FEBS Letters

FEBS PRESS science publishing by scientists

X-ray structure of linalool dehydratase/isomerase from *Castellaniella defragrans* reveals enzymatic alkene synthesis

Sina Weidenweber[1,*], Robert Marmulla[2,*], Ulrich Ermler[1] and Jens Harder[2]

1 Department of Molecular Membrane Biology, Max Planck Institute of Biophysics, Frankfurt am Main, Germany
2 Department of Microbiology, Max Planck Institute for Marine Microbiology, Bremen, Germany

Correspondence
U. Ermler, Department of Molecular Membrane Biology, Max Planck-Institute of Biophysics, D-60348 Frankfurt am Main, Germany
Fax: +49 6963031002
Tel: +49 6963031054
E-mail: ulrich.ermler@biophys.mpg.de
and
J. Harder, Department of Microbiology, Max Planck-Institute for Marine Microbiology, Celsiusstrasse 1, D-28359 Bremen, Germany
Fax: +49 421 2028 590
Tel: +49 421 2028 750
E-mail: jharder@mpi-bremen.de

*These authors contributed equally to this work.

(Received 24 February 2016, revised 31 March 2016, accepted 31 March 2016, available online 18 April 2016)

doi:10.1002/1873-3468.12165

Edited by Stuart Ferguson

Linalool dehydratase/isomerase (Ldi), an enzyme of terpene degradation in *Castellaniella defragrans*, isomerizes the primary monoterpene alcohol geraniol into the tertiary alcohol (*S*)-linalool and dehydrates (*S*)-linalool to the alkene β-myrcene. Here we report on the crystal structures of Ldi with and without terpene substrates, revealing a cofactor-free homopentameric enzyme. The substrates were embedded inside a hydrophobic channel between two monomers of the $(\alpha,\alpha)_6$ barrel fold class and flanked by three clusters of polar residues involved in acid-base catalysis. The detailed view into the active site will guide future biotechnological applications of Ldi, in particular, for industrial butadiene and isoprene production from renewable sources.

Keywords: $(\alpha,\alpha)_6$ barrel; acid/base catalysis; isoprenoids; linalool dehydratase/isomerase; monoterpene degradation

Essential oils produced by plants mainly consist of monoterpenes; volatile hydrocarbons with the composition $C_{10}H_{16}$ [1]. More than 127 Tg C·year^{-1} monoterpenes are emitted into the atmosphere and an unknown amount enters soils, lakes and rivers [2]. Up to 1 L monoterpenes·m^{-2} are in forest soils [3] and monoterpenes are among the most frequent indoor volatile hydrocarbons [4]. Microorganisms use this abundant carbon source, but little is known about degradation pathways and the enzymes involved [5]. Linalool dehydratase/isomerase (Ldi), a remarkable enzyme of terpene metabolism, was found in the betaproteobacterium *Castellaniella defragrans* 65Phen that mineralizes monoterpenes coupled to anaerobic denitrification [6]. The periplasmic enzyme reversibly catalyzes the isomerisation from the primary alkenol geraniol into the tertiary alkenol (*S*)-linalool and its dehydration to β-myrcene [6,7] (Fig. 1). Dehydratase and isomerase activities of the bifunctional Ldi were only detectable in the presence of a reductant

Abbreviations
Ldi, linalool dehydratase/isomerase; Lis, linalool isomerase.

(dithiothreitol) and in the absence of O_2 [6,7]. The dehydration reaction of Ldi or genetically engineered variants thereof can be used for the transformation of smaller substrate molecules into butadiene and isoprene according to patents and patent claims [8–13]. The market for these two compounds, precursors for many polymers (e.g. nylon, polyester, polyisoprene), amounts to billions of dollars annually [14].

For a profound understanding of alkene synthesis from an alcohol without adjacent activating groups, e.g. CoA thioesters, and for optimizing Ldi for biotechnological purposes by a more rational approach, detailed structural data of the enzyme are indispensable. In particular, Ldi does not share sequence similarities to any protein of sequence data banks except for a membrane-anchored linalool isomerase (Lis) without dehydratase activity (NCBI: EN087364) very recently discovered by our group [15]. The gene *ldi* codes for a preprotein with an N-terminal signal peptide for transport into the periplasmatic space and the purified enzyme lacked the predicted signal peptide [6]. Applying *ldi* expression in *Escherichia coli* [6], we now purified the Ldi protein with a new procedure and structurally characterized the enzyme by X-ray crystallography without and with substrate molecules bound to the enzyme.

Materials and methods

Cultivation of *E. coli* BL21(DE3) and expression of Ldi

Linalool dehydratase/isomerase was heterologously expressed in *E. coli* BL21(DE3) pET42-Ldi [6] in 500 mL rich medium (yeast extract 12 $g \cdot L^{-1}$, peptone 14 $g \cdot L^{-1}$, NaCl 5 $g \cdot L^{-1}$, $MgSO_4 \times 7$ H_2O 0.25 $g \cdot L^{-1}$, glycerol 2 $g \cdot L^{-1}$, kanamycin 50 $\mu g \cdot mL^{-1}$) at 37 °C in 1 L baffled flasks at 165 r.p.m. Ldi expression was induced by 1 mM Isopropyl-β-D-1-thiogalactopyranoside at an optical density (600 nm) around 1. Cultures were further cultivated for 4–5 h at 32 °C to an OD (600 nm) of approx. 3. Biomass was harvested by centrifugation (14 000 ***g***, 20 min, 4 °C), frozen in liquid nitrogen and stored at −80 °C.

