Supporting Information Appendix

Inside the closed complex of class I terpene synthases: identifying universal structural features governing catalysis

Patrick Schrepfer[1], Alexander Büttner[1], Christian Görner[1], Michael Hertel[1], Frank Wallrapp[2], Wolfgang Eisenreich[3], Volker Sieber[4], Robert Kourist[5] & Thomas Brück[1*]

[1] Department for Industrial Biocatalysis, Technische Universität München, Lichtenbergstr. 4, 85748 Garching, Germany

[2] Department for Bioinformatics and Computational Biology, Technische Universität München, Boltzmannstraße 3, 85748 Garching, Germany

[3] Chair of Biochemistry, Technische Universität München, Lichtenbergstr. 4, 85748 Garching, Germany

[4] Chair of Chemistry of Biogenic Resources, Technische Universität München, Schulgasse 16, 94315 Straubing, Germany [5]Junior Research Group for Microbial Biotechnology, Ruhr-Universität Bochum, Universitätsstraße 150, 44780 Bochum, Germany

Content

General experimental procedures

All chemicals were obtained from standard sources at the highest purity grade. NMR spectra were recorded in $CDCl_3$ with an Avance-III 500 MHz device (Bruker) at 300 K. ^1H NMR chemical shifts are given in ppm relative to $CHCl_3$ (δ=7.26 ppm) or CD_2HOD (δ=3.31 ppm) and CD_3OH (δ=4.87) (^1H NMR). ^{13}C NMR chemical shifts are given in ppm relative to $CDCl_3$ at δ=77.16 ppm or CD_3OH at δ=49.00 ppm. The 2D experiments (HSQC, HMBC, TOCSY, COSY and NOESY) were performed using standard Bruker pulse sequences and parameters.

GC-MS and GC-FID analysis of diterpene products from n-hexane extractions was conducted by a Trace GC Ultra with DSQII (Thermo Scientific). One µl sample was applied by TriPlus AS onto a SGE BPX5 column (30 m, I.D 0.25 mm, Film 0.25 µm). The initial column temperature was 50°C (maintained for 2.5 min). A temperature gradient was applied from 50°C – 320°C (10 °C/min), followed by 3 min maintenance at 320°C. MS data were recorded at 70 eV (EI), *m/z* (rel. intensity in %) as TIC, total ion current. The recorded *m/z* range was 50 – 650. Quantification was performed with flame ionization detector (FID) using 1 mg mL^{-1} α-humulene (Sigma-Aldrich, Germany) as an internal standard.

High-resolution mass spectra of diterpenes were determined with a Thermo Scientific DFS Magnetic Sector GC-HRMS system from n-hexane extractions. One µl sample was applied by TriPlus AS onto a SGE BPX5 column (30 m, I.D 0.25 mm, Film 0.25 µm). The initial column temperature was 50°C (maintained for 2.5 min). A temperature gradient was applied from 50°C – 320°C (10 °C min^{-1}), followed by 3 min maintenance at 320°C. The recorded *m/z* range was 50 – 650.

Circular dichroism (CD) spectroscopy was performed using a Chirascan plus spectropolarimeter (Applied Photophysics, United Kingdom). Samples were dissolved in acetonitrile and spectra were recorded in quartz cuvettes with 0.1 cm path length at 20 °C.

Glycerol content of the fed-batch bioprocess was quantified by HPLC, using an Agilent LC 1100 system (Agilent technologies, Waldbronn, Germany), equipped with an autosampler, column oven and a Shodex RI-101 detector (Showa Denko Europe GmbH, Munich). A Rezex ROA-Organic Acid H+ (8%) ion-exclusion column (300 mm, 7.8 mm internal diameter; Phenomenex LTD, Aschaffenburg, Germany) was used for the isocratic separation with 5 mM sulfuric acid at a flow rate of 0.5 ml min^{-1} at 70°C.

