

Activation-Independent Cyclization of Monoterpenoids

Gabriele Siedenburg,^a Dieter Jendrossek,^a Michael Breuer,^b Benjamin Juhl,^c Jürgen Pleiss,^c Miriam Seitz,^c Janosch Klebensberger,^c and Bernhard Hauer^c

Institute of Microbiology, University of Stuttgart, Stuttgart, Germany^a; Fine Chemicals and Biocatalysis Research, BASF SE, Ludwigshafen, Germany^b; and Institute of Technical Biochemistry, University of Stuttgart, Stuttgart, Germany^c

The biosynthesis of cyclic monoterpenes (C_{10}) generally requires the cyclization of an activated linear precursor (geranyldiphosphate) by specific terpene cyclases. Cyclic triterpenes (C_{30}), on the other hand, originate from the linear precursor squalene by the action of squalene-hopene cyclases (SHCs) or oxidosqualene cyclases (OSCs). Here, we report a novel terpene cyclase from *Zymomonas mobilis* (ZMO1548-Shc) with the unique capability to cyclize citronellal to isopulegol. To our knowledge, ZMO1548-Shc is the first biocatalyst with diphosphate-independent monoterpene cyclase activity. A combinatorial approach using site-directed mutagenesis and modeling of the active site with a bound substrate revealed that the cyclization of citronellal proceeds via a different mechanism than that of the cyclization of squalene.

Terpenoids are widespread secondary metabolites present in nearly all organisms. They have important or even essential functions in metabolism and act, for example, as membrane stabilizers, hormones, vitamins, or photoactive compounds. Terpenes are composed of isoprene units that are synthesized either via the mevalonate pathway (13) or via the 1-deoxy-D-xylulose-5-phosphate (DXP) route (12). The common linear precursor of all monoterpenes and higher terpenoids is geranyldiphosphate (30), independent from the pathway used for the synthesis of the isoprene backbone (isopentenylidiphosphate). Terpenoids are present in nature in a large variety of both linear and cyclic forms. The formation of the ring structure of cyclic terpenoids from diphosphate-activated linear precursors is catalyzed by terpene cyclases (terpene synthases) that are specific for either geranyldiphosphate (C_{10}) (monoterpene cyclases), farnesylidiphosphate (C_{15}) (sesquiterpene cyclases), or geranyl-geranyldiphosphate (C_{20}) (diterpene cyclases). Activation by diphosphate is essential for all cyclization reactions catalyzed by class I terpene cyclases. For all currently known diphosphate-dependent cyclizations, the reaction requires the Lewis acid (Mg^{2+})-catalyzed ionization of the diphosphate ester of the respective precursor. After isomerization and the subsequent elimination of the diphosphate group, a primary carbocation is generated, which subsequently attacks a nearby double bond, leading to the formation of the more stable carbenium ion (R_3C^+). Finally, the reaction is terminated by deprotonation, leading to a vast variety of cyclic terpenoids (1, 5, 6).

Cyclic triterpenes, such as hopanoids in bacteria or sterols in eukaryotes, arise either from the nonactivated linear precursor squalene (C_{30}) or from 2,3-oxidosqualene. The polycyclization reaction is catalyzed by squalene-hopene cyclases (SHCs), lanosterol synthases, or oxidosqualene cyclases, respectively (21, 25). The cyclization reaction of squalene involves the concerted formation of (i) five ring structures, (ii) 13 carbon-carbon bonds, and (iii) nine stereocenters, rendering it one of the most complex single-step reactions known in biochemistry. Remarkably, triterpene cyclases do not require a phosphorylated or otherwise activated precursor molecule. Instead, SHCs employ a Brønsted acid-driven catalysis mechanism for the polycyclization cascade (10). An aspartate of the DxDD motif that is strictly conserved in SHCs represents the protonating Brønsted acid (for a recent review of SHCs, see reference 28). SHCs are specific for triterpenes, and to

our knowledge, no SHC or any other biocatalyst that is able to catalyze a cyclization reaction of a nonactivated monoterpene has been described so far. There is, however, ample precedence for the chemocatalytic cyclization of nonactivated monoterpenoids (19, 35).

The selective synthesis of cyclic monoterpenoids such as isopulegol (2-isopropenyl-5-methyl-cyclohexanol) as an intermediate for menthol [5-methyl-2-(propan-2-yl)-cyclohexan-1-ol] synthesis from cheap and readily available precursors such as citral (3,7-dimethyl-octa-2,6-dienal) is of considerable biotechnological interest. Key steps of this route are the enantioselective reduction of citral to (*R*)-citronellal (3,7-dimethyl-oct-6-enal) and the subsequent cyclization to isopulegol. With enzymes like the so-called old yellow enzymes (e.g., OYE 2 [16]), the conversion of citral to citronellal has already been achieved, and ee values of >99% and product concentrations of 50 g/liter have been demonstrated (9, 15). Unfortunately, the strict dependence of class I terpene synthases on phosphorylated precursors for the subsequent cyclization reaction limits the technical applicability of this group of enzymes. We are therefore interested in the identification of biocatalysts with the ability to cyclize linear nonactivated monoterpenes. The cyclization of citronellal to isopulegol is thermodynamically possible, and chemical procedures for citronellal cyclization with inorganic catalysts have been described in the literature (3, 19, 35). Thus, it is reasonable to assume that activation-independent monoterpene cyclases exist or can be evolved from existing enzymes. Consequently, we decided to investigate the diversity of currently known SHCs for their potential to catalyze the cyclization of monoterpenoids, and we started with the SHC from *Alicyclobacillus acidocaldarius* (formerly *Bacillus acidocaldarius*) (Shc_{Aac}), which is the best-studied triterpene cyclase so far: its structure and details of its reaction mechanism have

Received 29 September 2011 Accepted 23 November 2011

Published ahead of print 9 December 2011

Address correspondence to Dieter Jendrossek, dieter.jendrossek@imb.uni-stuttgart.de.

Copyright © 2012, American Society for Microbiology. All Rights Reserved.

doi:10.1128/AEM.07059-11

been determined in the last decades (11, 14, 31, 32). Interestingly, Shc_{Aac} is rather unspecific with respect to carbon atom backbone lengths of accepted substrates and can cyclize a variety of alternative substrates, such as, e.g., homofarnesol to ambroxane, an important aroma chemical (11, 18; for a summary, see Table 2 in reference 28).