Activity assay

Enzymatic activity was determined by end-point measurement. Subfractions obtained during purification were dialyzed against Tris-Cl buffer (80 mM, pH 8.0). The assay was performed in 4 mL glass vials. About 300–500 µL of sample were treated with 5 mM DTT. Vials were closed with a butyl septum and flushed with N_2 gas for 3 min via needles. After a pre-incubation of 20 min at room temperature, substrate (200 mM (*R,S*)-linalool or geraniol in 200 µL 2,2,4,4,6,8,8-heptamethylnonane) was added via a needle. Incubation took place at 37 °C under mild shaking. Product formation was determined by gas chromatography with flame ionization detection (PerkinElmer Auto System XL, Überlingen, Germany). About 1 µL sample was injected onto an Optima-5 column (30 m × 0.32 mm, 0.25 µm film thickness; Macherey-Nagel, Düren, Germany) with hydrogen as carrier gas and the following temperature program: injection port 250 °C, detection port 350 °C, initial column temperature 40 °C for 2 min, increasing to 100 °C at a rate of 4 $°C \cdot min^{-1}$, keeping 100 °C for 0.1 min, followed by an increase to 320 °C at 45 $°C \cdot min^{-1}$ and hold for 3 min. The split ratio was set to 1 : 9.

Protein purification

E. coli BL21(DE3) pET42-Ldi biomass was suspended in Tris-Cl buffer (80 mM, pH 8) and thawed quickly at room temperature. Cell disintegration was performed by mechanical shearing using a French Press pressure cell at 8.26 MPa two times. The crude extract was clarified by ultracentrifugation (150 000 ***g***, 30 min, 4 °C) and the supernatant (soluble protein extract) was purified on an ÄKTA purifier system (GE Healthcare, Freiburg, Germany) at 6 °C. All buffers were prepared with pH values adjusted at room temperature (22 °C) and filtered (0.2 µm) prior use. Protein was detected by absorption measurement at 280 nm. SDS/PAGE was used to analyze purification samples and protein concentrations were determined according to the method described by Bradford using bovine serum albumin as a calibration standard [16].

Ldi was purified in a three step approach, including hydrophobic interaction, anion exchange and size-exclusion chromatography. Soluble protein extract from *E. coli* BL21 (DE3) pET42-Ldi was prepared in Tris-Cl buffer (80 mM, pH 8). The sample was adjusted to 5% ammoniumsulfate

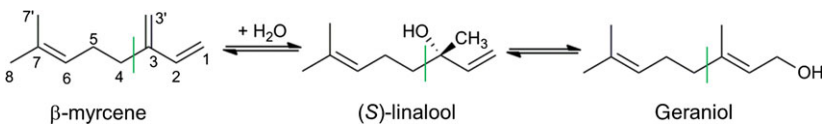

Fig. 1. Reaction of Ldi. Ldi catalyzes the hydration of myrcene to (*S*)-linalool and the hydroxyl transfer from (*S*)-linalool to geraniol. The three terpenes are composed of a methylpentene moiety (left of green line) and buta-1,3-diene (β-myrcene), but-1-en-3-ol (S-linalool) and but-2-en-1-ol (geraniol) moiety. In the free form β-myrcene is thermodynamically most stable, followed by (*S*)-linalool and geraniol [1,28].

(v/v) with a saturated ammonium sulfate solution to aid binding onto a phenyl sepharose column (dimensions: 8 × 100 mm; column volume 5 mL). The column was equilibrated with Tris-Cl buffer (80 mM, pH 8; 5% $(NH_4)_2SO_4$). After extensive washing with binding buffer, a gradual elution of bound protein was performed by applying buffers with decreasing hydrophobicity: Tris-Cl buffer (80 mM, pH 8) – Tris-Cl buffer (20 mM, pH 8, 0.1% Tween20 v/v) – pure water. The change from binding buffer to the first elution buffer was performed in a linear gradient (0–100% over 20 mL), while the other buffer changes were performed in an one-step gradient. Pooled fractions from the peak eluted with 20 mM Tris-Cl (pH 8.0; 0.1% Tween20) were concentrated with an AMICON filter unit (10 kDa cut-off; Merck-Millipore, Darmstadt, Germany) and dialyzed against NaH_2PO_4/Na_2HPO_4 buffer (50 mM, pH 7.0) for at least 12 h. This fraction was loaded onto a ResourceQ anion exchange column (CV 1 mL), equilibrated with NaH_2PO_4/Na_2HPO_4 buffer (50 mM, pH 7.0). After washing off unbound protein with binding buffer, bound protein was eluted with NaH_2PO_4/Na_2HPO_4 buffer (50 mM, pH 7.0) containg 1 M NaCl. The flow through was concentrated with an AMICON filter unit prior the application onto a size-exclusion chromatography column (Superdex 200, dimensions: 16 × 600 mm), equilibrated with Tris-Cl buffer (20 mM, pH 8.0), at a flow rate of 0.6 $mL \cdot min^{-1}$. Active fractions were dialyzed against 10 mM Tris-Cl buffer (pH 8) and subsequently concentrated to 10–15 $mg \cdot mL^{-1}$ protein for further use in crystallography.

Crystallization, X-ray data collection, phase determination and refinement

Ldi was crystallized at 4 °C under oxic as well as anoxic (97% N_2 : 3% H_2 v/v) conditions using the sitting/hanging drop vapor diffusion method. Crystallization conditions are given in Table 1. Diffraction data were collected at the PX II beamline at the Swiss Light Source (SLS, Villingen, Switzerland) and processed with XDS [17]. The multi-wavelength anomalous diffraction data set ldi13 was measured at the Hg (L-III) absorption edge using a Ldi crystal soaked with thimerosal (Table 1). The mercury positions were identified with SHELXC/D [18], the phases were calculated with SHARP [19] and improved by solvent flattening. The quality of the resulting electron density was sufficient for identifying five monomers of the Ldi pentamer but not for chain tracing. Non-crystallographic symmetry operators and the protein envelope were calculated with O [20]. For final structure determination the native data set ldi3 was included and the phases were calculated by the single isomorphous replacement with anomalous scattering method using SHARP. Phases were improved by five-fold averaging with DM [21] of the CCP4 package [22]. Automated model building was performed using BUCCANEER [23]. The model was manually completed and refined using COOT [24], REFMAC5 [25] and Refine in PHENIX [26] including the non-crystallographic symmetry and the translation/liberation/screw motion options. All crystals of Ldi were isomorphous. An elongated molecule present in all five subunits was assigned as partly attached precipitant PEG550 MME. Data collection and refinement statistics are summarized in Table 1. Protein fold comparison and superposition analysis were performed with DALI [27]. All Figures were prepared with PYMOL [28]. The structures without and with substrate were deposited under PDB ID: 5HLR and PDB ID: 5HSS, respectively.