Bacterial strains, genes and vectors

The *E. coli* strains XL-1 Blue and BL21(DE3) were used for cloning and diterpene production. All strains and plasmids were obtained from Novagen/Merck Millipore (Germany). Genes were synthesized by Life technologies GmbH (Thermo Fisher Scientific) featuring the appropriate restriction sites and adjusting codon usage for *E. coli*.

Genes and vectors used for *in vitro* production of taxa-4,11-diene

pET28b (+) vector was used for cloning and recombinant expression of a codon optimized version of the M60 truncation of taxadiene synthase from *Taxus brevifolia* (henceforth designated "TXS").[1] The Gene of taxadiene synthase (*txs*) (GenBank: AF326519.1) (**See Genes section**) was synthesized as M60 truncation including an NdeI restriction site at the 5`- and an XhoI restriction site at the 3`-end. The synthetic gene was introduced into the corresponding restriction site of pET28b (+) vector by standard cloning techniques.

Genes and vectors used for *in vivo* production of taxa-4,11-diene

The 1-deoxy-D-xylulose 5-phosphate (DXP) pathway was housed by the plasmids pColaDuet-1 and pCDFDuet-1, while the taxadiene biosynthesis genes were carried by the pETDuet-1 plasmid. To overexpress the DXP pathway in *E. coli* BL21(DE3), genes from *E. coli* of the 1-deoxy-D-xylulose 5-phosphate synthase (*dxs*) (GenBank: YP_001461602.1) (**See Genes section**), 1-deoxy-D-xylulose 5-phosphate reductoisomerase (*dxr*) (GenBank: NP_414715.1) (**See Genes section**), 2-C-methyl-D-erythriol 4-phosphate cytidyltransferase synthase (*ispD*) (GenBank: NP_417227.1) (**See Genes section**), 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (*ispF*) (GenBank: NP_289295.1) (**See Genes section**) and Isopentenyl-diphosphate delta isomerase (*idi*) (GenBank: NP_417365.1) (**See Genes section**) were synthesized. *lspD*/*ispF* was created as a bi-cistronic operon (**See Genes section**). The synthetic genes were introduced into the appropriate plasmids according to **Table S1** by standard cloning techniques. To achieve biosynthesis of taxadiene and its deprotonated intermediates the native geranylgeranyl diphosphate synthase (*crte*) (GenBank: M90698.1) was amplified from *Pantoea agglomerans* (ATCC 27155) using standard protocols. Primers used were 5'-AAA CCA TGG CAA TGG CAA CGG TCT GCG CA-3' and 5'-AAA GAA TTC TTA ACT GAC GGC AGC GAG TTT-3'. The genes of *crte* and synthetic *txs* were introduced into the appropriate plasmids according to **Table S2** by standard cloning techniques.

Site directed mutagenesis

Mutation primers were designed applying the Agilent Technologies QuikChange program (Table S4). For PCR the following mixture was used: 5 µL Pfu Ultra buffer (10x, Agilent), 1 µL forward/ reverse primer (2 µM stock) (**Table S5**), 1 µL dNTPs (10 µM stock), 200 ng template DNA (plasmid), 1 µL Pfu Cx Hotstart polymerase (Agilent) and

water to a final volume of 50 µL. After the PCR reaction, 2 µL of DpnI (20 u, NEB) were added and the mixture was incubated at 37 °C for at least 2 h. 2 µL of sample were transformed into XL1 blue competent cells (Agilent). After 1 h of incubation at 37 °C in SOB medium, cells were streaked out on LB (lysogeny broth)-kanamycin or in case of the *in vivo* approach -ampicillin agar plates. Colonies were grown overnight at 37°C. Single colonies were picked and inoculated in LB-kanamycin or -ampicillin medium overnight. Plasmids were isolated applying a plasmid miniprep kit (Agilent). The mutant TXS genes were verified via sequencing (MWG Eurofins, Ebersberg) using the T7 forward/T7 reverse primer set in case of pET28b (+) or DuetUp2 (Novagen)/T7 reverse primer set in case of pET duet vector.

