J Ind Microbiol Biotechnol (2015) 42:1341–1352
DOI 10.1007/s10295-015-1668-8

BIOCATALYSIS

Directed optimization of a newly identified squalene synthase from *Mortierella alpine* based on sequence truncation and site-directed mutagenesis

Di Huang[1,2,3] · Yongpeng Yao[1,2] · Hang Zhang[1,2] · Zhu Mei[1,2] · Ru Wang[1,2] · Lu Feng[1,2,3,4,5] · Bin Liu[1,2,3]

Received: 3 June 2015 / Accepted: 28 July 2015 / Published online: 15 August 2015

Abstract Terpenoids, a class of isoprenoids usually isolated from plants, are always used as commercial flavor and anticancer drugs. As a key precursor for triterpenes and sterols, biosynthesis of squalene (SQ) can be catalyzed by squalene synthase (SQS) from two farnesyl diphosphate molecules. In this work, the key SQS gene involved in sterols synthesis by *Mortierella alpine*, an industrial strain often used to produce unsaturated fatty acid such as γ-linolenic acid and arachidonic acid, was identified and characterized. Bioinformatic analysis indicated that MaSQS contained 416 amino acid residues involved in four highly conserved regions. Phylogenetic analysis revealed the closest relationship of MaSQS with *Ganoderma lucidum* and *Aspergillus*, which also belonged to the member of the fungus. Subsequently, the recombinant protein was expressed in *Escherichia coli* BL21(DE3) and detected by SDS-PAGE. To improve the expression and solubility of protein, 17 or 27 amino acids in the C-terminal were deleted. In vitro activity investigation based on gas chromatography–mass spectrometry revealed that both the truncated enzymes could functionally catalyze the reaction from FPP to SQ and the enzymatic activity was optimal at 37 °C, pH 7.2. Moreover, based on the site-directed mutagenesis, the mutant enzyme mMaSQSΔC17 (E186K) displayed a 3.4-fold improvement in catalytic efficiency (k_{cat}/K_m) compared to the control. It was the first report of characterization and modification of SQS from *M. alpine,* which facilitated the investigation of isoprenoid biosynthesis in the fungus. The engineered mMaSQSΔC17 (E186K) can be a potential candidate of the terpenes and steroids synthesis employed for synthetic biology.

Keywords Squalene synthase · *Mortierella alpine* · Truncated enzyme · GC–MS analysis · Site-directed mutagenesis

Di Huang and Yongpeng Yao have contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s10295-015-1668-8) contains supplementary material, which is available to authorized users.

✉ Bin Liu
liubin1981@nankai.edu.cn

1 TEDA Institute of Biological Sciences and Biotechnology, Nankai University, TEDA, Tianjin 300457, People's Republic of China

2 Key Laboratory of Molecular Microbiology and Technology, Ministry of Education, Tianjin 300071, People's Republic of China

3 Tianjin Key Laboratory of Microbial Functional Genomics, Tianjin 300457, People's Republic of China

4 State Key Laboratory of Medicinal Chemical Biology, Nankai University, Tianjin 300071, People's Republic of China

5 SynBio Research Platform, Collaborative Innovation Center of Chemical Science and Engineering (Tianjin), Tianjin 300072, People's Republic of China

Introduction

Terpenoids, which are usually isolated from plants, belong to a class of isoprenoid family with important medical and industrial properties. Sterols are a subgroup of the steroids and an important class of organic molecules. Due to their crucial functions in vivo and significant pharmaceutical application, the triterpenes and sterols biosynthesis have attracted increasing interest in recent years [1, 4, 7, 40]. For example, the engineered oleaginous yeast and fungi have

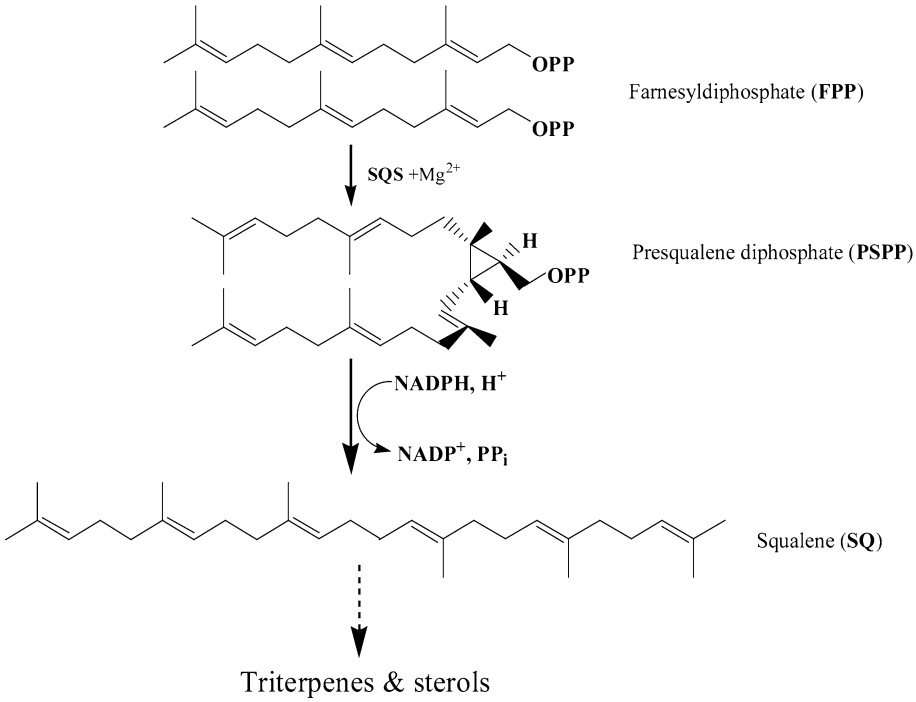


Fig. 1 Conversion of farnesyl diphosphate (FPP) to squalene (SQ)

been dedicated to the development of efficient synthesis of sterol compounds such as squalene, lanosterol, zymosterol, ergosterol [2]. As secondary metabolites, triterpenes and sterols are synthesized via the mevalonate pathway, which contains the sequential conversion of farnesyl diphosphate (FPP) to squalene (SQ) and then to 2,3-oxidosqualene, followed by a series of cyclization, oxidation, and reduction reactions [7, 19, 48].

Squalene synthase (SQS; EC 2.5.1.21), an enzyme of isoprenoid biosynthesis enzymes Class 1 superfamily, combines two molecules of FPP to form presqualene diphosphate (PSPP) and further to SQ by subsequent rearrangement and NADPH-dependent reduction step (Fig. 1) [6, 26]. These reactions, as the committed steps in triterpenoid and sterols biosynthesis, play an important role in the regulation of isoprenoid biosynthesis in eukaryotes [11, 26]. In addition, NADPH serves as an indispensable cofactor in SQ synthesis. To illustrate the function of NADPH in the reaction, two molecules of FPP and NADPH were added in the reaction buffer, and the reaction product was investigated for yeast SQS [25, 31]. When NADPH is present in the reaction buffer, PSPP is transformed directly to SQ. However, without adding NADPH cofactor, PSPP accumulated and released from the active site [15].

Mortierella alpine, an industrial strain often used to produce unsaturated fatty acid such as γ-linolenic acid and arachidonic acid [27, 39], is capable of synthesizing 13 kinds of important sterols, such as cholesta-5, 24-dienol (desmosterol), 24-methyl-desmosterol [8]. However, the gene involved in SQS synthesis in *M. alpine* has not been identified and characterized. Recently, the whole genome sequence of *M. alpine* (ATCC 32222) has been announced and characterized by our laboratory [47]. Our analysis suggested that there is a putative SQS gene in *M. alpina*. Therefore, it is of particular interest to identify the enzymatic processes for sterols synthesis.

As a key enzyme in the isoprenoid biosynthesis, SQS has been already found in many organisms. However, no research has been done on the SQS gene from *M. alpine* so far. In this work, a putative fungal SQS from *M. alpine* was identified, and characterized biochemically and biophysically. The soluble protein was obtained by truncating 17 or 27 amino acids in the C-terminal and the recombinant protein was expressed, purified and detected by in vitro activity investigation. The reaction conditions including pH and temperature were optimized. Based on site-directed mutagenesis, the mutant enzyme mMaSQSΔC17 (E186K) exhibited an improved catalytic efficiency. The characterized SQS in this work would provide potential application in synthetic biology to engineer a pathway in *E. coli* for production of important terpenoids and sterols.