Construction of amino-exchange mutants

Amino acid-exchange was performed by site-directed mutagenesis, using primers with defined mismatches (Table S1), the pET42-Ldi plasmid as a template and the Phusion Polymerase according to the manufacturer manual (Thermo Fisher Scientific, Dreieich, Germany). The complete plasmid was amplified by overlapping primers, ligated and transformed into *E. coli*. Mutants were verified by sequencing of the plasmid with standard T7-primers using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Life Technologies Corporation, Carlsbad, CA, USA) with the following program: 95 °C for 5 min, 99 cycles at 96 °C for 30 s, 55 °C for 20 s and 60 °C for 4 min. Fragments were analyzed on an ABI Prism 3130*xl* Genetic Analyzer (Applied Biosystems, Life Technologies Corporation). Verified Ldi mutants were grown in 100 mL LB medium in 250 mL flasks and protein expression was induced by addition of 1 mM isopropyl β-D-1-thiogalactopyranoside. Soluble protein extract was obtained as described above and used for evaluation of enzyme activity.

Modelling

The cytosolic domain of Lis (amino acid 139–644) was modelled with I-TASSER (Iterative threading assembly refinement, V 4.3) tool, using the structural data of an Ldi monomer as template [29].

Results

Upon crystallization the X-ray structure of Ldi was determined with the single isomorphous replacement anomalous scattering method using thimerosal as heavy metal compound and refined to R/R_{free} values of 19.0/22.5% at 1.9 Å resolution (Table 1).

The crystal structure of Ldi revealed a cyclic homopentameric protein complex (Fig. 2A), in agreement with size-exclusion chromatographic data in solution. The pentamer resembles a planar rosette with a central hole of ca. 30 Å in diameter and is held together by an extended interface between adjacent monomers. Architecturally, each monomer is built up

Table 1. Crystallization, soaking and crystallographic data.

Crystal	ldi3	ldi13			ldi31	ldi42
c_{ldi} (mg·mL^{-1})	13.5	13.5			15	12
Condition	30% (w/v) PEG 550 MME 100 mM bicine pH 9	25% (w/v) PEG 550 MME 100 mM bicine pH 9 100 mM NaCl			22% (w/v) PEG 550 MME 100 mM bicine pH 9	26% (w/v) PEG 550 MME 100 mM bicine pH 9
Co-crystallization	5 mM DTT	5 mM DTT			5 mM (*R,S*)-linalool, 5 mM DTT	4 mM DTT
Soaking	–	5 mM thiomersal, 3 h			5 mM (*R,S*)-linalool, 16 h	5 mM (*R,S*)-linalool, 5 mM DTT, 16 h (anaerobic)
Cryo	–	30 % (w/v) PEG 550 MME, 100 mM bicine pH 9, 100 mM NaCl			35 % (w/v) PEG 550 MME, 100 mM bicine pH 9, 5 mM (*R,S*)-linalool	35 % (w/v) PEG 550 MME, 100 mM bicine pH 9, 5 mM (*R,S*)-linalool, 5 mM DTT
Data collection						
Space group	$P2_12_12_1$	$P2_12_12_1$			$P2_12_12_1$	$P2_12_12_1$
Cell dimensions						
a, *b*, *c* (Å)	90.0, 105.5, 220.3	99.9, 106.1, 220.4			99.9, 106.7, 223.1	99.3, 106.3, 221.3
α, β, γ (°)	90, 90, 90	90, 90, 90			90, 90, 90	90, 90, 90
		Peak	*Inflection*	*Remote*		
Wavelength (Å)	0.99996	1.00724	1.00939	0.99999	1.00002	0.97928
Resolution (Å)	50–2.2 (2.3–2.2)	50–2.4 (2.5–2.4)	50–2.7 (2.8–2.7)	50–3.0 (3.1–3.0)	50–1.9 (2.0–1.9)	50–2.5 (2.6–2.5)
R_{sym}	10.9 (81.3)	13.1 (68.7)	12.3 (54.0)	14.9 (53.9)	9.8 (71.0)	14.0 (131.4)
$I/\sigma I$	10.4 (1.9)	7.9 (2.7)	8.2 (2.5)	7.3 (2.7)	10.5 (2.9)	13.7 (2.4)
Completeness (%)	99.6 (99.6)	99.8 (99.8)	98.7 (98.8)	98.9 (98.5)	99.9 (99.8)	99.8 (99.2)
Redundancy	3.5 (3.4)	4.6 (4.7)	3.3 (3.2)	3.3 (3.4)	5.9 (5.9)	11.7 (11.7)
Refinement						
Resolution (Å)	48.3–2.2 (2.23–2.20)				49.5–1.91 (1.93–1.91)	49.1–2.5 (2.52–2.50)
No. reflections	792246				1074363	957382
R_{work}/R_{free}	17.5 (28.3) 20.9 (32.5)				19.0 (27.5) 22.5 (32.7)	17.9 (37.3) 22.3 (42.7)
No. atoms						
Protein	14590				14602	14602
Ligand/ion	–				–	93
Water	371				580	127
B-factors						
Protein	38.4				35.7	59.3
Ligand/ion	–				–	89.9
Water	36.6				35.0	50.2
R.m.s deviations						
Bond lengths (Å)	0.003				0.011	0.008
Bond angles (°)	0.80				1.17	0.93
Ramachandran favored/ disallowed (%)	97.4/0.1				97.7/0.2	96.9/0.1