MATERIALS AND METHODS

Bacterial strains and biocatalysts. Experiments were performed with resting *Escherichia coli* cells or with purified ZMO1548-Shc protein. SHCs were overexpressed in *E. coli* BL21(DE3)(pLys) or *E. coli* Rosetta(pLys-RAR62) harboring pET16b constructs with the respective *shc* gene by growth in Luria-Bertani medium supplemented with ampicillin (100 µg/ml), chloramphenicol (34 µg/ml), and 0.5 mM isopropyl thio-β-D-galactoside at an optical density at 600 nm (OD₆₀₀) of 0.4 and additional growth for 4 h at 30°C. Cells were harvested by centrifugation and stored at −20°C until use. Purified ZMO1548-Shc protein was obtained from 240 g of cell biomass suspended in 1 liter of sodium phosphate buffer (50 mM NaP_i, 10 mM MgCl₂, 0.5% [vol/vol] Triton X-100 [pH 6.5]). Cells were disrupted by homogenization (APV-Gaulin Micron-Lab-40 high-pressure homogenizer, 140 MPa). Cellular debris were removed by centrifugation (1 h at 4°C at 9,150 × g). Crude extract (37-mg/ml protein content) was purified on a Q-Sepharose column (dimensions, 23 by 5 cm; flow rate, 10 ml/min; loading rate, 20 mg protein/ml in stationary phase) (GE Healthcare). Sepharose was equilibrated with a solution containing 50 mM KH₂PO₄ and 0.05% (vol/vol) Triton X-100 (pH 7.5); proteins were eluted by a linear gradient (0 to 100%) of 0.99 liters of equilibration buffer containing 1 M NaCl, followed by 0.72 liters of this buffer. Active fractions were pooled, combined with (NH₄)₂SO₄ (20%, wt/vol), and loaded onto a phenyl-Sepharose column (dimensions, 28 by 2.6 cm; flow rate, 10 ml/min) (GE Healthcare). This column was equilibrated with a solution containing 50 mM KH₂PO₄, 0.05% (vol/vol) Triton X-100, and 20% (wt/vol) (NH₄)₂SO₄ (pH 7.0); protein was eluted by a linear gradient (0 to 100%) of 0.5 liters of a solution containing 50 mM KH₂PO₄ and 0.05% (vol/vol) Triton X-100 (pH 7.0), followed by 0.9 liters of this buffer and 0.7 liters of a solution containing 50 mM KH₂PO₄, 0.05% (vol/vol) Triton X-100, and 20% (vol/vol) 2-propanol (pH 7.0). Finally, the active fractions of the phenyl-Sepharose chromatography column were pooled and loaded onto a Mono-Q HR 16/10 column (dimensions, 10 by 1.6 cm; flow rate, 4 ml/min) (GE Healthcare) equilibrated with a solution containing 50 mM KH₂PO₄ and 5% (vol/vol) Tween 20 (pH 7.5); protein was eluted with 40 ml equilibration buffer, followed by a linear gradient (0 to 100%) of 0.5 liters of a solution containing 50 mM KH₂PO₄, 5% (vol/vol) Tween 20, 0.5 M NaCl (pH 7.5), and 80 ml of the latter buffer. The purified protein was ≥90% pure and was stored in aliquots at −20°C.

Vector construction. The two *Zymomonas mobilis* squalene-hopene cyclase genes ZMO1548 (GenBank accession number YP_163283.1; region, 1578816...1580993) and ZMO0872 (accession number AAF203881_2; region, 1068...33044) and ZMO1548-Shc mutant variants were PCR cloned into pET16b by standard techniques. The constructs were confirmed by DNA sequencing and transformed into *E. coli* cells. Only clones with the correct sequence were used. The *shc* gene from *A. acidocaldarius* was cloned by an analogue procedure.

Cyclization assay. Recombinant *E. coli* cells were suspended in 20 mM Tris-HCl (pH 8.0) (3 ml per g wet cells). The cyclization mixture contained 250 µl of the cell suspension or 250 µl purified ZMO1548-Shc protein (0 to 2.6 mg/ml), 50 µl of 1 M citrate buffer (pH 4.5), 20 mM (final concentration) substrate, and water to 500 µl. In case squalene was cyclized, 1% (vol/vol) Triton X-100 was added. The reaction was terminated after 16 h at 30°C by extraction with heptane. The organic phase was analyzed by gas chromatography (GC). Controls were performed with *E. coli* cells harboring the empty vector and with heat-inactivated SHC-expressing cells. A slow chemical formation of isopulegol from citronellal at pH 4.5 in the absence of enzyme or cells was observed within a reaction

time of 16 h. This nonenzymatic activity was subtracted from all data obtained with biocatalysts.

Gas chromatography. Terpenoids were assayed qualitatively and quantitatively by gas chromatography using an Agilent 7890A GC apparatus equipped with a DB-5 column (20 m by 0.1 mm [internal diameter]; film thickness, 0.1 µm) and a flame ionization detector. Alternatively, a Shimadzu GC-MS QP 2010 system with an FS Supreme 5 column (30 m by 0.25 mm by 0.25 µm) was used for coupled GC and mass spectrometry (MS) analysis. Three microliters of the solvent extract was applied onto the column (split ratio, 1:5; helium flow rate, 0.25 or 0.5 ml/min; injector temperature, 250°C). For the separation of linear and cyclic monoterpenoids, the initial oven temperature (60°C) was raised to 130°C at a rate of 40°C/min, to 150°C at a rate of 2°C/min, and then to 200°C at a rate of 40°C/min. The retention times of terpenoids (Fig. 1A) were as follows: 7.55 min for (R,S)-citronellal, 7.70 min for isopulegol, 7.90 min for *neo*-isopulegol, 8.10 min for *iso*-isopulegol, 8.25 min for *neois*-isopulegol, and 9.91 min for 1-decanol (internal standard). The sum of all isopulegol isomers was calculated. For the detection of triterpenes, the injector temperature used was 300°C. The oven temperature was initially 60°C, raised to 220°C at a rate of 40°C/min, further increased to 310°C at a rate of 6°C/min, and kept constant for 10 min. Squalene and hopene eluted after 19.2 min and 26.9 min, respectively. Alternatively, GC-MS was used for coupled GC and MS analyses (split ratio, 1:20) (3 min at a temperature of 120°C with an increase to 135°C at a rate of 2°C/min and a further increase to 365°C at a rate of 10°C/min, followed by cooling to 300°C at a rate of 70°C/min).

GC-MS data were analyzed by using LabSolutions GCsolutions Postrun software. The substrate citronellal always contained minor amounts of isopulegol as impurities. The GC area values of the respective terpenoids were taken as 100%. The area values of isopulegol in the product were corrected by the amount of isopulegol already present in the substrate and the control, respectively.