Table S1: Plasmids used to construct the overexpressed DXP pathway in *E. coli* BL21(DE3)

Gene(s)	***Vector***	***Multiple Cloning Site***	***Restriction Sites***
dxr	pColaDuet-1	I	*Nco*I, *EcoR*I
dxs	pColaDuet-1	II	*Nde*I, *Xho*I
ispD/*ispF* operon	pCDFDuet-1	I	*Nco*I, *EcoR*I
idi	pCDFDuet-1	II	*Nde*I, *Xho*I

Table S2: Plasmids used to construct taxadiene biosynthesis in *E. coli* Bl21(DE3)

Gene(s)	***Vector***	***Multiple Cloning Site***	***Restriction Sites***
crtE	pET-Duet-1	I	*Nco*I, *EcoR*I
txs	pET-Duet-1	II	*Nde*I, *Xho*I

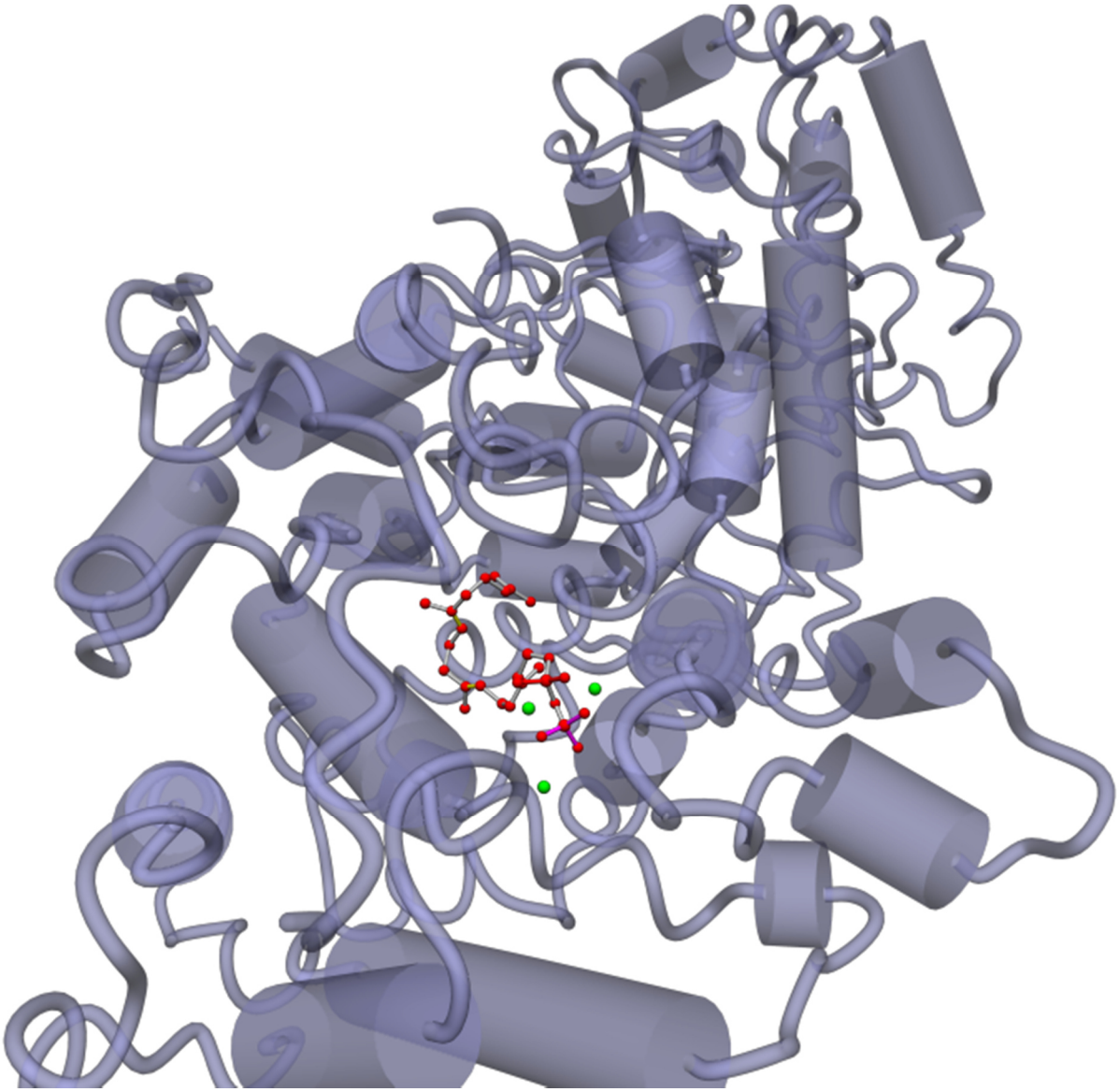


Fig. S1a: Closed conformation of TXS harboring productive GGPP

Closed conformation of TXS harboring productive **GGPP** (red). Magnesium ions are shown in green.