of a classical $(\alpha,\alpha)_6$ barrel fold composed of six inner helices (63–80, 124–139, 179–193, 238–251, 294–306 and 341–351) flanked by six lateral helices (25–35, 85–100, 146–163, 198–212, 255–266 and 309–322) oriented in an antiparallel fashion (Fig. 2B). The bottom of the barrel is locked by short loops connecting the inner and lateral helices and by the N (1–25)- and C (352–366)-terminal segments while its entrance is formed by variable segments linking the lateral and the next inner helices (Fig. 2B). A protein data bank research revealed a large number of proteins with an $(\alpha,\alpha)_6$ barrel fold; the highest structural similarities being found between Ldi and rhamnogalacturonyl hydrolase from *Bacillus subtilis* (2d8l) [30] as well as cellobiose-

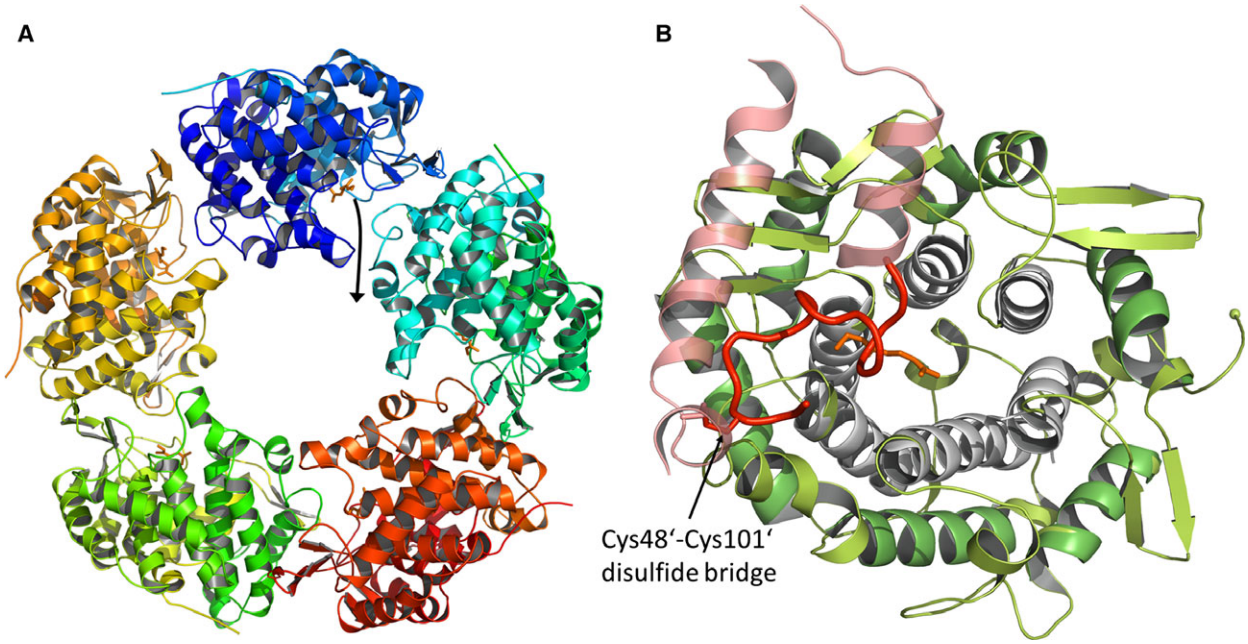

Fig. 2. Structure of Ldi. (A) The homopentamer. The terpene (orange), shown as stick model, is accessible by a channel whose entrance points toward the interior of the pentameric ring (black arrow). (B) The Ldi dimer. The Ldi monomer is built up of an $(\alpha,\alpha)_6$ barrel fold with the inner helices (63–80, 124–139, 179–193, 238–251, 294–306 and 341–351) in gray and the lateral helices (25–35, 85–100, 146–163, 198–212, 255–266 and 309–322) in green. The bottom of the barrel is locked by short loops connecting the inner and lateral helices and by the N (1–25)- and C (352–366)-terminal segments while the entrance of the barrel is formed by variable segments (light green) linking the lateral and the next inner helices. The cavity at the barrel entrance is capped by a loop (red) of the neighboring subunit (salmon). Thus, the dimer represents the functional unit.