Homology models and docking. The A chain of the *A. acidocaldarius* SHC structure (Shc_{Aac}) (Protein Data Bank [PDB] accession number 1UMP) was used as a template to build homology models for wild-type ZMO1548-Shc and the F486A and F486Y variants by using Modeler 9v6 (26). The target-template alignment was extracted from a multisequence alignment of the homologous family SHC_1 from the Triterpene Cyclase Engineering Database (<http://www.ttcad.uni-stuttgart.de/>), which contains Shc_{Aac} as well as ZMO1548-Shc. Additional secondary-structure restraints for the modeling process were derived from PSIPRED v3.0 (2), and the cocrystallized inhibitor 2-azasqualene was included in the modeling process. The very slow refinement method for the loops was selected, and three models were constructed for wild-type ZMO1548-Shc and each variant. The models were compared by using ProSA-web (34), and those with the lowest Z scores were selected for the docking experiments. The N-terminal region of ZMO1548-Shc and the loop region of G670 to R696 could not be aligned to a part of the template sequence, nor could any secondary-structure elements be predicted for the N-terminal region. The respective parts of the model therefore have no distinct structure and are energetically less favorable (the Z scores for the template, wild-type ZMO1548-Shc F486, F486A variant, and F486Y variant models were −13.43, −10.97, −10.69, and −10.78, respectively). However, neither of the two regions contributes directly to the substrate pocket, and the lower quality of the model in those regions does not impair the docking analysis, which focuses exclusively on the substrate pocket. Molecular docking of the various substrates and intermediates into the three ZMO1548-Shc models was performed by using FlexX (22). A pH of 4.5, which was also used for the conversion experiments, was assumed, and residues were protonated according to their pK_a values computed with MCCE (29). Squalene, its two-ring cyclization intermediate, citronellal, and isopulegol were docked by automatically selecting and placing the base fragment and default parameters. Residues of amino acids with charged side chains were protonated according to their pK_a values, computed with MCCE. Aspartate residues 352, 385, 401, 413, 464, 487, 539, 681, and 722; glutamate

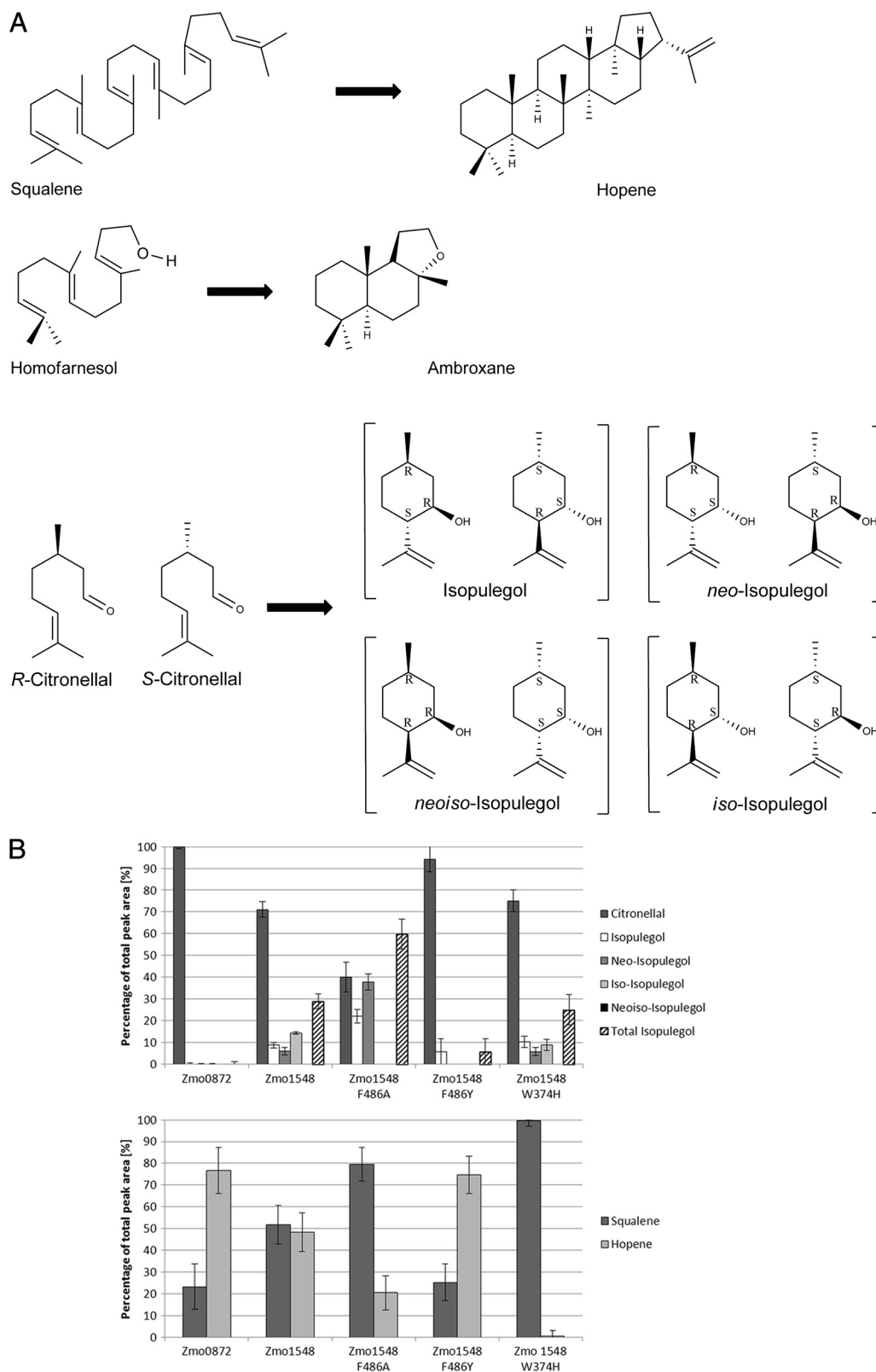


FIG 1 (A) Structures of linear substrates of *Z. mobilis* ZMO1548-Shc and their cyclization products. All isomers except for *neoiso*-isopulegol were detected as cyclization products from (*R,S*)-citronellal with ZMO1548-Shc as a catalyst. *neoiso*-Isopulegol is given for the completeness of isomer/enantiomer structures. (B) Cyclization of citronellal to isopulegol (top) and squalene to hopene (bottom) by *Z. mobilis* SHCs. Extracted starting material and products were analyzed by GC and/or by GC-MS. Error bars indicate standard deviations. Experiments were performed with at least three independent batches of *E. coli* cells (3 biological replicates and 12 technical replicates). The 100% value corresponds to 20 mM starting material.

residues 89, 95, 108, 122, 192, 250, 254, 262, 389, 395, 455, 463, 544, 546, 548, 575, 595, 636, 639, 660, 660, and 662; and K643 were all without a charge. Histidine residues 110, 284, 303, 311, 341, 358, 517, and 592 and H646 all had a +1 charge. Structure files were visualized by using programs such as PYMOL or CHIMERA. Structures of Shc models and docked substrates are available from the corresponding author upon request.