Materials and methods

Chemicals

Standard FPP, SQ and NADPH were purchased from Sigma-Aldrich (St. Louis, MO, USA). All other chemicals used were of high analytical grade and commercially available.

Fungal strain and culture conditions

Mortierella alpine with the accession number American Type Culture Collection Catalog 32222 was cultured and harvested as described by Wang et al. [46]. In detail, the strain was initially cultured on potato dextrose agar at 25 °C for approximately 5–7 days. When the growth was visible, the culture was inoculated into 200 mL of Kendrick medium containing 100 g/L glucose, 5 g/L yeast extract, 3 g/L $NaNO_3$, 1 g/L KH_2PO_4, and 0.5 g/L $MgSO_4{\cdot}7H_2O$, and incubated at 25 °C for 6 days with shaking. The mycelia were then collected by filtration through sterile cheesecloth and frozen immediately in liquid nitrogen for RNA extraction.

Construction of the fungal cDNA

Total RNA was extracted using Trizol Reagent (Invitrogen). RNA was subjected to RNase-free DNase digestion, and then purified using the RNeasy Mini kit (Qiagen) according to the manufacturer's instructions. The total RNA was reverse transcribed with the PrimeScript RT reagent kit (Takara Bio, Inc.) following the manufacturer's instructions, before the PCR amplification of SQS gene.

Cloning and site-directed mutagenesis of *M. alpine* SQS gene

Genes encoding potential SQS were identified using BLASTP (http://www.ncbi.nlm.nih.gov/blast). The NCBI accession number obtained was KT318395. The full length of *M. alpine* SQS gene was amplified by PCR using the oligonucleotide primers SQS-F and SQS-R. The primers SQS-F and SQSΔC17-R/SQSΔC27-R were designed to generate truncated sequence (MaSQSΔC17/MaSQSΔC27) by removing 17 or 27 amino acids from the C terminus. The PCR products were inserted into pET28a(+) expression vector (Novagen), generating plasmids pET-MaSQS, pET-MaSQSΔC17 and pET-MaSQSΔC27, which were sequenced to confirm the correct open reading frame inserted. To generate the truncated mutant mMaSQSΔC17 (E186K), the site-directed mutagenesis based on overlapped PCR was employed. Consequently, pET-MaSQSΔC17 (E186K) was constructed and further verified by sequencing. The expression plasmids were transferred into *E. coli* BL21(DE3) strain (Novagen) by electroporation. All designed primers are listed in Table 1.

Protein expression and purification

The *E. coli* BL21(DE3) strain carrying the corresponding plasmid was grown at 37 °C in LB medium containing 50 μg/mL kanamycin to OD_{600} of 0.6, at which point expression was induced by addition of 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and incubated for 20 h at 16 °C, 90 rpm. The collected cultures were centrifuged at 7000*g* for 5 min and washed with 50 mM Tris–HCl buffer, pH 8.0. After cell lysis by sonication, the supernatant was loaded onto a nickel ion affinity chromatography column (GE Healthcare) and purified according to the manufacturer's instructions. The protein concentration was determined by the Bradford method using BSA. The purified SQS was analyzed by 12 % SDS-PAGE and high-performance liquid chromatography–mass spectrometry (HPLC–MS) (Thermo).

Table 1 Primers used in this study

Primer name	Sequence
SQS-F	GGAATTC**CATATG**ATGGCTTCTGCTAT CCTCGCCT
SQS-R	CCC**AAGCTT**TTATGCGAGCGCATTGTT GATA
SQSΔC17-R	CCC**AAGCTT**TTAAAGCACGGTTCCG GTTTTG
SQSΔC27-R	CCC**AAGCTT**TTATCCAGCTTGCAT GCGCT
SQSΔC17(E186K)-F	CTGGGATTGAGCAAGATGTTCAGCGCC
SQSΔC17(E186K)-R	GGCGCTGAACATCTTGCTCAATCCCAG

Restriction cleavage sites are bold

Determination of enzymatic activity

The SQS activity of the recombinant protein was measured on the basis of the conversion of FPP to SQ. SQS was assayed by modification of the procedure reported by Lee, Poulter [21]. The assay was performed in 50 mM MOPS buffer (pH 7.0), containing 40 μM FPP, 30 μg purified SQS, 10 mM $MgCl_2$, 1 mM NADPH, 1 mM dithiothreitol (DTT), 0.2 mg BSA, in a final volume of 200 μL. The reaction was carried out at 37 °C for 1 h and was slightly shaken every 10 min. The reaction was initiated by adding FPP to the assay cocktail. In the negative control, the SQS solution was replaced by assay buffer. After incubation, the reaction mixture was terminated and extracted with 200 μL n-hexane. The n-hexane solution was dried by superfluous sodium sulfates anhydrous, and was then concentrated at room temperature until the total volume reached about 20 μL. The catalytic product SQ (1 μL) was detected by gas chromatography–mass spectrometry (GC–MS). Diluted standard SQ (1 μL) was directly subjected to GC–MS detection. The activity was determined by monitoring the conversion of NADPH to NADP during enzyme–substrate reaction. One unit of enzyme is defined as the amount of the enzyme that liberates 1 μmol of NADP per minute under the assay conditions.

SQ detection by GC–MS

GC–MS (QP2010, Shimadzu, Kyoto, Japan) equipped with an Rtx-5 Amine column was operated at an ionization voltage of 70 eV with a scan range of m/z 50–450. The oven temperature was programmed at an initial temperature of 150 °C for 5 min; then increased to 280 °C at the rate of 15 °C/min and maintained for 7 min with helium. To protect GC–MS detector, the compounds were exposed to the detector after 12.5 min so that the impurities were excluded. The mass spectra of eluted peaks in the reaction were compared with standard SQ.

Temperature and pH profiles of MaSQS activity

Assays were performed in 50 mM MOPS buffer, containing 40 μM FPP, 30 μg of SQS, 10 mM $MgCl_2$, 1 mM NADPH, 1 mM DTT, 1 mg/mL BSA. The reactions were performed as described for the standard assay except that the total volume of reaction mixture was changed to 50 μL. For determination of the optimum pH, enzyme activities were measured with pH ranging from 6.0 to 9.0 at 37 °C. In the case of the optimum temperature determination, SQS reactions (pH 7.2) were carried out at 15, 25, 30, 37, 40, 45, 55 °C, respectively. These experiments were replicated at least three times.

Measurement of kinetic parameters

To measure the K_m, V_{max} and k_{cat} values of FPP, the assay buffer contained 1.0–10 μM FPP at a fixed concentration of NADPH (1 mM). To measure the K_m and V_{max} values of NADPH, the assay buffer contained 50–300 μM NADPH at a fixed concentration of FPP (10 μM). Assays were performed in 50 mM MOPS buffer, pH 7.2, containing 10 mM $MgCl_2$, 1 mM DTT, 1 mg/mL BSA, 100 nM enzyme in a total volume of 50 μL at 37 °C for one minute. The conversion of NADPH to NADP was monitored by capillary electrophoresis (CE). The kinetic parameters were determined by fitting the initial rates of reaction measured at various concentrations of one substrate at a fixed concentration of the other to the Lineweaver–Burk equation using linear regression.

Results

Bioinformatic analysis

Mortierella alpine is a filamentous fungus commonly found in soil which is able to produce lipids in the form of triacylglycerols that account for up to 50 % of its dry weight. Analysis of the *M. alpina* genome suggests that there is an SQS gene involved in sterols synthesis of interest. Thus, this sequence was chosen for further investigation.

