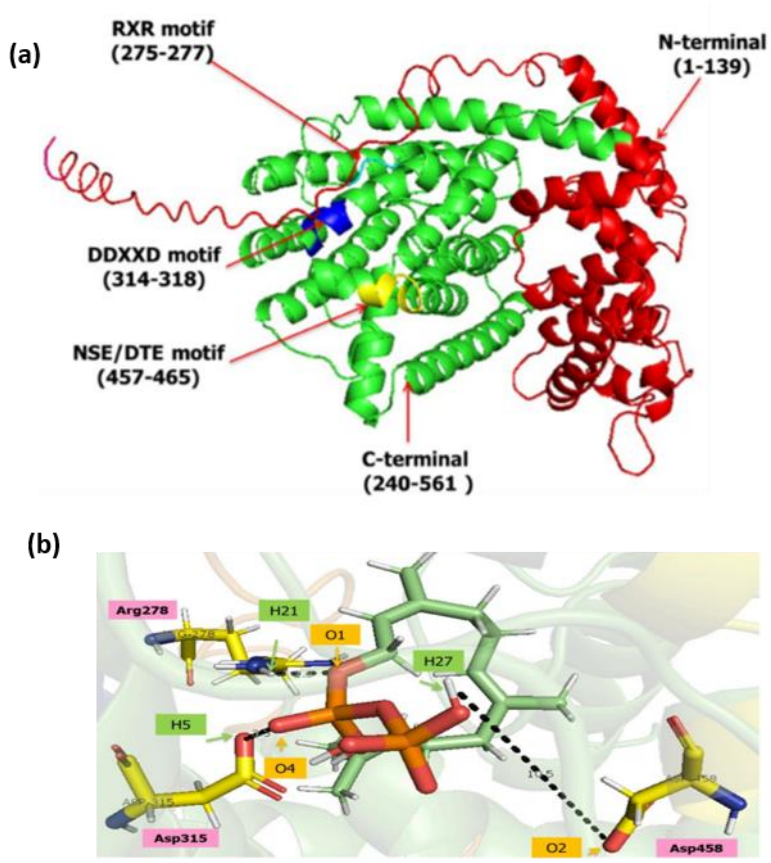

Figure 3. A three-dimension structure of PbSTS1 predicted by I-TASSER (a). The substrate binding model in catalytic site of PbSTS1 generated by Pymol (b).

Recombinant protein production and purification

In this study, the recombinant PbSTS1 protein was successfully expressed by 0.5mM IPTG induction and molecular mass approximately 64 kDa was found in after induction (red arrow) while did not appeared the expected PbSTS1 protein band in before induction (lane1). The recombinant PbSTS1 was found the highest in purified first fraction (lane3). The other reported sesquiterpene synthase were similarly molecular weight with PbSTS1 such as β-sesquiphellandrene synthase form *Persicaria minor* and patchoulol synthase from *Pogostemon cablin*. These enzyme were encode protein at 65 and 64.2 kDa respectively[15-16].

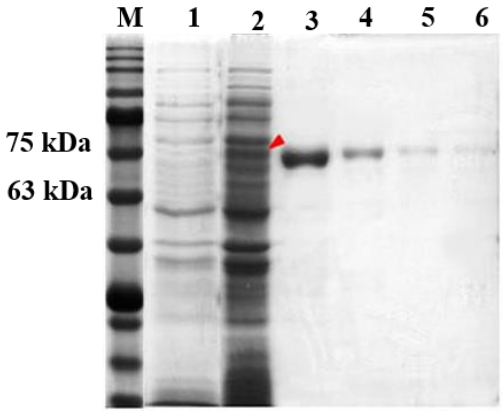

Figure 4. SDS-PAGE analysis of recombinant protein production in *Escherichia coli* (BL21star). The red arrow showed the expected band of recombinant PbSTS1. Lane M; protein molecular marker (BLUeye Prestained Protein Ladder, GeneDireX®), Lane1; protein before induction, Lane2; protein after induction (crude protein), Lane3 to lane6; purified protein fraction 1 to 4 respectively.