RESULTS AND DISCUSSION

The so far best studied SHC of *A. acidocaldarius* (Shc_{Aac}) was expressed in recombinant *E. coli* cells and analyzed for its potential to cyclize terpenoids other than squalene. We were able to confirm hopene-forming SHC activity from squalene by Shc_{Aac} and also detected ambroxane as a cyclization product from homofarnesol. The level of homofarnesol cyclase activity was extremely low in comparison to its activity on squalene. However, attempts to demonstrate substantial cyclase activity with citronellal or related monoterpenoids failed (data not shown). Apparently, wild-type Shc_{Aac} has no detectable monoterpene cyclase activity.

Data mining of published genome sequences revealed that hundreds of potential *shc* genes are annotated in databases. An exciting example comes from *Zymomonas mobilis*. Interestingly, and in contrast to most genome-sequenced bacteria, *Z. mobilis* has two *shc* genes, namely, ZMO0872 (GenBank accession number [AAF203881_2](#)) and ZMO1548 (accession number YP_163283.1). The hopene-forming activity of ZMO0872-Shc was proven previously (24), but the function of ZMO1548-Shc has never been investigated. Furthermore, an analysis of the cyclic triterpene content of *Z. mobilis* revealed that this organism has unusual hopanoids in its hydrocarbon fraction in addition to hopene (7). These observations suggest unusual properties of the involved SHCs and prompted us to investigate the activities of both *Z. mobilis* SHCs in the cyclization of short linear terpenoids as alternative substrates. For this, both *Z. mobilis shc* genes were cloned and heterologously expressed in *E. coli*, and terpene cyclase activities with nonactivated terpenoids were determined. A substantial amount of the applied squalene was converted to hopene by recombinant *E. coli* cells expressing ZMO1548-Shc or ZMO0872-Shc, which confirmed the functional expression of both genes (Fig. 1B). Remarkably, the formation of ambroxane from homofarnesol and isopulegol from citronellal was catalyzed by ZMO1548-Shc-expressing resting *E. coli* cells but not by *E. coli* cells harboring ZMO0872-Shc. Our experiments demonstrated the formation of 5 mM isopulegol (25%) from a starting concentration of 20 mM citronellal in the assay within an incubation period of 16 h. No significant isopulegol-forming activity was detected in cells expressing ZMO0872-Shc or in any of the corresponding controls (heat-inactivated *E. coli* cells expressing ZMO1548-Shc, *E. coli* cells harboring the empty vector, or *E. coli* cells expressing ZMO1548-Shc in the absence of a substrate). The identity of isopulegol was confirmed by GC-MS analysis and chromatographic comparisons with authentic material (data not shown). Therefore, ZMO1548 is the first biocatalyst showing monoterpene cyclase activity with a not-activated substrate.

The cyclization reactions described above were successfully repeated with isolated ZMO1548-Shc protein that was obtained from cell extracts by classic chromatographic methods (Fig. 2). The cyclization reaction did not require any supplements except enzyme, water, substrate, and appropriate physical conditions (pH 4.5 to 6.0 and a temperature of 20°C to 40°C). However, the

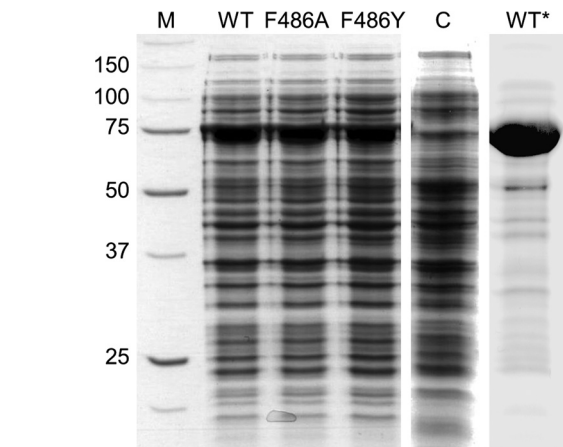


FIG 2 Purity of biocatalysts used in this study. Reducing standard SDS-PAGE was performed. Lanes from left to right are molecular mass marker proteins, with the indicated molecular masses in kilodaltons (lane M); *E. coli* expressing wild-type ZMO1548-Shc (lane WT); *E. coli* expressing ZMO1548-Shc F486A (lane F486A); *E. coli* expressing ZMO1548-Shc F486Y (lane F486Y); *E. coli* with an empty pET16b vector (lane C); and 26 μ g of purified ZMO1548-Shc (lane WT*). For whole-cell analysis, 17 μ g of cell biomass in denaturation solution was heated to 95°C for 5 min and applied to the gel after the removal of solids by centrifugation.

activity obtained with the purified enzyme was greatly reduced in comparison to the activities obtained with resting cells. About 1 mg purified ZMO1548-Shc protein was needed to convert 1 mM citronellal to isopulegol within 16 h. This equates to a more-than-10-fold reduction in activity compared to results obtained by the use of resting cells. The isopulegol-forming activity of purified ZMO1548-Shc was dependent on the reaction time, the amount of the Shc catalyst, and the citronellal concentration (reaction time of 0 to 24 h, 0 to 2.6 mg ZMO1548-Shc per ml assay mixture, and 0 to 40 mM citronellal tested). The strong difference in activity between the isolated enzyme and resting cells can most likely be explained by the finding that *in vivo*, SHCs, as far as investigated, are attached to the cytoplasmic membrane, which is also true for ZMO1548-Shc-expressing *E. coli* cells, as the majority of the enzyme was found in the membrane fraction (data not shown). It is known that Shc_{Aac} is a monotopic enzyme and that the catalytic cleft of the enzyme is facing toward the cytoplasmic membrane (20, 31, 33). From these data, it can be concluded that the hydrophobic substrate is most likely delivered to the enzyme via the cytoplasmic membrane. Therefore, the absence of the native environment of the purified SHC protein apparently has a negative influence on the overall activity of the enzyme. Because of the higher level of activity of ZMO1548-Shc in the whole-cell system, all subsequent experiments were performed with resting cells. SDS-PAGE was used to determine the amount of expressed SHC protein in the different batches of the recombinant strains to ensure that comparable amounts of SHC protein were used in all assays.

Mutagenesis of the active site of ZMO1548-Shc. Shc_{Aac} is the only SHC for which a structure of a complex with the inhibitory substrate analogue 2-azasqualene has been determined (23, 32). The orientation of the substrate in the active site ensures close proximity to the molecule's reactive residues and determines the stereochemical course of the cyclization. All attempts to obtain refracting crystals from purified ZMO1548-Shc have failed so far.