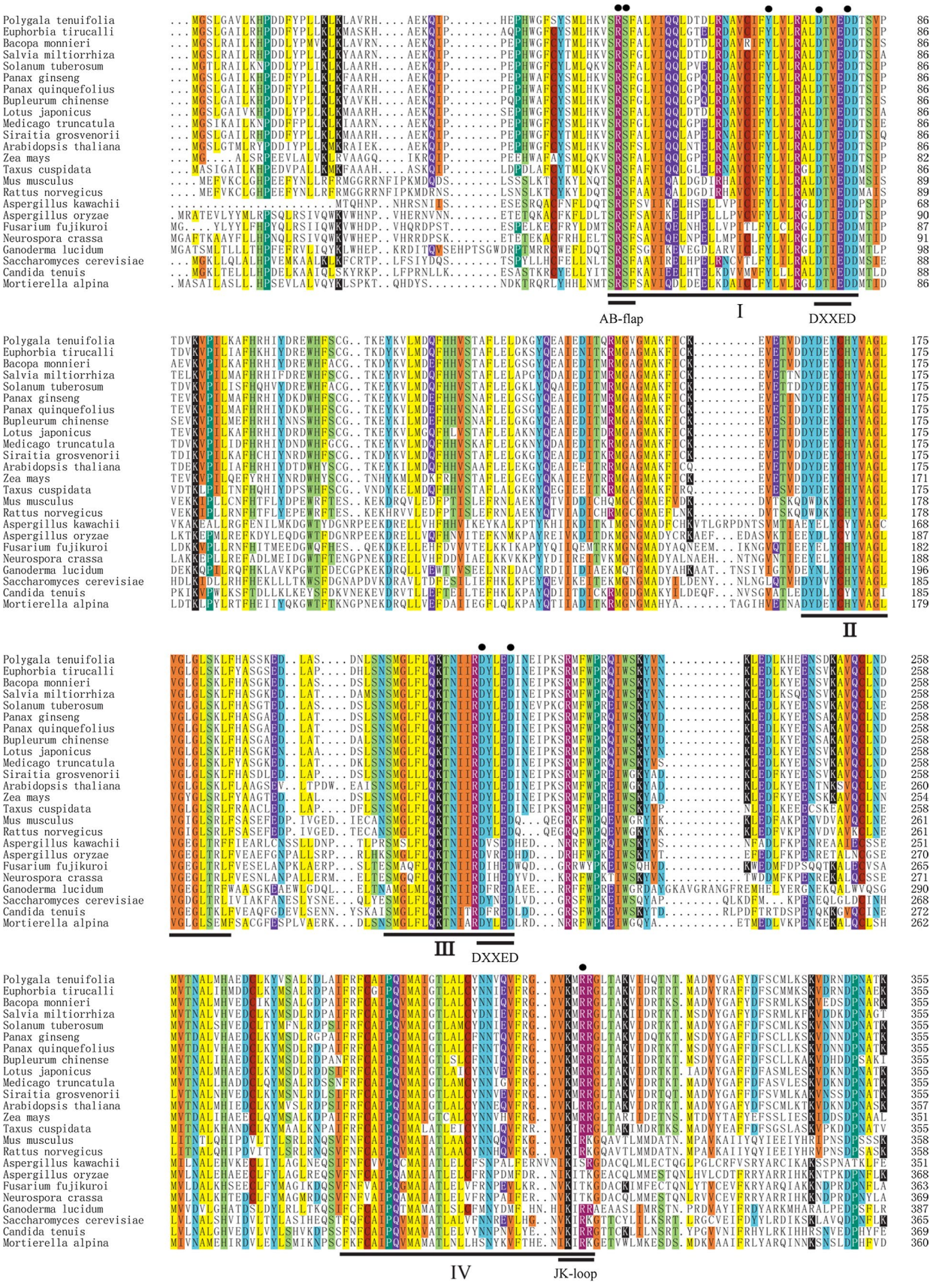
Fig. 2 Alignment of SQS proteins from *M. alpine* and other representative SQS proteins from animals, plants and fungi. Four conserved regions (I–IV) are defined. Two DXXED motifs as the substrate binding sites are shown. The *circles* indicate residues involved in NADPH recognition. The AB-flap and JK-loop are indicated below the sequences ▶

The full length of MaSQS is predicted to encode a protein of 416 amino acid residues with a MW of 47.0 kDa and theoretical pI of 5.93. To predict domain, we made a comparison between the candidate amino acid sequence and biochemically characterized SQSs from fungi, animals and plants. Multiple sequence alignment of MaSQS and these SQS proteins indicated that MaSQS shares a high amino acid identity (ranging from 42 to 51 %) with known SQS sequences (Fig. 2). It could be seen that there existed four highly conserved regions between *M. alpine* and other organisms (animals, plants and fungi). The amino acid sequence of MaSQS shares 48 and 49 % identity with *Mus musculus* and *Rattus norveqicus* SQS proteins, respectively, which is a little high compared with that from plants, and is most similar (51 %) to that from *Neurospora crassa*. Interestingly, the AB-flap and the JK-loop were also included, which were assumed to regulate the binding of substrates (prenyl donor and acceptor) and NADPH, respectively. The amino acid sequence alignment also revealed several distinctive domains distributed between the conserved domains.

A phylogenetic tree (Fig. 3) analysis suggested that the MaSQS shared the closest relationship with *Ganoderma lucidum* and *Aspergillus*, which also belonged to the member of the fungus. Besides, different forms of SQS could evolve from a single ancestral gene. Furthermore, three-dimensional models' prediction for MaSQS was performed by SWISS-MODEL Workspace Web Server using crystal structure of human SQS (PDB ID code 1EZFB) as template. The result showed that SQS protein was mainly consisted of fourteen α-helices and two β-pleated sheets. The main α-helices are numbered with A–N from N-terminal to C-terminal (Fig. 4). To understand the function of this enzyme, we cloned and purified the protein and performed in vitro catalytic assay.

Cloning, solubilization, and purification of MaSQS

Initially, full-length MaSQS was found exclusively in inclusion bodies as an inactive enzyme. The soluble recombinant MaSQS could not be obtained. Therefore, we analyzed the *M. alpine* SQS protein sequence using TMHMM software (http://www.cbs.dtu.dk/services/TMHMM/). The result indicated that 17 amino acids of the C-terminal of *M. alpine* SQS were assumed to locate the membrane-spanning


AB-flap
I
DXXED
II
III
DXXED
IV
JK-loop

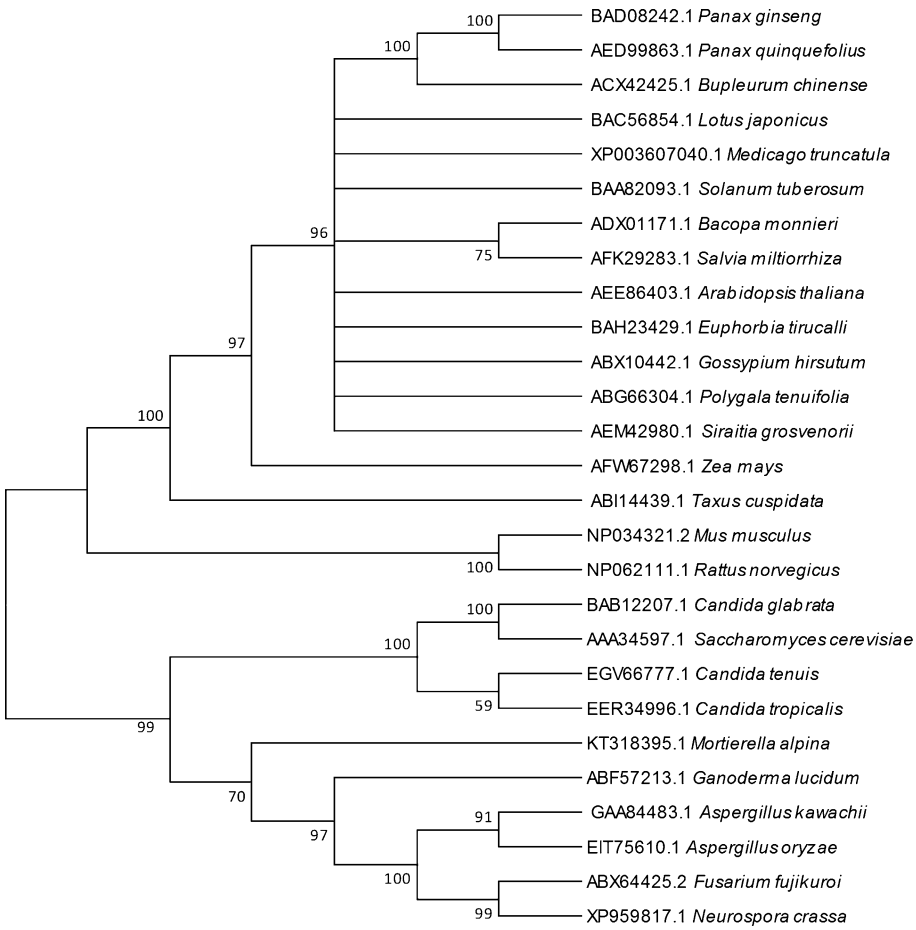


Fig. 3 Phylogenetic analysis of SQS amino acid sequences using the neighbor-joining (NJ) method. Twenty-four squalene synthase sequences from different species were conducted from GenBank. The accession numbers are indicated in the *front* of each species. *Numbers at nodes* represent the percentage bootstrap values based on 1000 replicates. Only values above 50 % are considered significant and are shown here