Enzyme kinetics

Base on indirect detection method, the kinetic constants (K_m, V_{max}, K_{cat} and K_{cat}/K_m) were determined by incubating fixed amount of enzyme with varied concentration of FPP as a substrate (25 to 100 µM). The K_m and V_{max} were derived from Lineweaver burk plot and found to be 32.57 µM and 7.9 µmole/min respectively (Figure 5). The K_{cat} value was calculated at 6.4 s^{-1}. The K_{cat} value correlated to maximum of substrate molecules converted to product per active site per unit of time. Therefore, the K_{cat}/K_m referred to catalytic efficiency and found to be 0.196 $s^{-1}\mu M^{-1}$. The kinetic parameters were compared with the reported sesquiterpene synthase namely 5-epi-aristolochene synthase (TEAS) from *Nicotiana tabacum*[17] which also used FPP as a substrate. The comparative kinetic parameters were showed that K_m value of PbSTS1 was higher than TEAS that indicated PbSTS1 lower affinity to bind the FPP than TEAS. In the other hand, the K_{cat}/K_m of PbSTS1 was higher than TEAS that interpreted whether PbSTS1 higher efficiency to convert FPP to product sesquiterpene more than TEAS.

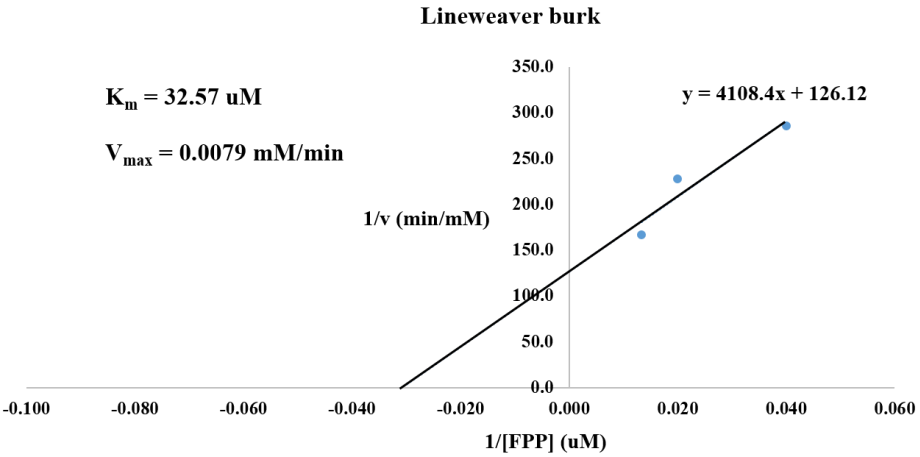

Figure 5. The enzyme kinetic constants were presented in Km and Vmax values which derived from lineweaver bruk plot

Enzyme activity and GC-MS analysis

In mention above, whole PbSTS1 amino acid sequence was probable to be germacrene D synthase. To investigate the PbSTS1 activity, enzyme activity was performed by incubating with FPP as substrate and confirmed its product by GC-MS. The results showed the only one

product from the reaction (Figure 6A) and then identified as germacrene D by comparison of the mass fragmentation patterns (Figure 6C), while negative control did not appeared any product (Figure 6B). For this reason, PbSTS1 absolutely sure to be germacrene D synthase in *P. betle*.