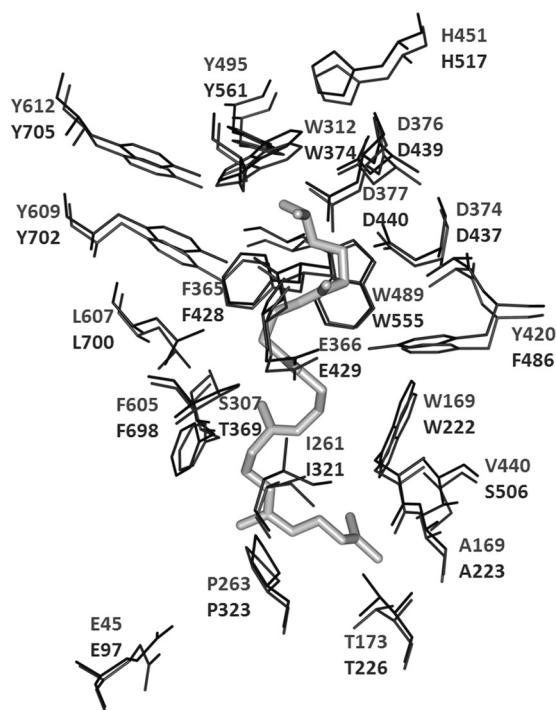


FIG 3 Comparison of the active sites of *A. acidocaldarius* SHC and *Z. mobilis* SHC1548. Shc_{Aac} and ZMO1548-Shc amino acids are given (Shc_{Aac} at the top and ZMO1548-Shc at the bottom). The squalene analogue 2-azasqualene is shown in gray. Note the almost identical positions of residues in the two SHCs.

ZMO1548-Shc is related to Shc_{Aac} (44% and 54% overall amino acid identity and similarity, respectively). The amino acids of the active site and substrate binding pocket are almost identical in both SHCs (Fig. 3). Alignments of 325 SHC amino acid sequences from the Triterpene Cyclase Engineering Database (<http://www.ttcdd.uni-stuttgart.de/>) revealed that a phenylalanine in the substrate binding site (F486 in ZMO1548-Shc) is conserved in 312 (96%) SHC sequences, highlighting the importance of this residue. Remarkably, Shc_{Aac} has a tyrosine at the equivalent position (Y420), and this is the only difference between Shc_{Aac} and ZMO1548-Shc in the inner part of the binding pocket. This part of the binding pocket is likely to interact with the first isoprene units of squalene. Mutagenesis of this residue (Y420A) led to the formation of bicyclic and tricyclic products from squalene with Shc_{Aac} (8). We concluded that this position is not essential for the initiation of the cyclization reaction but apparently is important for the continuation of the polycyclization reaction. Mutagenesis of this residue therefore should have only little effect on the cyclization of monoterpenoids. To test this hypothesis, we constructed F486Y and F486A mutants, functionally expressed wild-type ZMO1548-Shc and both ZMO1548-Shc variants in *E. coli* cells, determined the cyclase activities with squalene and citronellal, and compared the activities of the variants with those of the wild-type enzyme. Surprisingly, the activity with citronellal as a cyclization substrate was almost completely abolished in cells expressing the F486Y variant of the enzyme, but the squalene cyclase activity was increased by about 50% (Fig. 1B). Moreover, the citronellal cyclase activity of cells expressing the F486A mutant was increased nearly 3-fold (12 mM isopulegol formed from 20 mM substrate) compared to cells expressing the wild-type enzyme,

while the cyclase activity with squalene was strongly reduced. These results confirmed the importance of position 486 for squalene cyclase activity and demonstrated that a replacement of the large aromatic ring by a smaller residue increases activity with citronellal.

Modeling of squalene and citronellal binding to the active site of ZMO1548-Shc. All three SHC variants of ZMO1548-Shc F486 (F486 [wild type], F486Y, and F486A) were more or less effectively able to catalyze the cyclization of squalene but strongly differed in their citronellal cyclase activities. To investigate the molecular basis of the experimentally determined differences in substrate specificity, homology models were generated for all these SHC variants using Shc_{Aac} as a template. The structures obtained were then used to individually model the interactions of the three ZMO1548-Shc variants with squalene and with citronellal by the docking of the substrates in their ground state and as two-ring reaction intermediates of squalene. Binding of squalene and bending of the first two isoprene units into the hopene ring structure were found for all three mutants (Fig. 4a), which was similar to the conformation previously described for the Shc_{Aac} complex with 2-azasqualene (31, 32). No spatial hindrance by any residues of the active site was detectable, and the squalene molecule was stabilized by multiple hydrophobic interactions between the methyl groups of squalene and the aromatic residues of the substrate binding site: W222, W374, F428, and W555 in the inner part of the binding pocket near to the protonating DxDD motif and F182, F503, F668, and F698 in the outer part of the binding pocket (Fig. 4a). The docking score, which estimates the free energy of binding, was lower for the more active F486Y variant (−5.1 kJ/mol) than for the wild type (−3.5 kJ/mol) and was highest for the F486A mutant (−2.9 kJ/mol) (Table 1). This finding is in good agreement with the observed differences in the conversion activities toward squalene (F486Y > F486WT > F486A) (Fig. 1B). It was proposed previously that the formation of the ring structure of hopene proceeds stepwise, because the Y420A variant of SHC_{Aac} resulted in a bicyclic product, which was considered an intermediate of the cyclization reaction (17). Upon the modeling of the bicyclic intermediate of squalene in complex with the SHC variants (Fig. 4a), all docking poses found for the wild-type enzyme and for the F486Y variant were in identical positions between residues 486 and Y702. Significantly more interactions were found between the bicyclic intermediate and the F486Y variant. This is in agreement with the increased SHC activity of the F486Y mutant. In contrast, upon docking to the F486A variant, the bicyclic intermediate resulted in different poses, and no interaction with alanine at position 486 was found. A productive orientation of the bicyclic intermediate is expected to be essential for the formation of hopene, and a weakening of the interaction with the intermediate would therefore result in a reduced formation of hopene by the F486A variant. These results are in good agreement with the experimentally determined order of SHC activity, F486Y > F486WT > F486A, with squalene as a substrate.