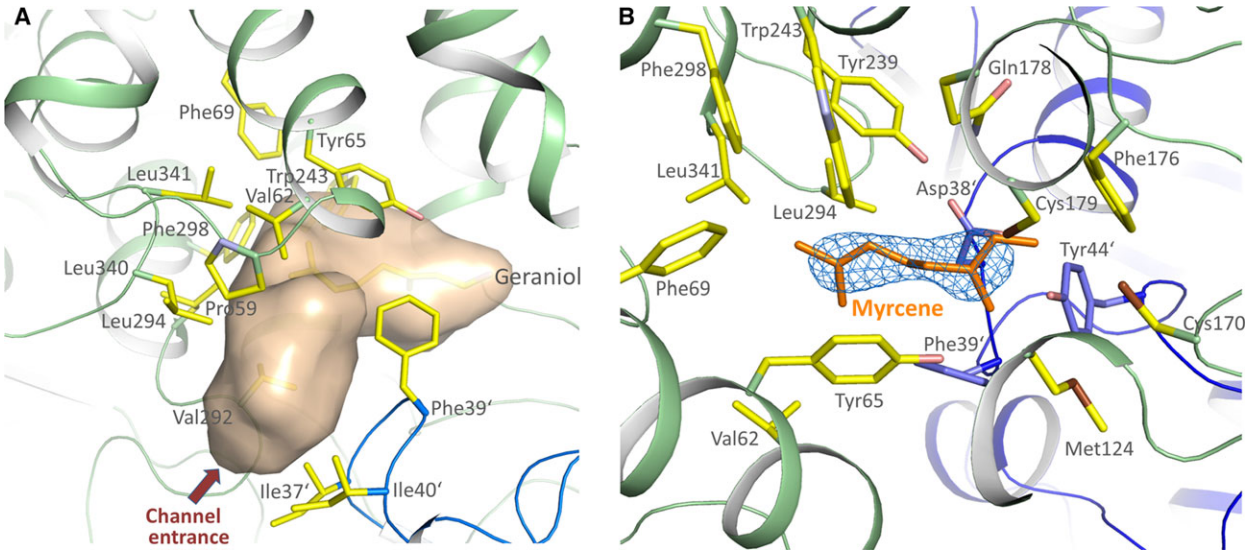

Fig. 3. Substrate binding. (A) Channel architecture. The substrate binding site is embedded inside of 15 Å long and narrow L-shaped channel (shown as salmon surface) accommodated between two subunits (green, blue) of the pentamer. The entrance of the channel is lined up by several nonpolar residues including Ile37′, Ile40′, Pro59, Val62, Val292 and Leu340. (B) The Ldi-myrcene structure. The electron density of myrcene is contoured at 0.8 σ. Its buta-1,3-diene moiety is located at the bottom of the channel flanked by Tyr65, Met124, Cys170, Phe176, Gln178, Cys179, Tyr239, Asp38′, Phe39′ and Tyr 44′ and its methylpentene moiety is surrounded by various hydrophobic residues including Val62, Tyr65, Phe69, Tyr239, Trp243, Leu294, Phe298, and Leu341.

2-epimerase from *Ruminococcus albus* (3vw5) [31], with sequence identities of 16.5% and 15.8%, and with rmsds of 3.3 and 3.4 Å, respectively, using 77% and 75% of the C_α atoms for calculation. Interestingly, $(\alpha,\alpha)_6$ barrel folds are frequently found in isomerases and in enzymes of the terpenoid metabolism, i.e. terpenoid cyclases and protein prenyl transferases [32], but in contrast to Ldi these enzymes operate as a monomer.

The substrate binding site in $(\alpha,\alpha)_6$ barrel enzymes is embedded inside a cavity at the barrel entrance whereas the substrate cavity in Ldi is partly capped by the irregular segment 36′–52′ of the neighbor monomer (marked by an apostrophe) of the homopentamer. Notably, this capping segment is fixed to the $(\alpha,\alpha)_6$ barrel core by an intra-subunit disulfide bond between Cys48′ and Cys101′ (Fig. 2B). Site-specific exchange of one of these cysteines completely inactivated the enzyme (Table S1). Despite the attached neighboring subunit the bottom of the cavity is accessible from bulk solvent via a narrow channel, ca. 15 Å long and 7 Å wide, which is placed at the interface between the two subunits (Fig. 3A). The entrance of the channel is lined up by several nonpolar residues including Ile37′, Ile40′, Pro59, Val62, Val292 and Leu340 which may attract the hydrophobic terpenes. This type of active site architecture is new in $(\alpha,\alpha)_6$ barrel proteins and therefore unpredictable without structural data. Rhamnogalacturonyl hydrolase [30], structurally most related to Ldi, and various other $(\alpha,\alpha)_6$ barrel proteins are endowed with a more polar and exposed substrate binding site only partly shielded. An alternative strategy of active site encapsulation has been reported for monomeric squalene cyclases [33] that use an additional domain for this purpose.

To study the substrate binding and the catalytic mechanism, Ldi crystals were soaked with (*R*,*S*)-linalool and dithiothreitol under enzyme activity