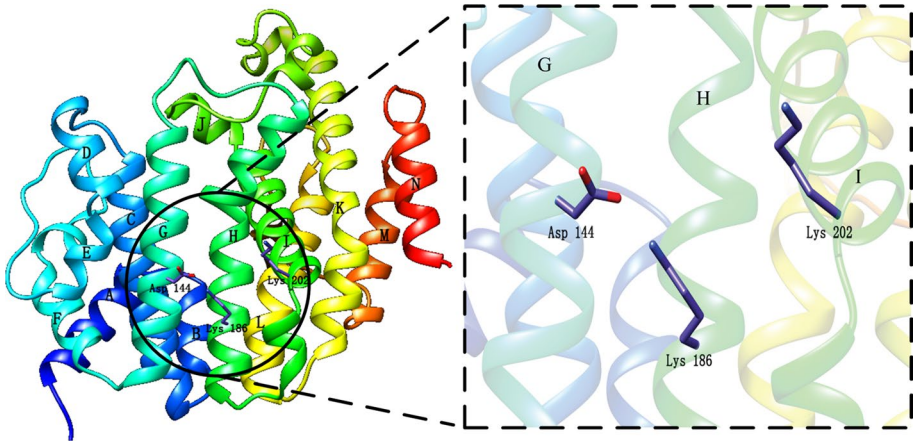


Fig. 4 Cartoon display of the 3-D structure of MaSQS predicted by SWISS-MODEL Workspace Web Server using crystal structure of human squalene synthase (Database PDB: 1EZFB) as template. Modeled residue range was 32–370. The figure was made using Chimera (http://www.cgl.ucsf.edu/chimera). The main α-helices are numbered with A–N from N-terminal to C-terminal

α-helix domain. Therefore, these 17 amino acids of the C-terminal were removed from the protein and the resulting plasmid pET-MaSQSΔC17 was constructed to express truncated MaSQS. The recombinant enzyme was expressed as His-tagged fusion proteins in *E. coli* BL21(DE3) induced by IPTG, and purified to near homogeneity by nickel ion affinity chromatography. As expected, the expressed protein was almost expressed in soluble form. The purified MaSQSΔC17 was detected by 12 % SDS-PAGE and MW of purified protein was about 46 kDa (Fig. 5), which corresponded well to the theoretically calculated mass of 47 kDa based on the 399 amino acids and 6 histidine residues. In addition, 27 amino acids of the C-terminal were deleted and the purified recombinant protein also exhibited the soluble form (Fig. 5), indicating that these amino acids in the C-terminal played a pivotal role in the water solubility

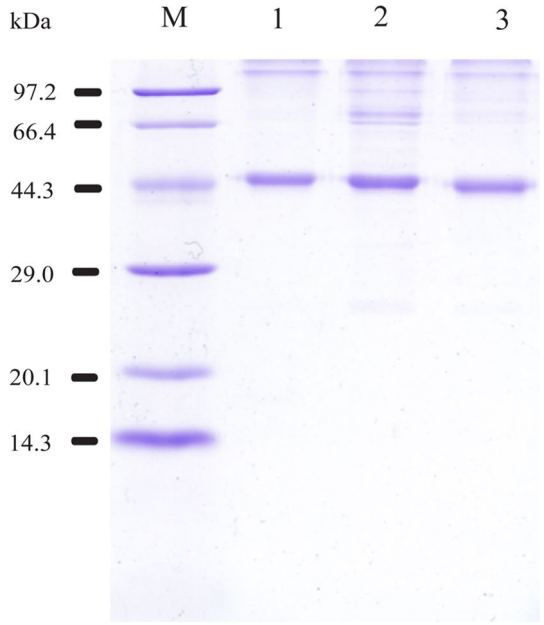


Fig. 5 SDS-PAGE analysis of the purified truncated squalene synthases and corresponding mutant from *M. alpine*. *Lane M*, protein marker; *lane 1*, MaSQSΔC17; *lane 2*, mMaSQSΔC17; *lane 3*, MaSQSΔC27

or inclusion bodies. Moreover, the optimal expression for both MaSQSΔC17 and MaSQSΔC27 was achieved at 0.1 mM IPTG at 16 °C, 90 rpm for 20 h. Besides, the yield of purified MaSQSΔC17 and MaSQSΔC27 was 2.28 and 1.80 mg/mL, respectively.

Activity of MaSQSΔC17 and MaSQSΔC27

FPP substrate was incubated with the purified MaSQSΔC17 (MaSQSΔC27), then the reaction products were detected by GC–MS, as shown in Fig. 6. No such product peak was detected in the negative control (Fig. 6a). MaSQSΔC17 generated a significant product peak with a GC retention time of 16.5 min (Fig. 6c), corresponding to the position of the authentic SQ sample (Fig. 6b). Besides, the fragmentation patterns of catalytic products (Fig. 6e) and authentic SQ (Fig. 6d) were consistent. A comparison of the retention time and the corresponding full-scan mass spectra of the samples with those of authentic SQ confirmed that the recombinant truncated enzyme from BL21(DE3)/pET-MaSQSΔC17 exerted the SQS activity. In addition to MaSQSΔC17, the recombinant protein MaSQSΔC27 also displayed the enzyme activity (Fig. S1).