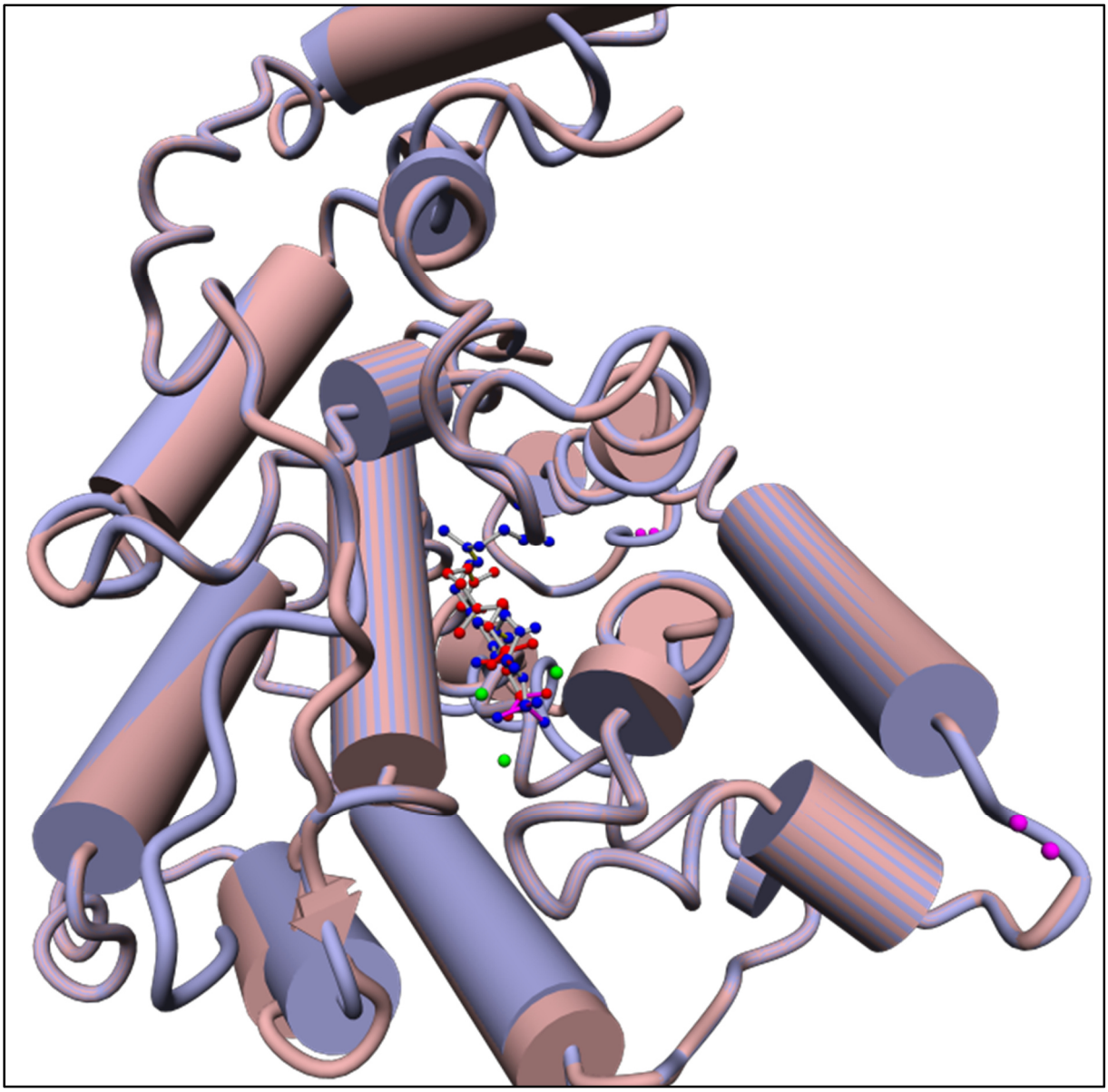


Fig. S1b: Structural superposition of BPPS and TXS

Structural superposition of the closed conformation of BPPS (PDB: 1N20)[2] (light red) with TXS (light blue). The alignment exhibits a root mean square deviation (RMSD) of 0.304Å over 331 AAs with 31.42 % sequence identity. In BPPS and TXS only the amino acids (AA) characterizing the class I domains (BPPS: AA54-99 + AA 272-598; TXS: AA80-130 + AA537-862) are shown. The substrate of BPPS 3-aza-2,3-dihydrogeranyl diphosphate is shown in red, the substrate of TXS geranylgeranyl pyrophosphate (**GGPP**) is shown in blue. Magnesium ions are shown in green.