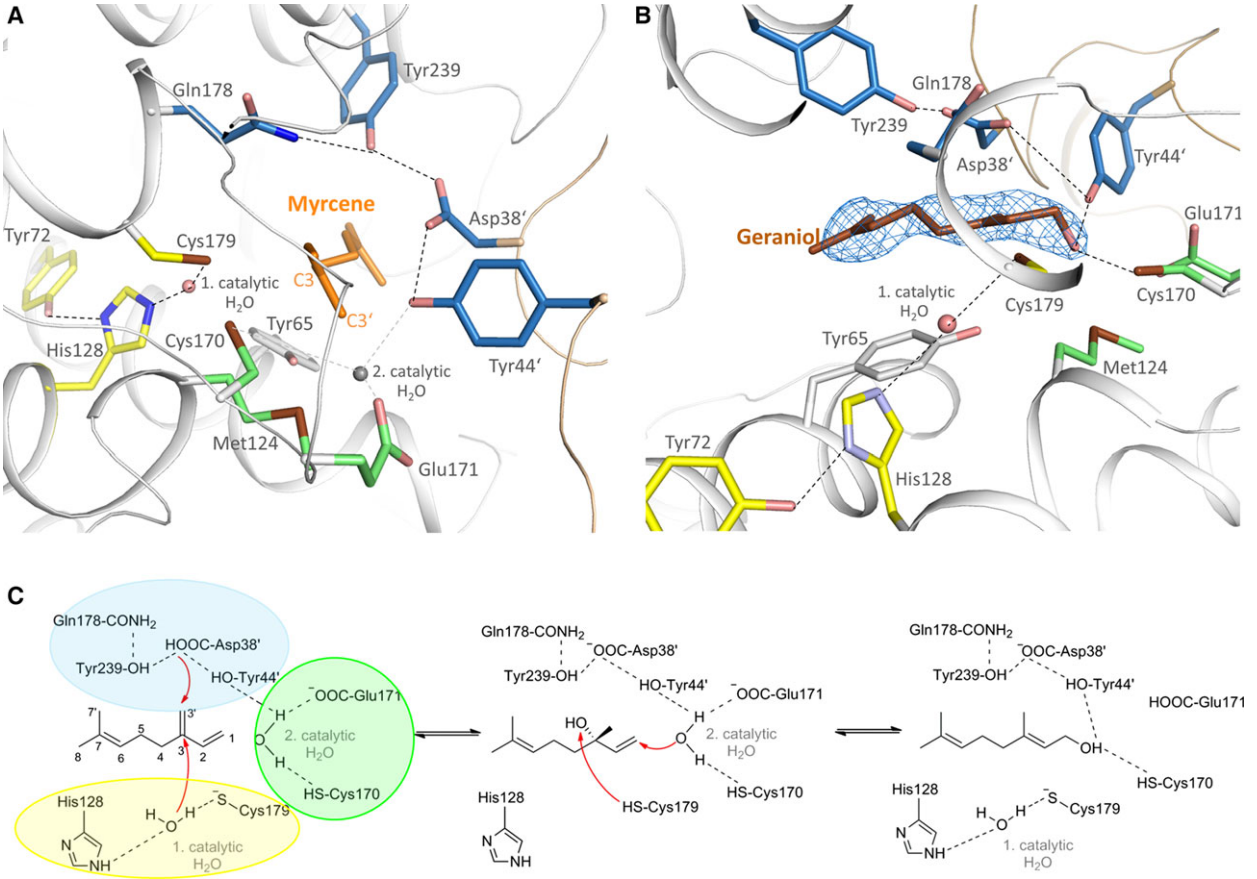

Fig. 4. Catalytic reaction of Ldi. (A) The dehydration reaction. The active site is subdivided into three catalytic clusters termed cluster I (yellow), II (blue), and III (green) grouped around myrcene (orange). The 2. catalytic H_2O (grey) was modelled between Tyr44′, Cys170 and Glu171. (B) The hydroxyl mutase reaction. The channel with the highest occupied geraniol and the three catalytic clusters were shown. The electron density of geraniol has a contour level of $\sigma = 0.8$. (C) Postulated catalytic mechanism.

conditions and structurally analyzed at 2.45 Å resolution (Table 1). Significant extra electron density was clearly visible at the inferred substrate binding site at the bottom of the five channels in the asymmetric unit (Figs 3 and 4). Although the terpene binding sites were incompletely and partially inhomogeneous occupied, the best fit was realized with β-myrcene in two channels and geraniol in the three remaining channels (Figs 3B and 4B). Because the crystals were soaked with linalool at pH 9, the pH optimum for enzymatic activity, this finding implicates a dehydratase or isomerase reaction in the Ldi crystal (Fig. 1). The monoterpenes are embedded into the channel such that the unpolar methylpentene tail points to the channel entrance while the buta-1,3-diene or but-2-en-1-ol heads of β-myrcene and geraniol, respectively, are directed to the channel bottom (Fig. 3), the site of the dehydratase and hydroxyl mutase reaction. For myrcene, its buta-1,3-diene moiety is flanked by Tyr65, Met124, Cys170, Phe176, Gln178, Cys179, Tyr239, Asp38′, Phe39′ and Tyr 44′ and its methylpentene moiety is surrounded by various hydrophobic residues including Val62, Tyr65, Phe69, Tyr239, Trp243, Leu294, Phe298, Leu341, Asp38′ and Phe39′. Tyr65 stacks against the entire terpene molecule and a site-directed exchange to phenylalanine only moderately reduced the enzymatic activity, suggesting that Tyr65 is not involved in acid/base catalysis.

As no organic cofactor or metal ions were found, the reaction apparently proceeds by an acid/base mechanism using protonable amino acid side chains. The functional groups of the monoterpenes comprising the hydroxy groups of (*S*)-linalool and geraniol and the C3-C3′ double bond of β-myrcene are contacted by several polar side chains subdivided into three catalytic clusters (Fig. 4). Cluster I includes Cys179, a firmly bound H_2O (1. catalytic H_2O), His128 and Tyr72 arranged in a line hydrogen-bonded to each other. Notably, Cys179 is located at the N-terminus of helix 178–192 which might stabilize a deprotonated thiolate as described for diaminopimelate epimerase [34]. The terpene is contacted from the opposite side of the channel by cluster II composed of Asp38′, Tyr44′, Tyr239 and Gln178. Moreover, the terpene-free Ldi structure, determined at higher resolution, reveals additional visible H_2O adjacent to cluster II that might move towards C3 and C3′ of the terpene. Both clusters contact the C3-C3′ group of the terpenes and this ensures the regioselectivity of the hydration reaction. Cluster III at the apex of the channel mainly consists of Cys170, Glu171 and Tyr44′. These residues might activate one H_2O modelled into an empty space in front of the β-myrcene C1 (and equivalently of the (*S*)-linalool C1) (Fig. 4A). This 2. catalytic H_2O, which partly overlaps with the C1 hydroxy group of geraniol of the superimposed Ldi-geraniol complex, is involved in the hydroxyl mutase reaction (Fig. 4B). Site-directed mutagenesis of His128, Cys170 or Cys179 completely inactivated the enzyme.