The strong increase in the squalene cyclase activity of the F486Y variant can be explained by the facilitated binding and stabilization of the bicyclic intermediate (Fig. 4a). Interestingly, the enzyme-bound bicyclic intermediate of squalene is 6 Å away from the protonation site of the DxDD motif. This is in contrast to the experimentally determined structure of the substrate analogue 2-azasqualene or the modeled structure of squalene in its ground state, where the distance was 3 Å. This finding suggests the follow-

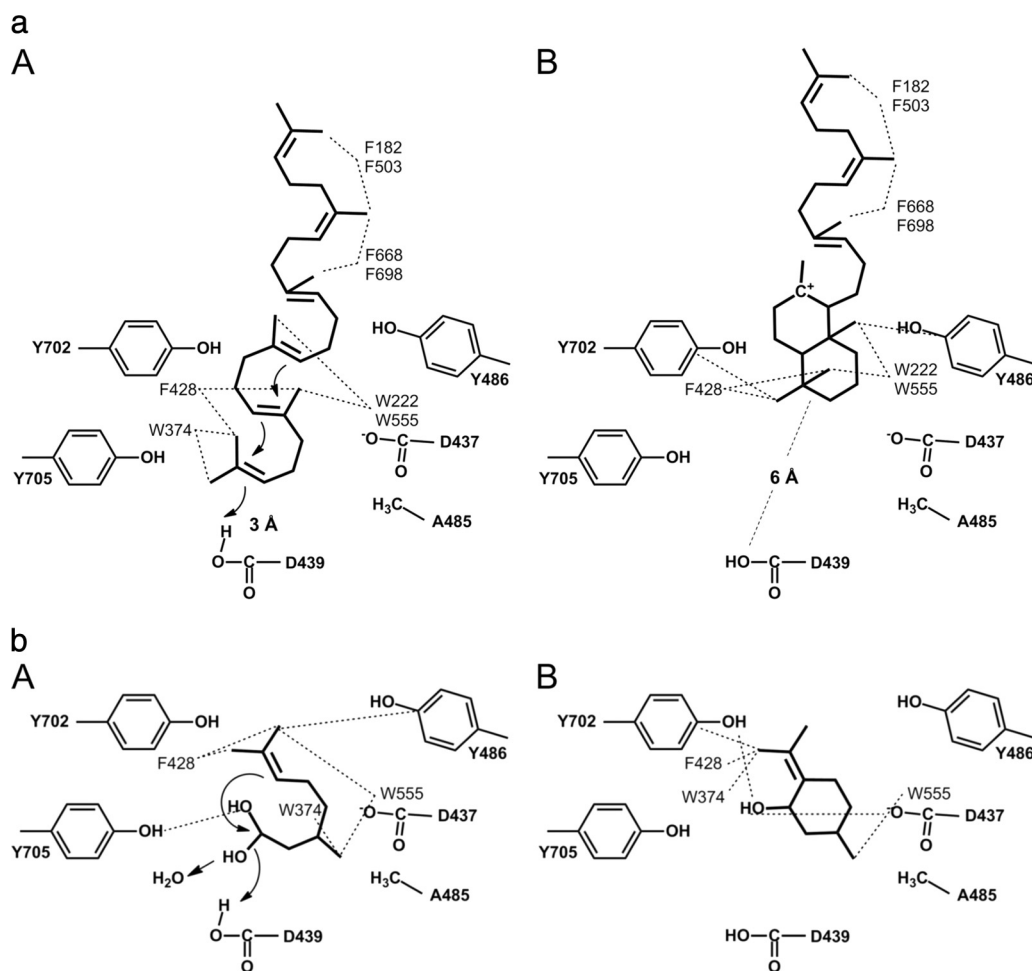


FIG 4 (a) Interactions of squalene and its two-ring intermediate in the ZMO1548-Shc F486Y substrate pocket. (A) A terminal double bond of squalene is located close to the catalytic aspartate (D439) of the DxDD motif. The first two isoprene units meander in the inner part of the substrate pocket and enable the formation of two six-member rings in a chair configuration by small molecular changes (↷). Squalene is stabilized in this position by hydrophobic interactions (····) with several aromatic residues that are located above and below the plane of the scheme: W222, W374, F428, and W555 in the inner part of the pocket and F182, F503, F668, and F698 in the shallow part of the pocket. (B) The localization and stabilization of the isoprene units of the two-ring intermediate are similar to those of squalene in the shallow part of the substrate pocket. The two formed six-member rings in the inner part of the pocket are located further away (6 Å) from the catalytic aspartate than squalene (3 Å). The two rings are located between the two residues 702 and 486 that form a neck in the substrate pocket and are stabilized by hydrophobic interactions with W222, F428, W555, and residues 486 and 702. For distance values not indicated in the figure, see the text. (b) Interactions of hydrated citronellal and isopulegol in the ZMO1548-Shc F486Y substrate pocket. (A) An OH group of the hydrated citronellal is located close to the catalytic aspartate (D439), while the whole molecule preforms a ring conformation that enables the formation of isopulegol in a chair configuration by small molecular changes (↷). Citronellal is stabilized in this position by hydrophobic interactions with several aromatic residues, most of them located above or below the plane of the scheme (W374, F428, F/Y486, and W555), and a single hydrogen bond (····) between Y705 and an OH group of the hydrated citronellal. (B) Isopulegol is stabilized in the inner part of the substrate pocket by hydrophobic interactions with W374, F428, W555, and Y702 and by a single hydrogen bond between the isopulegol OH group and Y702 or D437.

ing reaction mechanism for the initiation of squalene cyclization. Squalene binds in a preformed bicyclic conformation close (3 Å) to the catalytic aspartate, where it is activated by protonation from the DxDD motif. Once the terminal double bond of squalene is protonated, the first two hopene rings are formed. The bicyclic intermediate then moves away from the active site (6 Å) and is stabilized by the binding of the methyl groups to the aromatic rings of Y486 and Y702 (calculated distances of 4.0 to 4.5 Å).

Few differences were found between the ZMO1548-Shc F486 variants with respect to citronellal binding to the substrate pocket. Similarly, the product isopulegol was docked into the three variants with few differences in the resulting poses. The molecule was stabilized by hydrophobic interactions with W374, F428, W555,

and Y702 and a hydrogen bond with Y702 or D437 in the inner part of the substrate binding site (Fig. 4b). The docking score for isopulegol slightly favored the F486A variant over the wild type and the F486Y variant, which would be in agreement with the observed activities (Table 1 and Fig. 1). In summary, the docking score upon the binding of the substrate or of the product to the active site only partially explains the determined considerable differences in the cyclase activities of the three ZMO1548-Shc variants toward citronellal. However, citronellal poses were found in more areas of the binding pocket leading to the initiation site for the more active F486A variant than for the wild type and the F486Y variant. In the cases of the less active F486 wild-type and F486Y variant, most poses were found in the inner part of the

TABLE 1 Docking scores of citronellal, isopulegol, squalene, and the two-ring intermediate of squalene docked into wild-type ZMO1548-Shc and the F486Y and F486A variants^a

Compound	Docking score (kJ/mol)		
	F486 wild type	F486A variant	F486Y variant
(<i>R</i>)-Citronellal	−5.7	−5.7	−5.7
Isopulegol	−8.3	−9.1	−7.8
Squalene	−3.5	−2.9	−5.1
Two-ring intermediate	−3.7	−4.6 ^b	−5.0

^a The docking scores for (*R*)-citronellal and squalene are given for the first pose that is productively oriented toward the DxDD motif in the inner part of the binding pocket. The docking scores for isopulegol and the two-ring intermediate are given for the first-ranked pose.