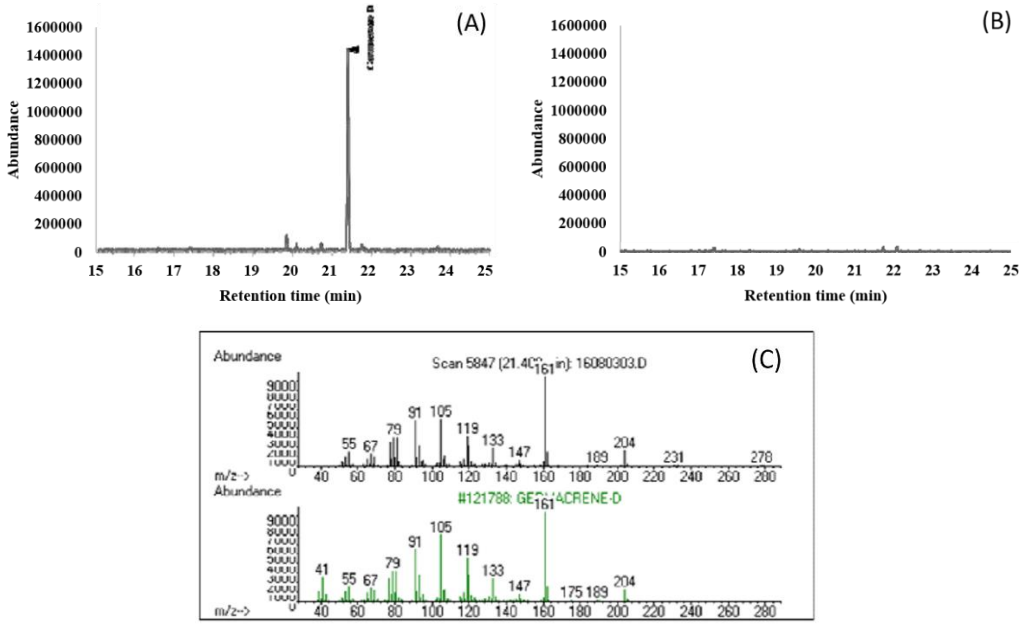

Figure 6. GC-MS analysis results showed (A) Chromatogram of PbSTS1's product, (B) negative control and (C) germacrene D mass spectral comparison.

Conclusion

Sesquiterpenes are a little constitutive producing in *Piper betle* L. and capable to induce by several stimulus such as pathogen and herbivore. Thus, the gene which encodes sesquiterpene synthase probably active and suitable used for the biotechnology. In this study, the first reported full length PbSTS1 amino acid sequence accession number ANA50340 was a functionally active sequence to encode a sesquiterpene synthase namely germacrene D synthase, and then capable to synthesize specific product that identified as germacrene D. These results suggested the biotechnology is the alternative process that could be utilize for enhance and reliable source of sesquiterpene production.

Reference

1. Fazal F, Mane PP, Rai MP, Thilakchand KR, Bhat HP, Kamble PS and et al. Chin J Integr Med. 2014:1-11.
2. Tholl D. Curr Opin Plant Biol. 2006; 9(3):297-304.
3. Taniguchi S, Miyoshi S, Tamaoki D, Yamada S, Tanaka K, Uji Y, Tanaka S, Akimitsu K, Gomi K. J Plant Physiol. 2014;171:625–632.
4. Ee S-F, Mohamed-Hussein Z-A, Othman R, Shaharuddin NA, Ismail I, Zainal Z. Sci World J. 2014: 11.
5. Kang TH, Lee YM, Lee WJ, Hwang EI, Park KD, Choi GJ and et al. J Microbiol Biotechnol. 2017;27(7):1272-1275
6. Makabe H, Maru N, Kuwabara A, Kamo T, Hirota M. Nat Prod Res. 2006;20(7):680-5.
7. Dahham SS, Tabana YM, Iqbal MA, Ahamed MB, Ezzat MO, Majid AS and et al. Molecules. 2015; 20(7):11808-29.
8. Paddon CJ, Westfall PJ, Pitera DJ1, Benjamin K, Fisher K, McPhee D. Nature. 2013; 496: 528-532.

9. Lücker J, Bowen P, Bohlmann J. Phytochemistry. 2004; 65(19): 2649-59.
10. Jin J, Kim MJ, Dhandapani S, Tjhang JG, Yin JL, Wong L. J Exp Bot. 2015;66(13):3959-75.
11. Ruan JX, Li JX, Fang X, Wang LJ, Hu WL, Chen XY. Front Plant Sci. 2016; 7: 638.
12. Ee S-F, Mohamed-Hussein Z-A, Othman R, Shaharuddin NA, Ismail I, Zainal Z. Sci World J. 2014: 11.
13. Bohlmann J, Meyer-Gauen G, Croteau R. Proc Natl Acad Sci USA. 1998; 95:4126–4133.
14. Lopez-Gallego F, Agger SA, Pella DA, Distefano MD, Schmidt-Dannert C. Chembiochem. 2010; 11(8): 1093–1106.
15. López-Gallego F, Wawrzyn GT, Schmidt-Dannert C. JAEM. 2010; 76(23):7723–7733.
16. Ker DS, Pang SL, Othman NF, Kumaran S, Tan EF, Krishnan T. PeerJ. 2017; 5: e2961.
17. Frister T,Hartwig S, Alemdar S, Schnatz K, Thöns L, Scheper T. Appl. Biochem. Biotechnol. 2015; 176:2185–2201.
18. Külheim C, Padovan A, Hefer C, Krause ST, Köllner TG, Myburg AA. BMC Genomics. 2015; 16:450.

Acknowledgement

This study was financially supported by faculty of Medical Science, Naresuan University and thanks Dr. Krit Tantanarat for competent cell (*Escherichia coli* (BL21star™)).