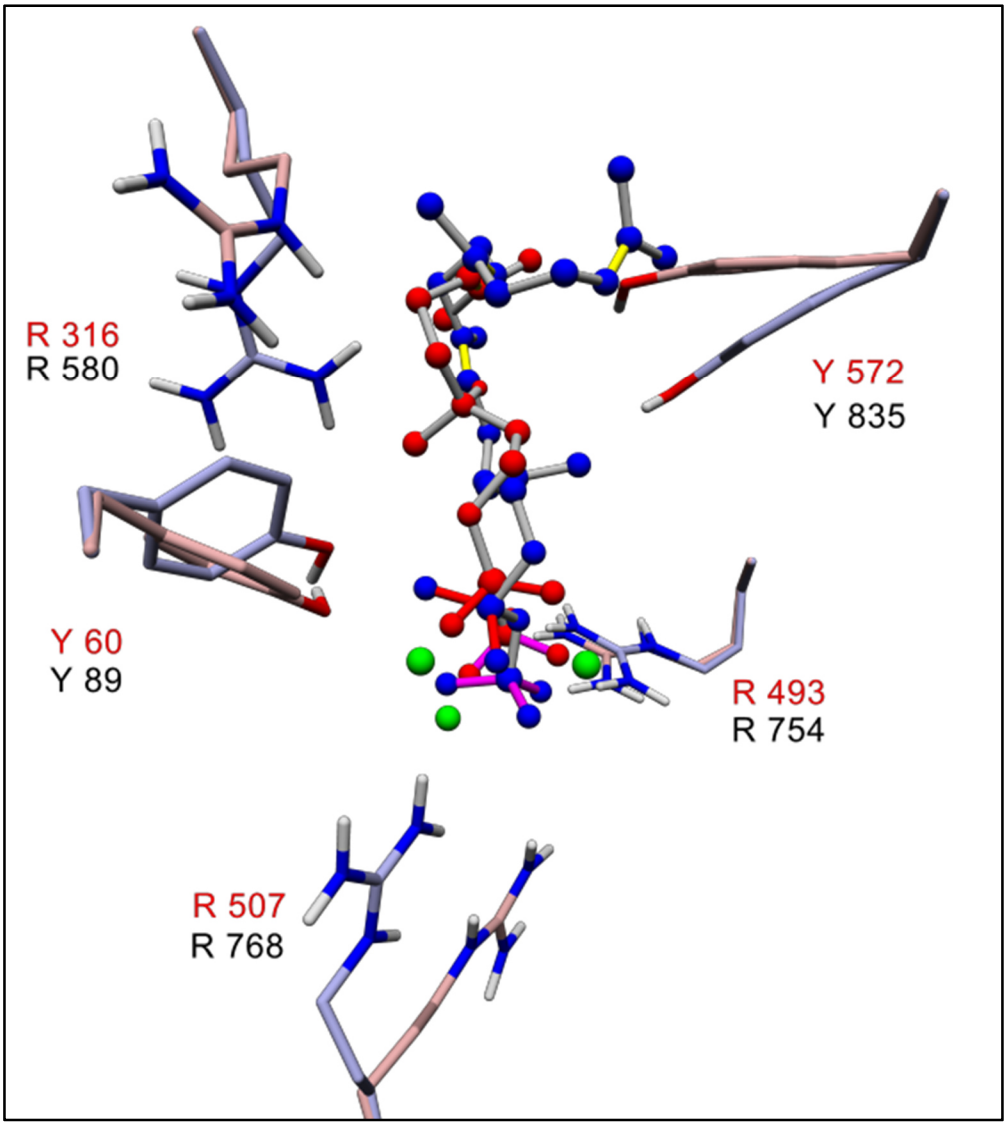


Fig. S1c: Structural superposition of the H-bond donor amino acid network of BPPS and TXS

Structural superposition of the closed conformation of BPPS (PDB: 1N20) (light red) with TXS (light blue). The amino acids representing the H-bond donor network of BPPS and TXS are shown (BPPS residues in red, the corresponding TXS residues in black). Substrate colours are corresponding to **Fig. S1a**.