^b The docking score for the two-ring intermediate docked into the F486A variant is not directly comparable to the other docking scores for the two-ring intermediate, as the pose is in a different orientation than that of the other two and is less likely to allow a continuation of the cyclization of squalene.

binding pocket, with only low docking scores, and not in the parts of the binding pocket leading to the initiation site. Furthermore, the OH group of the F486Y variant was stabilizing citronellal in the inner part of the pocket in a nonproductive orientation, which is in agreement with the observed low level of activity of the F486Y variant with citronellal. These findings suggest that there are at least two factors that determine the citronellal cyclization activity of wild-type and mutant ZMO1548-Shc: the binding of the substrate in a productive ring conformation and substrate stabilization during diffusion through the substrate channel to the active site.

Protonation of citronellal differs from that of squalene. The most remarkable result of the docking of the monoterpene substrate citronellal into the substrate binding site of the three SHC variants was the finding that the aldehyde group of citronellal was preferably placed near the catalytic aspartate of the DxDD motif, resulting in an reverse orientation of citronellal relative to that of squalene (Fig. 4b). Consequently, the carbonyl group is expected to represent the target of the DxDD protonation site. The citronellal molecule was stabilized in the inner part of the binding pocket by hydrophobic interactions with residues W374, F428, F486, and W555 as well as by a single hydrogen bond with Y705, and the ring structure of the product isopulegol was preformed. In contrast, attempts to model citronellal in the same orientation as that of squalene, i.e., with the terminal C=C double bond oriented toward the active-site aspartate, did not result in poses with

a preformed chair conformation of the first isoprene unit (data not shown). These results were achieved for citronellal as aldehyde as well as for its hydrated hemiacetal form. As a consequence of our findings, the citronellal carbonyl group and not the C=C double bond is orientated toward the protonating DxDD motif. Thus, the reaction mechanisms of citronellal and squalene cyclizations must be different (Fig. 5). The SHC-catalyzed cyclization of citronellal to isopulegol is similar to the mechanism of chemical citronellal cyclization and to Prins cyclization (19, 30), where a carbonyl group reacts with the C=C double bond. The Lewis acid for citronellal cyclization to isopulegol in chemical synthesis is provided by heterogeneous catalysts like nickel-sulfated zirconia, and menthol is subsequently formed by hydrogenation of isopulegol (4). In the biocatalytic variant, the enzyme's acidic aspartate (Brønsted acid) delivers the proton to the carbonyl oxygen, thus polarizing the adjacent carbon (Fig. 5). Consequently, this adds to the C=C bond forming the cyclohexane ring and a carbocation at the exocyclic tertiary carbon atom. The elimination of a proton from the neighboring methyl group forms isopulegol. To our knowledge, this is the first example that an enzyme catalyzes the formation of a covalent bond in two different substrate molecules by partially different mechanisms.

To find experimental support for our docking-based assumption that the citronellal carbonyl group is the protonation target during citronellal cyclization and not the terminal methyl groups/C=C double bond, we constructed a W374H variant and determined its cyclase activity with squalene and citronellal. Our docking results suggest that W374 stabilizes both terminal methyl groups of enzyme-docked squalene by a hydrophobic interaction (Fig. 4a). Obviously, W374 cannot stabilize the carbonyl group of citronellal if citronellal binds in an orientation opposite that of squalene because of the strong polarity of the carbonyl group. Therefore, a mutation of residue W374 should affect the cyclization of squalene but should have no or only little effect on citronellal cyclization. This was the case exactly when a W374H mutant was constructed and assayed (Fig. 1B): while the squalene-hopene cyclase activity of the W374H mutant was abolished (<1% residual activity), the isopulegol-forming activity from citronellal was only little affected (16% reduced activity compared to that of the wild type). These results are in agreement with analogous findings on the cyclization activity of Shc_{Aac} mutants with squalene and oxidosqualene (27).

Conclusions. In this contribution, we identified the ZMO1548 shc gene product from *Z. mobilis* as a novel type of terpene cyclase

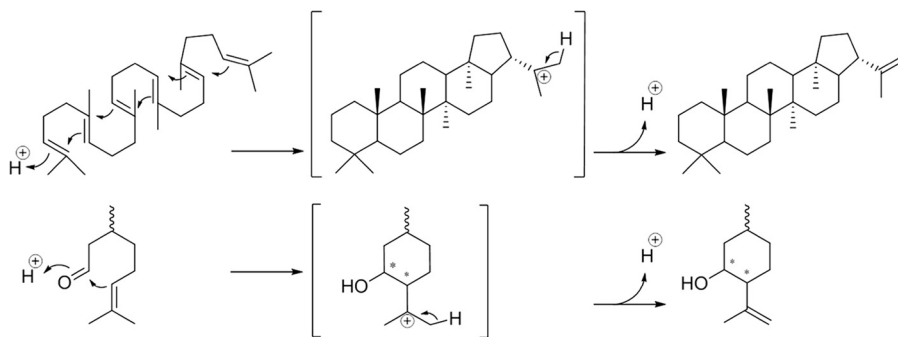


FIG 5 Comparisons of squalene cyclization to hopene cyclization (top) and of citronellal cyclization to isopulegol cyclization (bottom). For details, see the text.

with unexpected substrate promiscuity. ZMO1548-Shc has the unique ability to cyclize substrates of different lengths (monoterpenes to triterpenes) and to accept different molecule classes as substrates, such as hydrocarbons (squalene), alcohols (homofarnesol), and aldehydes (citronellal). Most remarkably, the cyclization of all substrates, including the monoterpene citronellal, was independent of activation by diphosphate, in contrast to conventional monoterpene cyclases. To our knowledge, ZMO1548-Shc is the first biocatalyst with activation-independent monoterpene cyclase activity described so far.

ACKNOWLEDGMENTS

We acknowledge the support by Silvia Racolta for the sequence analysis of the Triterpene Cyclase Engineering Database and G. Sprenger for helpful discussion.

This work was supported by the Bundesministerium für Bildung und Forschung (BMBF) (reference number 315406).