Discussion

In principle, all three catalytic clusters have the competence to perform the protonation/deprotonation and/or hydration/dehydration steps required for the dehydratase and isomerase reaction. Therefore, the postulated catalytic mechanism (Fig. 4C) is not only based on the presented structural data of LDI but also on biochemical and primary structure data of Lis, the enzyme that only catalyzes the hydroxyl mutase reaction [15] (Table S2). The potentially monomeric enzyme only contains the catalytic clusters I and III as Cys170, Cys179 and Glu171 are conserved. Asp38′ and Tyr44′ of the capping segment including Cys48′ and Cys101′ are absent (Fig. 2B). Accordingly, the hydroxyl mutase reaction from geraniol to (*S*)-linalool of Ldi and of Lis might proceed by a nucleophilic attack of the 1. catalytic H_2O of cluster I on C3, an allylic rearrangement and a release of the C1-hydroxyl group (Fig. 4B) presumably assisted by Glu171 of cluster III acting as proton donor. The enzymatically crucial His128 is exchanged in Lis by a tryptophane and might therefore only serve in the hydroxyl mutase reaction as hydrogen bond donor but not as an acid/base catalyst. Dehydration of (*S*)-linalool to β-myrcene in Ldi consequently proceeds by deprotonation of C3′ via Asp38′ or Tyr44′ of cluster II essentially formed by the partner subunit and by dehydroxylating C3 via protonation, releasing the 1. catalytic H_2O of cluster I in an *anti*-selective manner (Fig. 4A,C). Both reactions have to be synchronized to achieve the previously observed stereospecificity of (*S*)-linalool formation [7]. The proposed scenario implicates that C3 points to cluster I and C3′ to cluster II. The Ldi-terpene complex was refined in several orientations of β-myrcene and geraniol but the quality of the electron density does not unambiguously allow the determination of the orientation of the roughly rotationally symmetric monoterpenes. However, the fitted terpene conformation is in line with the proposed enzymatic mechanism.

Hydration of an electron-rich carbon double bond with the poor nucleophile water is a chemically challenging reaction [35] and therefore nature normally prefers a conjugated electron-poor double bond using the Michael reaction for alcohol formation using

enzymes like aspartatases/fumarases, enoyl-CoA hydratases and aconitases [36–38]. Isomerization between (*S*)-linalool and geraniol is a basic chemical reaction, which was also found in a few other intramolecular hydroxyl transferases as i.e. isochorismate synthase [39]. Hydratases and isomerases use frequently histidines and aspartates/glutamates but rarely cysteines as (de)protonable groups. As experimentally verified by site-directed mutagenesis studies (Table S1) Cys170 and Cys179 of Ldi play a crucial role which is in accordance with its inactivity under aerobic and non-reducing conditions probably due to cysteine sulfoxide or Cys170-Cys179 disulfide formation (distance 4.1 Å) (Fig. 4). The use of cysteines in Ldi for acid/base catalysis is not without precedents. In glutamate racemase two cysteines act as acid/base [40] and in malate isomerase one cysteine acts as nucleophile forming a covalent thioether bond and one as proton donor [41].

A sustainable world requires the transformation of renewable resources into intermediates of the fine chemical industry. Low cost-production of butane-1,4-diol from feedstocks by engineered microorganisms has already been developed [42]. For butadiene and isoprene, several patents used linalool dehydratases or derivatives thereof for the final step in the synthesis of the olefins [8–13]. The presented structural data of Ldi revealed the terpene binding site between two monomers inside a hydrophobic channel, unprecedentedly for $(\alpha,\alpha)_6$ barrel proteins, and three catalytic clusters involved in catalysis (Figs 2–4). This knowledge provides a rational basis for expanding the substrate specificity of the enzyme and for increasing its turnover rate. From the new platform more targeted biotechnological avenues can be explored.

Acknowledgements

This work was supported by the German Research Foundation (DFG) in the SPP1319 and the Max-Planck Society. We thank Hartmut Michel and Friedrich Widdel for continuous support and the staff of PXII at the Swiss Light Source (Villigen) for help during data collection.

Authors' contributions

SW, RM, UE and JH designed the experiments. RM performed the protein expression, protein purification and generation and analysis of mutants. SW performed the protein crystallography and data evaluation. SW, RM, UE and JH drafted the manuscript. All authors read and approved the final manuscript.

References

1 Dudareva N, Klempien A, Muhlemann JK and Kaplan I (2013) Biosynthesis, function and metabolic engineering of plant volatile organic compounds. *New Phytol* **198**, 16–32.
2 Günther A, Hewitt CN, Erickson D, Fall R, Geron C, Graedel T, Harley P, Klinger L, Lerdau M, Mckay WA *et al.* (1995) A global model of natural volatile organic compound emission. *J Geophys Res Atmos* **100**, 8873–8892.
3 Ziemann PJ and Atkinson R (2012) Kinetics, products, and mechanisms of secondary organic aerosol formation. *Chem Soc Rev* **41**, 6582–6605.
4 Wilt FM, Miller GC, Everett RL and Hackett M (1993) Monoterpene concentrations in fresh, senescent, and decaying foliage of single-leaf pinyon (*Pinus monophylla* Torr and Frem: Pinaceae) from the Western Great-Basin. *J Chem Ecol* **19**, 185–194.
5 Marmulla R and Harder J (2014) Microbial monoterpene transformations – a review. *Front Microbiol* **5**, 346.
6 Brodkorb D, Gottschall M, Marmulla R, Lüddeke F and Harder J (2010) Linalool dehydratase-isomerase, a bifunctional enzyme in the anaerobic degradation of monoterpenes. *J Biol Chem* **285**, 30436–30442.
7 Lüddeke F and Harder J (2011) Enantiospecific (*S*)-(+)-linalool formation from beta-myrcene by linalool dehydratase-isomerase. *Z Naturforsch C* **66**, 409–412.
8 Pearlman PS, Chen C, Botes AL and van Eck Conradie A (2013) Methods for biosynthesizing 1,3 butadiene. WO2013082542.
9 Campbell P, Bredow S, Zhou H, Doneske S and Monticello DJ (2013) Microorganisms and processes for the production of isoprene. WO2013173437.
10 Coelho PS, Farrow MF and Smith MA (2014) De novo metabolic pathways for isoprene biosynthesis. WO2014066892.
11 Marlière P (2014) Production of volatile dienes by enzymatic dehydration of light alkenols. WO2014033129.
12 Koch DJ, Lopes MSG, Gouvea IE, Zeidler AFB, Dumaresq ASR and Rodrigues MEP (2015) Modified microorganisms and methods of using same for anaerobic coproduction of isoprene and acetic acid. WO2015002977.
13 Botes AL and van Eck Conradie A (2015) Methods for biosynthesis of isoprene. WO2015021045.
14 Bailey MP (2014) The future of butadiene. *Chem Eng* **121**, 19–24.
15 Marmulla R, Šafarić, B, Markert S, Schweder T and Harder J (2016) Linalool isomerase, a membrane-anchored enzyme in the anaerobic monoterpene degradation *Thauera linaloolentis* 47 Lol. *BMC Biochem* **17**, 6.