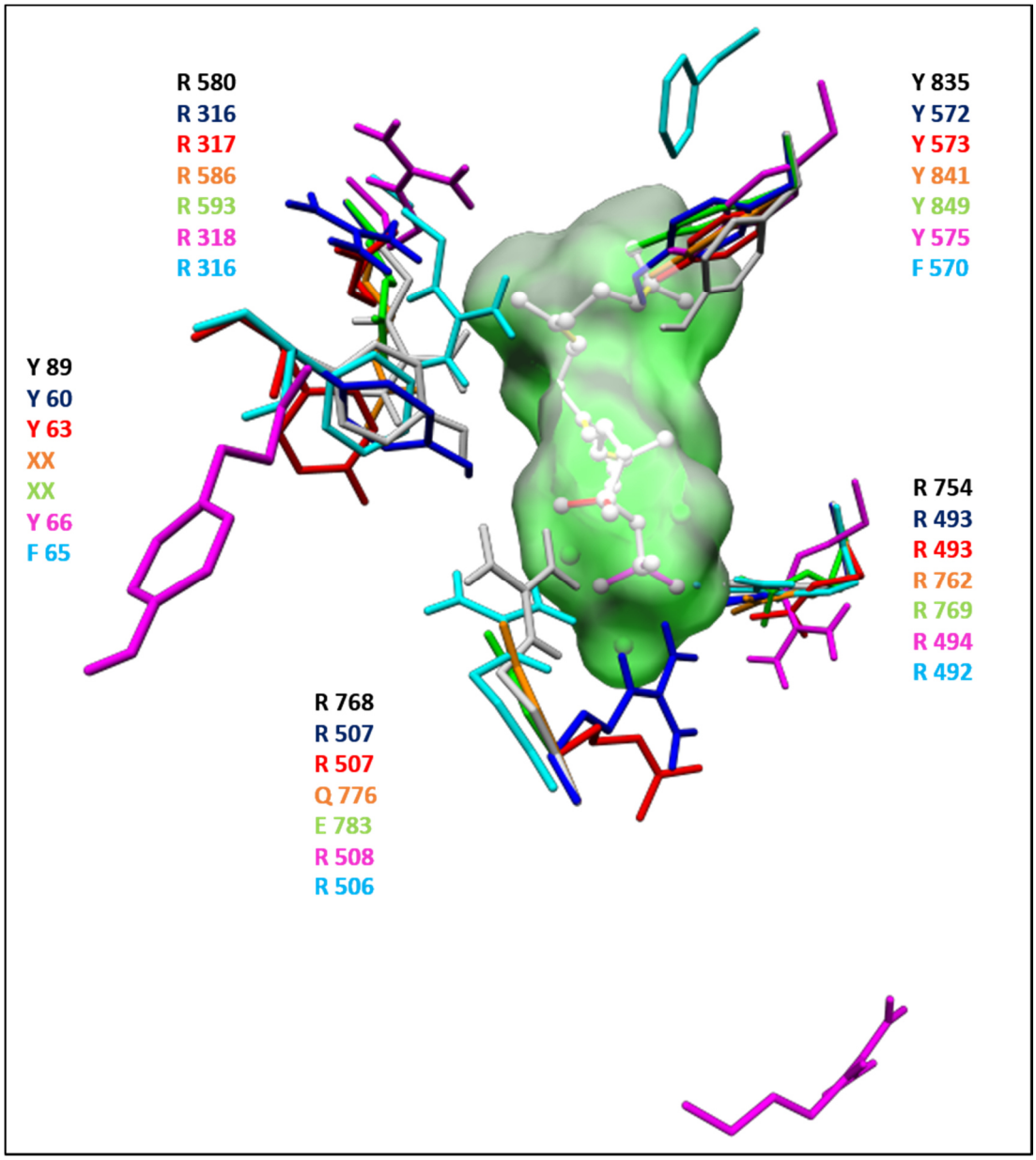


Fig. S1d: Structural superposition of eukaryotic mono- and diterpene synthases

A structural superposition of the H-bond donor amino acid network of eukaryotic mono- and diterpene synthases is shown. TXS in closed conformation (diterpene synthase: residues in gray, labelling in black, **GGPP** in gray with molecular surface), BPPS in closed conformation (monoterpene synthase: blue (PDB: 1N20)), limonene synthase from *M. spicata* in closed conformation (mono terpene synthase: red (PDB: 2ONG)), class I domain of abietadiene synthase from *A. grandis* (bifunctional diterpene synthase: orange (PDB: 3S9V)), class I domain of isopimaradiene synthase from *P. abies* (bifunctional diterpene synthase: green (homology model, this work)), casbene synthase from *J. curcas* (diterpene synthase: magenta (open conformation homology model, this work)) and cembratriene-ol synthase (CBTS) from *N. tabacum* (diterpene synthase: light blue (closed conformation homology model, this work)). Note that in the bifunctional diterpene synthases the corresponding second tyrosine (XX) is part of the class II domain and thus could not be taken into account.