REFERENCES

- Bohlmann J, Meyer-Gauen G, Croteau R. 1998. Plant terpenoid syntheses: molecular biology and phylogenetic analysis. *Proc. Natl. Acad. Sci. U. S. A.* 95:4126–4133.
- Bryson K, et al. 2005. Protein structure prediction servers at University College London. *Nucleic Acids Res.* 33:W36–W38.
- Chuah GK, Liu SH, Jaenicke S, Harrison LJ. 2001. Cyclisation of citronellal to isopulegol catalysed by hydrous zirconia and other solid acids. *J. Catal.* 200:352–359.
- Cortes CB, Galvan VT, Pedro SS, Garcia TV. 2011. One pot synthesis of menthol from ($\pm$)-citronellal on nickel sulfated zirconia catalysts. *Catal. Today* 172:21–26.
- Croteau RB, Davis EM, Ringer KL, Wildung MR. 2005. (–)-Menthol biosynthesis and molecular genetics. *Naturwissenschaften* 92:562–577.
- Davis EM, Croteau R. 2000. Cyclization enzymes in the biosynthesis of monoterpenes, sesquiterpenes, and diterpenes. *Top. Curr. Chem.* 209:53–95.
- Douka E, Koukoku A, Drinas C, Grosdemange-Billiard C, Rohmer M. 2001. Structural diversity of the triterpenic hydrocarbons from the bacterium *Zymomonas mobilis*: the signature of defective squalene cyclization by the squalene/hopene cyclase. *FEMS Microbiol. Lett.* 199:247–251.
- Füll C, Poralla K. 2000. Conserved tyr residues determine functions of *Alicyclobacillus acidocaldarius* squalene-hopene cyclase. *FEMS Microbiol. Lett.* 183:221–224.
- Hall M, Hauer B, Stuermer R, Kroutil W, Faber K. 2006. Asymmetric whole-cell bioreduction of an α,β -unsaturated aldehyde (citral): competing prim-alcohol dehydrogenase and C-C lyase activities. *Tetrahedron Asymmetry* 17:3058–3062.
- Hoshino T, Kumai Y, Kudo I, Nakano S-I, Ohashi S. 2004. Enzymatic cyclization reactions of geraniol, farnesol and geranylgeraniol, and those of truncated squalene analogs having C20 and C25 by recombinant squalene cyclase. *Org. Biomol. Chem.* 2:2650–2657.
- Hoshino T, Sato T. 2002. Squalene-hopene cyclase: catalytic mechanism and substrate recognition. *Chem. Commun. (Camb.)* 2002:291–301.
- 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. U. S. A.* 97:13172–13177.
- Miziorko HM. 2011. Enzymes of the mevalonate pathway of isoprenoid biosynthesis. *Arch. Biochem. Biophys.* 505:131–143.
- Morikubo N, et al. 2006. Cation- π interaction in the polyolefin cyclization cascade uncovered by incorporating unnatural amino acids into the catalytic sites of squalene cyclase. *J. Am. Chem. Soc.* 128:13184–13194.
- Müller A, Hauer B, Rosche B. 2006. Enzymatic reduction of the α,β -unsaturated carbon bond in citral. *J. Mol. Catal. B Enzym.* 38:126–130.
- Müller A, Hauer B, Rosche B. 2007. Asymmetric alkene reduction by yeast old yellow enzymes and by a novel *Zymomonas mobilis* reductase. *Biotechnol. Bioeng.* 98:22–29.
- Nakano S-I, Ohashi S, Hoshino T. 2004. Squalene-hopene cyclase: insight into the role of the methyl group on the squalene backbone upon the polycyclization cascade. Enzymatic cyclization products of squalene analogs lacking a 26-methyl group and possessing a methyl group at C7 or C11. *Org. Biomol. Chem.* 2:2012–2022.
- Neumann S, Simon H. 1986. Purification, partial characterization and substrate specificity of a squalene cyclase from *Bacillus acidocaldarius*. *Biol. Chem. Hoppe Seyler* 367:723–729.
- Nie Y, Chuah GK, Jaenicke S. 2006. Domino-cyclisation and hydrogenation of citronellal to menthol over bifunctional Ni/Zr-beta and Zr-beta/Ni-MCM-41 catalysts. *Chem. Commun. (Camb.)* 2006:790–792.
- Oliaro-Bosso S, et al. 2005. Access of the substrate to the active site of squalene and oxidosqualene cyclases: comparative inhibition, site-directed mutagenesis and homology-modelling studies. *Biochem. Soc. Trans.* 33:1202–1205.
- Poralla K. 2002. The changing path of hopanoid research: from condensing lipids to new membrane enzymes, p 263–283. *In* Braun V, Götz F (ed), *Microbial fundamentals of biotechnology*. Wiley-VCH, Weinheim, Germany.
- Rarey M, Kramer B, Lengauer T, Klebe G. 1996. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* 261:470–489.
- Reinert DJ, Balliano G, Schulz GE. 2004. Conversion of squalene to the pentacarboxylic hopene. *Chem. Biol.* 11:121–126.
- Reipen IG, Poralla K, Sahm H, Sprenger GA. 1995. *Zymomonas mobilis* squalene-hopene cyclase gene (shc): cloning, DNA sequence analysis, and expression in *Escherichia coli*. *Microbiology* 141:155–161.
- Sahm H, Rohmer M, Bringer-Meyer S, Sprenger GA, Welle R. 1993. Biochemistry and physiology of hopanoids in bacteria. *Adv. Microb. Physiol.* 35:247–273.
- Sali A, Blundell TL. 1993. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234:779–815.
- Sato T, Hoshino T. 1999. Functional analysis of the DXDDTA motif in squalene-hopene cyclase by site-directed mutagenesis experiments: initiation site of the polycyclization reaction and stabilization site of the carbocation intermediate of the initially cyclized A-ring. *Biosci. Biotechnol. Biochem.* 63:2189–2198.
- Siedenburg G, Jendrosseck D. 2011. Squalene-hopene cyclases. *Appl. Environ. Microbiol.* 77:3905–3915.
- Song Y, Mao J, Gunner MR. 2009. MCCE2: improving protein pKa calculations with extensive side chain rotamer sampling. *J. Comput. Chem.* 30:2231–2247.
- Sprenger GA, et al. 1997. Identification of a thiamin-dependent synthase in *Escherichia coli* required for the formation of the 1-deoxy-D-xylulose 5-phosphate precursor to isoprenoids, thiamin, and pyridoxol. *Proc. Natl. Acad. Sci. U. S. A.* 94:12857–12862.
- Wendt KU, Lenhart A, Schulz GE. 1999. The structure of the membrane protein squalene-hopene cyclase at 2.0 Å resolution. *J. Mol. Biol.* 286:175–187.
- Wendt KU, Poralla K, Schulz GE. 1997. Structure and function of a squalene cyclase. *Science* 277:1811–1815.
- Wendt K, Schulz G, Corey E, Liu D. 2000. Enzyme mechanisms for polycyclic triterpene formation. *Angew. Chem. Int. Ed. Engl.* 39:2812–2833.
- Wiederstein M, Sippl MJ. 2007. ProSA-web: interactive Web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* 35:W407–W410.
- Yadav GD, Nair J. 2000. Isomerization of citronellal to isopulegol using electrically engineered sulfated zirconia-carbon molecular sieve composite catalysts, UDCaT-2. *Langmuir* 16:4072–404079.