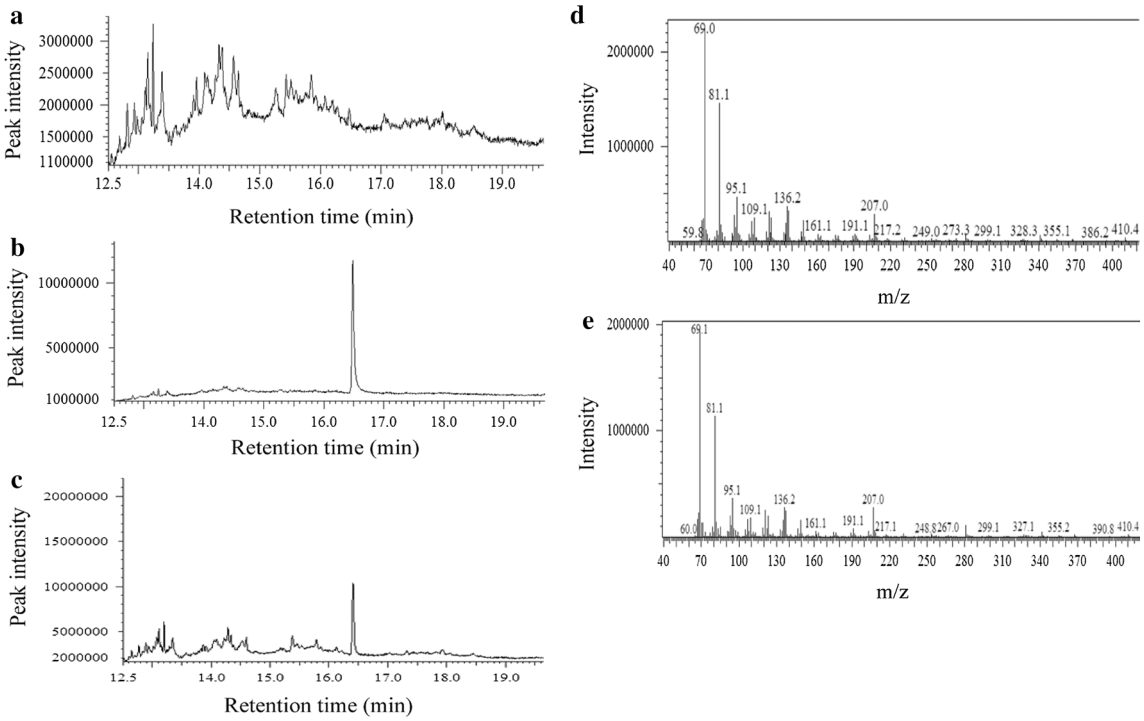


Fig. 6 GC–MS detection of the catalyzed product of MaSQSΔC17. **a** Total ion chromatograms of the sample extracted from an in vitro reaction mixture without squalene synthase (negative control). **b** Total ion chromatograms of authentic squalene. **c** Total ion chromatograms of squalene generated by MaSQSΔC17. **d** Fragmentation patterns of authentic squalene. **e** Fragmentation patterns of squalene generated by MaSQSΔC17

The specific activities of MaSQSΔC17 and MaSQSΔC27 are 35.5 U/mg and 30.1 U/mg. Therefore, the relatively higher activity of MaSQSΔC17 was further characterized biochemically.

Effects of temperature and pH on MaSQSΔC17 activity

Temperature and pH optima on MaSQS activity were investigated. As shown in Fig. 7a. MaSQSΔC17 was active at each of temperature tested (ranging from 15 to 55 °C), and achieved the greatest activity at 37 °C (Fig. 7a), a little higher than the optimal growth temperature of *M. alpine*. Although the activity decreased quickly from 37 to 55 °C, the enzyme still maintained 50 % specific activity upon incubation at 45 °C. In addition to temperature, MaSQSΔC17 showed highly active activity between pH 7.0 and 7.5, with an optimal activity at pH 7.2 in 50 mM MOPS buffer (Fig. 7b), which is consistent with the value reported for the yeast enzyme [51]. However, when pH was over 9.0 or below 6.0, approximately half of the enzyme activity lost.

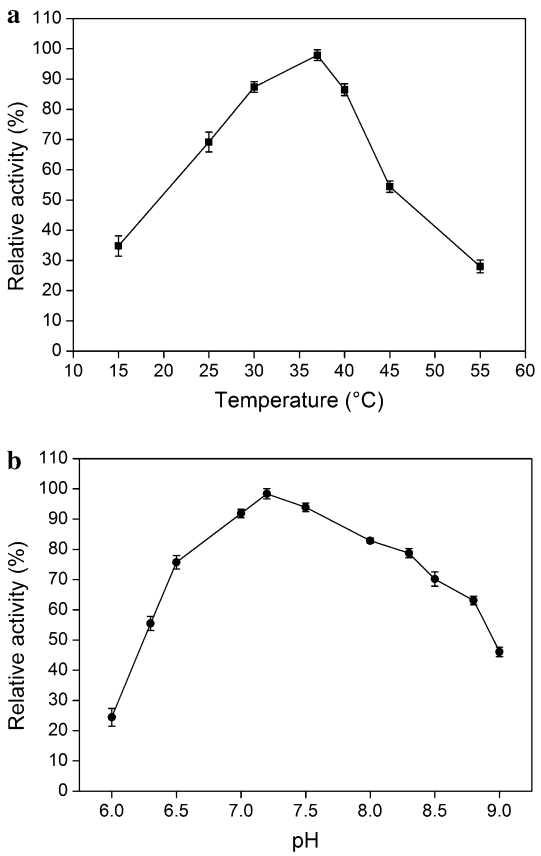


Fig. 7 Effects of temperature and pH on MaSQS activity. **a** Determination of the optimal temperature. **b** Determination of the optimal pH

Kinetic properties of MaSQSΔC17

Standard procedures were used to determine kinetic parameters. It should be noted that the kinetics of the reaction catalyzed by MaSQSΔC17 presented a good fit with the Michaelis–Menten model, and linear regression of the experimental data to the Lineweaver–Burk equation. As shown in Table 2, the K_m and V_{max} of MaSQSΔC17 for FPP were 5.20 μM and 1.68 nmol min^{-1} mg^{-1} protein, respectively. However, the K_m and V_{max} of MaSQSΔC17 for NADPH were 241.95 μM and 2.43 μM min^{-1} mg^{-1} protein, respectively. Moreover, the k_{cat}^{FPP} value of MaSQSΔC17 was 0.48 s^{-1} and the k_{cat}^{FPP}/K_m^{FPP} was 0.09 μM^{-1} s^{-1}.

Site-directed mutation of MaSQSΔC17

SQS had an obvious amino acid substitution at position E186 in region II while others are K or R (Fig. 2). Thus, E186K mutant enzyme (mMaSQSΔC17) was built based on overlapped PCR and the catalytic activity was investigated. The purified mMaSQSΔC17 was also detected by 12 % SDS-PAGE (Fig. 5), with the yield of 1.947 mg/mL. In vitro enzyme activity assay indicated that the mutant enzyme also exhibited the catalytic activity (Fig. S1). The K_m, k_{cat} and V_{max} of mMaSQSΔC17 for FPP were 2.68 μM, 0.83 s^{-1} and 2.55 μM min^{-1} mg^{-1} protein, respectively (Table 2). The kinetic data showed that site-directed mutation of MaSQSΔC17 resulted in a 72.9 % increase in k_{cat}^{FPP} and a 48.5 % decrease in K_m^{FPP}. The k_{cat}^{FPP}/K_m^{FPP} value for mutant was 0.31 μM^{-1} s^{-1}, 3.4-fold relative to the value of MaSQSΔC17. It was indicated that the mutant enzyme mMaSQSΔC17 exerted an enhanced catalytic efficiency compared to MaSQSΔC17.

Discussion

Mortierella alpine is an oleaginous fungus used in the industrial scale production of arachidonic acid. To investigate the potential strategies for enhanced lipid production, many strategies have been proposed, including the development of microbial fermentation processes, molecular breeding of strains by manipulating a variety of genes encoding fatty acid desaturases and elongases involved in polyunsaturated fatty acids biosynthesis [33], selection and characterization of promoters based on genomic approach for the molecular breeding [29], reconstruction and analysis of a genome-scale metabolic model [49]. Due to wide significance of *M. alpine* as an industrial fungus, it is tempting to decipher the biosynthetic pathway of triterpenes and sterols, but also exploit the full biotechnological potential of this fungus. SQ is a polyunsaturated, triterpenic hydrocarbon widely applied in nutritional [44], cosmetic

Table 2 Kinetic constants of squalene synthase from different species and sources

SQS source	SQS type	K_m^{FPP} (μM)	K_m^{NADPH} (μM)	V_{max}^{FPP} (nmol min^{-1} mg^{-1})	V_{max}^{NADPH} (nmol min^{-1} mg^{-1})	k_{cat} (s^{-1})	k_{cat}/K_m^{FPP} (μM^{-1} s^{-1})	Reference
Yeast	C-terminally truncated	2.5	530			0.53	0.21	[23]
Methylococcus capsulatus	Full length	13.6	75.4			0.55	0.04	[28]
Trypanosoma cruzi	Glycosomal	2.8	40	0.97	0.43			[43]
	Microsomal/mitochondrial	2.3	33	1.18	0.76			[43]
	N- and C-terminally truncated	5.3	23	1428.56	1853.24	1.05	0.20	[34]
Leishmania mexicana	Microsomal/mitochondrial	2.8	57	0.98	0.84			[43]
	Glycosomal	3.2	62	0.75	0.58			[43]
Thermosynechococcus elongatus	Full length	0.97	241			1.74	1.80	[21]
M. alpine	C-terminally truncated	5.20	241.95	1.68	2.43	0.48	0.09	This work
	Mutant mMaSQS (E186K)	2.68	168.34	2.55	3.32	0.83	0.31	This work

[13], pharmaceutical [9] and medical industry [32]. It has been reported that yeast can exclusively produce isoprenoids, including squalene, through manipulating the key enzymes consisting of acetyl-CoA carboxylase, HMG-CoA reductase, squalene epoxidase, squalene synthase, and ATP citrate lyase [45]. As a key precursor of isoprenoids, SQ is synthesized from FPP by SQS (Fig. 1), which has been identified in human beings [36, 38], animals [14, 24, 34], plants [11, 12, 18, 26, 42, 50], fungi [16, 20, 51, 52], bacteria [21]. SQS is commonly considered to be a crucial enzyme and a potential regulatory branch node that controls carbon flux towards the triterpenes and sterols biosynthesis. Besides, SQS from *Candida tropicalis*, which was cloned and expressed as recombinant protein in *Pichia pastoris*, has been identified as one of the immunogenic proteins that could be a potential diagnostic candidate for the pathogenic fungus *C. tropicalis* [20].