16 Bradford MM (1976) Rapid and sensitive method for quantitation of microgram quantities of protein utilizing principle of protein-dye binding. *Anal Biochem* **72**, 248–254.
17 Kabsch W (2010) Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125–132.
18 Schneider TR and Sheldrick GM (2002) Substructure solution with SHELXD. *Acta Crystallogr D Biol Crystallogr* **58**, 1772–1779.
19 Vonrhein C, Blanc E, Roversi P and Bricogne G (2007) Automated structure solution with autoSHARP. *Methods Mol Biol* **364**, 215–230.
20 Jones TA, Zou JY, Cowan SW and Kjeldgaard M (1991) Improved methods for building protein models in electron-density maps and the location of errors in these models. *Acta Crystallogr A* **47**, 110–119.
21 Cowtan KD (1994) Joint CCP4 and ESF-EACBM newsletter on protein crystallography.
22 Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235–242.
23 Cowtan K (2006) The Buccaneer software for automated model building. *Acta Crystallogr D Biol Crystallogr* **62**, 1002–1011.
24 Emsley P and Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126–2132.
25 Murshudov GN, Vagin AA and Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240–255.
26 Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213–221.
27 Holm L and Rosenstrom P (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res* **38**, W545–W549.
28 Schrödinger LLC (2010) The PyMOL molecular graphics system, version 1.4.1.
29 Yang J, Yan R, Roy A, Xu D, Poisson J and Zhang Y (2015) The I-TASSER Suite: protein structure and function prediction. *Nat Methods* **12**, 7–8.
30 Itoh T, Ochiai A, Mikami B, Hashimoto W and Murata K (2006) Structure of unsaturated rhamnogalacturonyl hydrolase complexed with substrate. *Biochem Biophys Res Commun* **347**, 1021–1029.
31 Fujiwara T, Saburi W, Inoue S, Mori H, Matsui H, Tanaka I and Yao M (2013) Crystal structure of *Ruminococcus albus* cellobiose 2-epimerase: structural insights into epimerization of unmodified sugar. *FEBS Lett* **587**, 840–846.
32 Wendt KU and Schulz GE (1998) Isoprenoid biosynthesis: manifold chemistry catalyzed by similar enzymes. *Structure* **6**, 127–133.
33 Wendt KU, Poralla K and Schulz GE (1997) Structure and function of a squalene cyclase. *Science* **277**, 1811–1815.
34 Pillai B, Cherney MM, Diaper CM, Sutherland A, Blanchard JS, Vederas JC and James MN (2006) Structural insights into stereochemical inversion by diaminopimelate epimerase: an antibacterial drug target. *Proc Natl Acad Sci USA* **103**, 8668–8673.
35 Resch V and Hanefeld U (2015) The selective addition of water. *Catal Sci Technol* **5**, 1385–1399.
36 Agnihotri G and Liu H-W (2003) Enoyl-CoA hydratase: reaction, mechanism, and inhibition. *Bioorg Med Chem* **11**, 9–20.
37 Lauble H, Kennedy MC, Beinert H and Stout CD (1994) Crystal structures of aconitase with trans-aconitate and nitrocitrate bound. *J Mol Biol* **237**, 437–451.
38 Puthan Veetil V, Fibriansah G, Raj H, Thunnissen A-MWH and Poelarends GJ (2012) Aspartase/fumarase superfamily: a common catalytic strategy involving general base-catalyzed formation of a highly stabilized aci-carboxylate intermediate. *Biochemistry* **51**, 4237–4243.
39 Sridharan S, Howard N, Kerbarh O, Błaszczyk M, Abell C and Blundell TL (2010) Crystal structure of *Escherichia coli* enterobactin-specific isochorismate synthase (EntC) bound to its reaction product isochorismate: implications for the enzyme mechanism and differential activity of chorismate-utilizing enzymes. *J Mol Biol* **397**, 290–300.
40 Glavas S and Tanner ME (2001) Active site residues of glutamate racemase. *Biochemistry* **40**, 6199–6204.
41 Fisch F, Fleites CM, Delenne M, Baudendistel N, Hauer B, Turkenburg JP, Hart S, Bruce NC and Grogan G (2010) A covalent succinylcysteine-like intermediate in the enzyme-catalyzed transformation of maleate to fumarate by maleate isomerase. *J Am Chem Soc* **132**, 11455–11457.
42 Yim H, Haselbeck R, Niu W, Pujol-Baxley C, Burgard A, Boldt J, Khandurina J, Trawick JD, Osterhout RE, Stephen R *et al.* (2011) Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol. *Nat Chem Biol* **7**, 445–452.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:
Table S1. Site-specific amino acid exchange in Ldi.
Table S2. Active site-specific amino acids of Ldi and the corresponding amino acids of Lis in a superimposed structural model.