Multiple sequence alignment among the *M. alpine* SQS and other characterized SQS proteins revealed the presence of certain conserved motifs (Fig. 2), which are important for catalysis [10, 30]. Regions I, II, and III are involved in the first half-reaction, the condensation of two molecules of FPP to give PSPP [10]. Region I is highly conserved with an aspartate-rich motif (DXXED), which is considered to be responsible for the binding of the substrate FPP through the magnesium ion [5, 28]. Region II is thought to catalyze two molecules of FPP to produce PSPP. It should be noted that region II of *M. alpine* SQS has an obvious amino acid substitution at position E186 compared to K or R of others. Therefore, the site-directed mutagenesis of E186K was investigated in this work to confirm its key role in the enzyme catalysis. Region III is supposed to bind the diphosphate units in FPP via bridging magnesium ion [30]. Region IV is believed to be responsible for the catalysis of the second step for the rearrangement of PSPP to SQ with an NADPH binding motif [10, 21, 30]. Besides, although amino acid residues in the C-terminal region show a low level of sequence identity among all SQS proteins (Fig. 2), this region is very hydrophobic among all SQS enzymes and thus functions as an anchor in the putative endoplasmic reticulum membrane [53], which was also predicted by TMHMM and SPLIT 4.0 bioinformatics programs. Furthermore, the phylogenetic tree suggested that the fungal enzymes were clustered into one group (Fig. 3), which is in accordance with the classical taxonomy.

The gene encoding SQS was cloned from *M. alpine* in this study. However, the whole sequence resulted in inclusion bodies and did not generate soluble protein which hindered the enzyme activity. In prokaryotes, SQS was often water soluble and did not require any detergent for its higher activity [28]. Nevertheless, eukaryotic SQSs are associated with microsomes [11, 16, 23, 24, 26], causing the protein resistant to solubilization. Therefore, the elimination of membrane-anchoring and organelle-targeting domains is beneficial for the expression of genes from higher eukaryotes [37]. According to the membrane-bound characterization of SQS, some work have attempted to design some truncated enzymes whose solubilization and activity were significantly improved. The C-terminal deletion of residues resulted in functionally soluble SQS enzyme activity and SQS protein accumulation in capsicum, tobacco and yeast. In addition to generating soluble SQSs, the genetic truncation of the hydrophobic region retains most of the enzyme activity [23, 51]. For example, enzymatic truncation of rat SQS revealed that deletion of 33 amino acids in the N-terminal did not reduce the enzyme

activity despite that the N-terminal domains of SQS were not hydrophobic [35]. In our study, 17 or 27 amino acids in the C-terminal were deleted and the soluble proteins were obtained, indicating the importance of correct protein folding and confirmation. In vitro activity analysis by GC–MS showed that the truncated MaSQSΔC17 or MaSQSΔC27 could synthesize the formation of SQ from FPP (Fig. 6 and Fig. S1), suggesting that the truncated enzyme exhibited a pivotal role in maintaining the catalytic activity and stability. Although further truncation was not investigated for MaSQS, it was predicted that three additional helixes and partial JK-loop removal in the C-terminal regions could lead to the inactivity of SQS due to the loss of NADPH binding (Fig. 2), which influence the dehydrosqualene synthase activity [17]. Moreover, the truncated MaSQSΔC17 was subjected to site-directed mutagenesis and the mutant mMaSQSΔC17 displayed an improved enzyme activity, indicating that the site E186 played a significant role in the catalytic ability. In addition, the comparison of the enzyme kinetics parameter k_{cat}^{FPP}/K_m^{FPP} value between MaSQSΔC17 and mMaSQSΔC17 revealed that the mutant enzyme exerted an increased catalytic efficiency. Different from other known characterized fungal SQS, MaSQSΔC17 has an optimal temperature at 37 °C and shows a broader stability range from 30 to 40 °C (Fig. 7a). In contrast, the optimal pH of MaSQSΔC17 for enzyme activity is similar to that of other fungal SQS [21], with a similar pH stability range (pH 6.5–8.5), especially in alkaline environment (Fig. 7b). In addition, MaSQSΔC17 has a strict dependent on Mg^{2+}, with the maximal activity at 20 mM (Fig. S2). However, no activity was observed without supplementing Mg^{2+}, which was in agreement with Lee, Poulter [21]. This characterization is typical for enzymes in the isoprenoid biosynthetic pathway utilizing diphosphate substrates.

The functional changes of this mutation can be modeled since a reliable 3D model is available. Previous investigation has characterized structure–function maps for SQS to identify the functional residues and/or domains responsible for botryococcene and SQ biosynthesis in the green alga *Botryococcus braunii* race B in the catalytic cascade [3]. Recently, the structures of human SQS and its mutants in complex with several substrate analogs and intermediates coordinated with Mg^{2+} or Mn^{2+} have been presented, and four major steps (substrate binding, condensation, intermediate formation and translocation) of the ordered sequential mechanisms involved in the isoprenoid biosynthetic pathway have been identified [22]. In this study, cartoon display of the 3-D structure of MaSQS has been predicted by SWISS-MODEL Workspace Web Server using crystal structure of human SQS (PDB ID code 1EZFB) as template (Fig. 4). The change of amino acid from glutamic to lysine results in the substitution of a negatively charged branched side chain with a positively charged branched side chain. Amino acid site 186 is located within the C-terminal part of α-helix-H which contains the catalytic center. Besides, amino acid site 186 is close to the C-terminal part of α-helix-G containing the residue Asp-144 and C-terminal part of α-helix-I containing the residue Lys-202. The positively charged carboxyl side chain of Lys-186 is more strongly attracted to the negatively charged amino side chain of Asp-144 on the C-terminal part of α-helix-G and keeps at a distance with the residue Lys-202 on the C-terminal part of α-helix-I which is connected with substrate binding site (Asp-220 and Asp-224) on α-helix-J (Fig. 4). Thus, the increased activity of E186K was due to the changed enzyme structure which creates a bigger interspace for substrate binding site on α-helix-J to incorporate FPP. In addition, the enzyme kinetics also provided the evidence about an active site for SQS.

Different sources of SQS have been identified, in which some were cloned, purified, and kinetic properties were carefully studied [21, 23, 34, 41, 43]. However, except for yeast and *Thermosynechococcus elongatus*, most of them are from animal sources which are difficult to use in engineering application. In our study, the k_{cat}^{FPP}/K_m^{FPP} value of mMaSQSΔC17 was 0.31 $\mu M^{-1}\ s^{-1}$, which was lower than *T. elongates* whose k_{cat}^{FPP}/K_m^{FPP} value was 1.80 $\mu M^{-1}\ s^{-1}$ (Table 2). However, in order to obtain soluble active enzyme SQS cloned from *T. elongates*, high concentrations of glycerol (20–50 % vol/vol) had to be added to the disruption and purification buffers, which complicated protein purification process. Besides, the k_{cat}^{FPP}/K_m^{FPP} value of mMaSQSΔC17 was higher than yeast whose k_{cat}^{FPP}/K_m^{FPP} value was 0.21 $\mu M^{-1}\ s^{-1}$ [23], indicating the higher catalytic efficiency of mMaSQSΔC17 compared to SQS of yeast.

In conclusion, this is the first report on the gene cloning and characterization of SQS of *M. alpine*. The truncated SQS (MaSQSΔC17) possessed activity of converting two molecules of FPP to SQ, while the mutant enzyme mMaSQSΔC17 via site-directed mutagenesis further exhibited an enhanced catalytic efficiency compared to MaSQSΔC17. Here, our results facilitate the investigation of isoprenoid biosynthesis in the fungus *M. alpine*. It raises a possibility that MaSQS can be potentially employed in synthetic biology to incorporate in a module of a chassis host such as *E. coli* for production of valuable terpenoids and sterols. Further efforts are underway to enhance the thermostability and catalytic efficiency by rational protein engineering. In addition, cofactor NADPH quantification has also been taken into account, which would be overexpressed to further enhance the SQS expression level. Moreover, SQS will be expressed in other expression systems like *Actinomycetes*/*Streptomyces*, *Pichia pastoris*

or *Saccharomyces,* etc., which probably will increase the expression level.

Acknowledgments This research was financially supported by the National 973 Project of China (Nos. 2012CB721101, 2012CB721001 and 2013CB733904), the National 863 Project of China (No. 2012AA022108), the Program of National Natural Science Foundation of China (Nos. 31270003, 31400081, 81471904, 31371259, 31270133 and 31470194), the Fundamental Research Funds for the Central Universities (No. 65141028) and Research Project of Chinese Ministry of Education (No. 113015A).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Ačimovič J, Rozman D (2013) Steroidal triterpenes of cholesterol synthesis. Molecules 18(4):4002–4017. doi:10.3390/molecules18044002
2. Bailey RB, Trueheart J, Madden KT (2010) Production of sterols in oleaginous yeast and fungi. United States patent 20100305341 A1 & US 8367395 B2
3. Bell SA, Niehaus TD, Nybo SE, Chappell J (2014) Structure-function mapping of key determinants for hydrocarbon biosynthesis by squalene and squalene synthase-like enzymes from the green alga *Botryococcus braunii* race B. Biochemistry 53(48):7570–7581. doi:10.1021/bi501264s
4. Benveniste P (2004) Biosynthesis and accumulation of sterols. Annu Rev Plant Biol 55(1):429–457. doi:10.1146/annurev.arplant.55.031903.141616
5. Bhat WW, Lattoo SK, Razdan S, Dhar N, Rana S, Dhar RS, Khan S, Vishwakarma RA (2012) Molecular cloning, bacterial expression and promoter analysis of squalene synthase from *Withania somnifera* (L.) Dunal. Gene 499(1):25–36. doi:10.1016/j.gene.2012.03.004
6. Blagg BS, Jarstfer MB, Rogers DH, Poulter CD (2002) Recombinant squalene synthase. A mechanism for the rearrangement of presqualene diphosphate to squalene. J Am Chem Soc 124(30):8846–8853. doi:10.1021/ja020411a
7. Darnet S, Rahier A (2004) Plant sterol biosynthesis: identification of two distinct families of sterol 4alpha-methyl oxidases. Biochem J 378:889–898. doi:10.1042/BJ20031572
8. David Nes W, Nichols SD (2006) Phytosterol biosynthesis pathway in *Mortierella alpina*. Phytochemistry 67(16):1716–1721. doi:10.1016/j.phytochem.2006.02.023
9. Fox C (2009) Squalene emulsions for parenteral vaccine and drug delivery. Molecules 14(9):3286–3312. doi:10.3390/molecules14093286
10. Gu P, Ishii Y, Spencer TA, Shechter I (1998) Function-structure studies and identification of three enzyme domains involved in the catalytic activity in rat hepatic squalene synthase. J Biol Chem 273(20):12515–12525. doi:10.1074/jbc.273.20.12515
11. Hanley KM, Nicolas O, Donaldson TB, Smith-Monroy C, Robinson GW, Hellmann GM (1996) Molecular cloning, in vitro expression and characterization of a plant squalene synthetase cDNA. Plant Mol Biol 30(6):1139–1151. doi:10.1007/BF00019548
12. Huang Z, Jiang K, Pi Y, Hou R, Liao Z, Cao Y, Han X, Wang Q, Sun X, Tang K (2007) Molecular cloning and characterization of the yew gene encoding squalene synthase from *Taxus cuspidata*. J Biochem Mol Biol 40(5):625–635
13. Huang ZR, Lin YK, Fang JY (2009) Biological and pharmacological activities of squalene and related compounds: potential uses in cosmetic dermatology. Molecules 14(1):540–554. doi:10.3390/molecules14010540
14. Inoue T, Osumi T, Hata S (1995) Molecular cloning and functional expression of a cDNA for mouse squalene synthase. Biochim Biophys Acta 1260(1):49–54. doi:10.1016/0167-4781(94)00178-6
15. Jarstfer MB, Zhang DL, Poulter CD (2002) Recombinant squalene synthase. Synthesis of non-head-to-tail isoprenoids in the absence of NADPH. J Am Chem Soc 124(30):8834–8845. doi:10.1021/ja020410i
16. Jennings SM, Tsay YH, Fisch TM, Robinson GW (1991) Molecular cloning and characterization of the yeast gene for squalene synthetase. Proc Natl Acad Sci USA 88(14):6038–6042
17. Katabami A, Li L, Iwasaki M, Furubayashi M, Saito K, Umeno D (2015) Production of squalene by squalene synthases and their truncated mutants in *Escherichia coli*. J Biosci Bioeng 119(2):165–171. doi:10.1016/j.jbiosc.2014.07.013
18. Kim OT, Seong NS, Kim MY, Hwang B (2005) Isolation and characterization of squalene synthase cDNA from *Centella asiatica* (L.) Urban. J Plant Biol 48(3):263–269. doi:10.1007/BF03030521
19. Lange BM, Rujan T, Martin W, Croteau R (2000) Isoprenoid biosynthesis: the evolution of two ancient and distinct pathways across genomes. Proc Natl Acad Sci USA 97(24):13172–13177. doi:10.1073/pnas.240454797
20. Lee PY, Yong VC, Rosli R, Gam LH, Chong PP (2014) Cloning, expression and purification of squalene synthase from *Candida tropicalis* in *Pichia pastoris*. Protein Expr Purif 94:15–21. doi:10.1016/j.pep.2013.10.012
21. Lee S, Poulter CD (2008) Cloning, solubilization, and characterization of squalene synthase from *Thermosynechococcus elongatus* BP-1. J Bacteriol 190(11):3808–3816. doi:10.1128/jb.01939-07
22. Liu C-I, Jeng W-Y, Chang W-J, Shih M-F, Ko T-P, Wang AH-J (2014) Structural insights into the catalytic mechanism of human squalene synthase. Acta Crystallogr Sect D Biol Crystallogr 70(2):231–241. doi:10.1107/S1399004713026230
23. LoGrasso PV, Soltis DA, Boettcher BR (1993) Overexpression, purification, and kinetic characterization of a carboxyl-terminal-truncated yeast squalene synthetase. Arch Biochem Biophys 307(1):193–199. doi:10.1006/abbi.1993.1578
24. McKenzie TL, Jiang G, Straubhaar JR, Conrad DG, Shechter I (1992) Molecular cloning, expression, and characterization of the cDNA for the rat hepatic squalene synthase. J Biol Chem 267(30):21368–21374
25. Mookhtiar KA, Kalinowski SS, Zhang D, Poulter CD (1994) Yeast squalene synthase. A mechanism for addition of substrates and activation by NADPH. J Biol Chem 269(15):11201–11207
26. Nakashima T, Inoue T, Oka A, Nishino T, Osumi T, Hata S (1995) Cloning, expression, and characterization of cDNAs encoding *Arabidopsis thaliana* squalene synthase. Proc Natl Acad Sci USA 92(6):2328–2332
27. Nie Z-K, Deng Z-T, Zhang A-H, Ji X-J, Huang H (2013) Efficient arachidonic acid-rich oil production by *Mortierella alpina* through a three-stage fermentation strategy. Bioprocess Biosystems Eng 37(3):505–511. doi:10.1007/s00449-013-1015-2
28. Ohtake K, Saito N, Shibuya S, Kobayashi W, Amano R, Hirai T, Sasaki S, Nakano C, Hoshino T (2014) Biochemical characterization of the water-soluble squalene synthase from *Methylococcus capsulatus* and the functional analyses of its two DXXD(E)D motifs and the highly conserved aromatic amino acid residues. FEBS J 281(24):5479–5497. doi:10.1111/febs.13090

29. Okuda T, Ando A, Sakuradani E, Kikukawa H, Kamada N, Ochiai M, Shima J, Ogawa J (2014) Selection and characterization of promoters based on genomic approach for the molecular breeding of oleaginous fungus *Mortierella alpina* 1S-4. Curr Genet 60(3):183–191. doi:10.1007/s00294-014-0423-4
30. Pandit J, Danley DE, Schulte GK, Mazzalupo S, Pauly TA, Hayward CM, Hamanaka ES, Thompson JF, Harwood HJ Jr (2000) Crystal structure of human squalene synthase. A key enzyme in cholesterol biosynthesis. J Biol Chem 275(39):30610–30617. doi:10.1074/jbc.M004132200
31. Radisky ES, Poulter CD (2000) Squalene synthase: steady-state, pre-steady-state, and isotope-trapping studies. Biochemistry 39(7):1748–1760. doi:10.1021/bi9915014
32. Reddy LH, Couvreur P (2009) Squalene: a natural triterpene for use in disease management and therapy. Adv Drug Del Rev 61(15):1412–1426. doi:10.1016/j.addr.2009.09.005
33. Sakuradani E, Ando A, Ogawa J, Shimizu S (2009) Improved production of various polyunsaturated fatty acids through filamentous fungus *Mortierella alpina* breeding. Appl Microbiol Biotechnol 84(1):1–10. doi:10.1007/s00253-009-2076-7
34. Sealey-Cardona M, Cammerer S, Jones S, Ruiz-Perez LM, Brun R, Gilbert IH, Urbina JA, Gonzalez-Pacanowska D (2007) Kinetic characterization of squalene synthase from *Trypanosoma cruzi*: selective inhibition by quinuclidine derivatives. Antimicrob Agents Chemother 51(6):2123–2129. doi:10.1128/aac.01454-06
35. Shechter I, Klinger E, Rucker ML, Engstrom RG, Spirito JA, Islam MA, Boettcher BR, Weinstein DB (1992) Solubilization, purification, and characterization of a truncated form of rat hepatic squalene synthetase. J Biol Chem 267(12):8628–8635
36. Soltis DA, McMahon G, Caplan SL, Dudas DA, Chamberlin HA, Vattay A, Dottavio D, Rucker ML, Engstrom RG, Cornell-Kennon SA, Boettcher BR (1995) Expression, purification, and characterization of the human squalene synthase: use of yeast and baculoviral systems. Arch Biochem Biophys 316(2):713–723. doi:10.1006/abbi.1995.1095
37. Stevens D, Wang X-Z, Rice J, Nye B, Broadwell D, Glassbrook N, Sevala V, Crawford J, Stewart S (2003) Methods for determining squalene synthase activity. United States Patent 20030157583
38. Summers C, Karst F, Charles AD (1993) Cloning, expression and characterisation of the cDNA encoding human hepatic squalene synthase, and its relationship to phytoene synthase. Gene 136(1-2Che):185–192. doi:10.1016/0378-1119(93)90462-C
39. Takeno S, Sakuradani E, Tomi A, Inohara-Ochiai M, Kawashima H, Shimizu S (2005) Transformation of oil-producing fungus, *Mortierella alpina* 1S-4, using Zeocin, and application to arachidonic acid production. J Biosci Bioeng 100(6):617–622. doi:10.1263/jbb.100.617
40. Thimmappa R, Geisler K, Louveau T, O'Maille P, Osbourn A (2014) Triterpene biosynthesis in plants. Annu Rev Plant Biol 65(1):225–257. doi:10.1146/annurev-arplant-050312-120229
41. Thompson JF, Danley DE, Mazzalupo S, Milos PM, Lira ME, Harwood HJ Jr (1998) Truncation of human squalene synthase yields active, crystallizable protein. Arch Biochem Biophys 350(2):283–290. doi:10.1006/abbi.1997.0502
42. Uchida H, Yamashita H, Kajikawa M, Ohyama K, Nakayachi O, Sugiyama R, Yamato KT, Muranaka T, Fukuzawa H, Takemura M, Ohyama K (2009) Cloning and characterization of a squalene synthase gene from a petroleum plant, *Euphorbia tirucalli* L. Planta 229(6):1243–1252. doi:10.1007/s00425-009-0906-6
43. Urbina JA, Concepcion JL, Rangel S, Visbal G, Lira R (2002) Squalene synthase as a chemotherapeutic target in *Trypanosoma cruzi* and *Leishmania mexicana*. Mol Biochem Parasitol 125(1–2):35–45. doi:10.1016/S0166-6851(02)00206-2
44. Uylaşer V, Yildiz G (2014) The historical development and nutritional importance of olive and olive oil constituted an important part of the Mediterranean diet. Crit Rev Food Sci Nutr 54(8):1092–1101. doi:10.1080/10408398.2011.626874
45. Walker KA, Knuth ME, Fong NM, Beetham PR (2009) Production of squalene using yeast. United States Patent 20090298150 A1
46. Wang H, Chen H, Hao G, Yang B, Feng Y, Wang Y, Feng L, Zhao J, Song Y, Zhang H, Chen YQ, Wang L, Chen W (2013) Role of the phenylalanine-hydroxylating system in aromatic substance degradation and lipid metabolism in the oleaginous fungus *Mortierella alpina*. Appl Environ Microbiol 79(10):3225–3233. doi:10.1128/AEM.00238-13
47. Wang L, Chen W, Feng Y, Ren Y, Gu Z, Chen H, Wang H, Thomas MJ, Zhang B, Berquin IM, Li Y, Wu J, Zhang H, Song Y, Liu X, Norris JS, Wang S, Du P, Shen J, Wang N, Yang Y, Wang W, Feng L, Ratledge C, Chen YQ (2011) Genome characterization of the oleaginous fungus *Mortierella alpina*. PLoS One 6(12):e28319. doi:10.1371/journal.pone.0028319
48. Withers S, Keasling J (2007) Biosynthesis and engineering of isoprenoid small molecules. Appl Microbiol Biotechnol 73(5):980–990. doi:10.1007/s00253-006-0593-1
49. Ye C, Xu N, Chen H, Chen YQ, Chen W, Liu L (2015) Reconstruction and analysis of a genome-scale metabolic model of the oleaginous fungus *Mortierella alpina*. BMC Syst Biol 9:1. doi:10.1186/s12918-014-0137-8
50. Ye Y, Wang R, Jin L, Shen J, Li X, Yang T, Zhou M, Yang Z, Chen Y (2014) Molecular cloning and differential expression analysis of a squalene synthase gene from *Dioscorea zingiberensis*, an important pharmaceutical plant. Mol Biol Rep 41(9):6097–6104. doi:10.1007/s11033-014-3487-9
51. Zhang D, Jennings SM, Robinson GW, Poulter CD (1993) Yeast squalene synthase: expression, purification, and characterization of soluble recombinant enzyme. Arch Biochem Biophys 304(1):133–143. doi:10.1006/abbi.1993.1331
52. Zhao MW, Liang WQ, Zhang DB, Wang N, Wang CG, Pan YJ (2007) Cloning and characterization of squalene synthase (SQS) gene from *Ganoderma lucidum*. J Microbiol Biotechnol 17(7):1106–1112
53. Zhao RY, Xiao W, Cheng HL, Zhu P, Cheng KD (2010) Cloning and characterization of squalene synthase gene from *Fusarium fujikuroi* (Saw.) Wr. J Ind Microbiol Biotechnol 37(11):1171–1182. doi:10.1007/s10295-010-0764-z