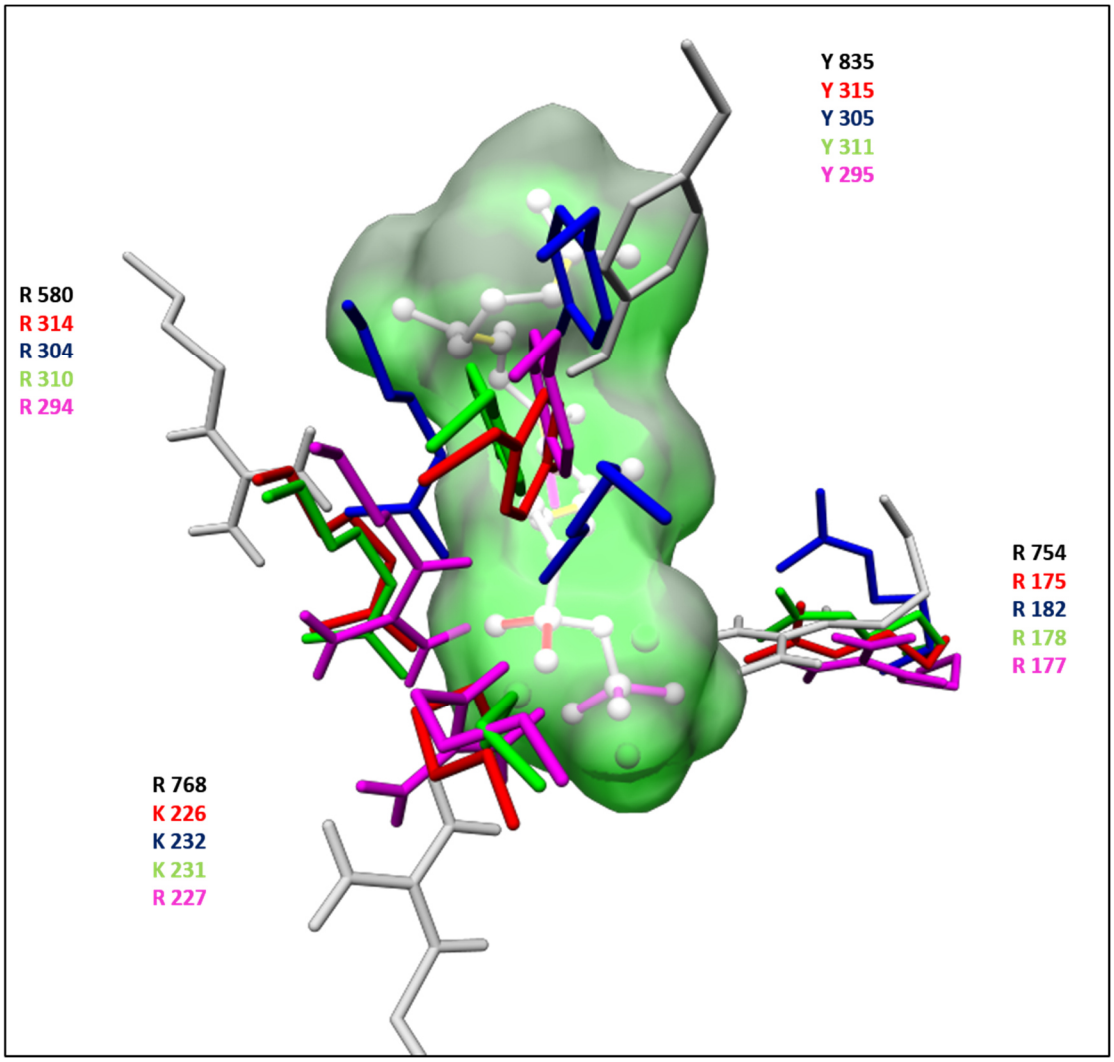


Fig. S1e: Structural superposition of bacterial and fungal sesqui- and diterpene synthases with TXS

A structural superposition of the H-bond donor amino acid network of bacterial and fungal sesqui- and diterpene synthases is shown. TXS in closed conformation (diterpene synthase: residues in gray, labelling in black, **GGPP** in gray with molecular surface), aristolochene synthase from *A. terreus* in closed conformation (fungal sesquiterpene synthase: red (PDB: 2OA6)), trichodiene synthase from *F. sporotrichioides* in closed conformation (fungal sesquiterpene synthase: blue (PDB: 1JFG)), selinadiene synthase form *S. pristinaespiralis* in closed conformation (bacterial sesquiterpene synthase: green (PDB: 4OKZ)) and cyclooctat-5-en-7-ol (CotB2) from *S. melanosporofaciens* (bacterial diterpene synthase: magenta, closed conformation model, this work: open conformation crystal structure PDB: 4OMG)). Note that bacterial and fungal sesqui- and diterpene synthase comprise a different overall fold with respect to TXS. Nevertheless the class I domain active site cavities correspond to that of their eukaryotic counterparts. The second corresponding tyrosine residue (Y89 in TXS) does not exist in bacterial and fungal sesqui- and diterpene synthases.

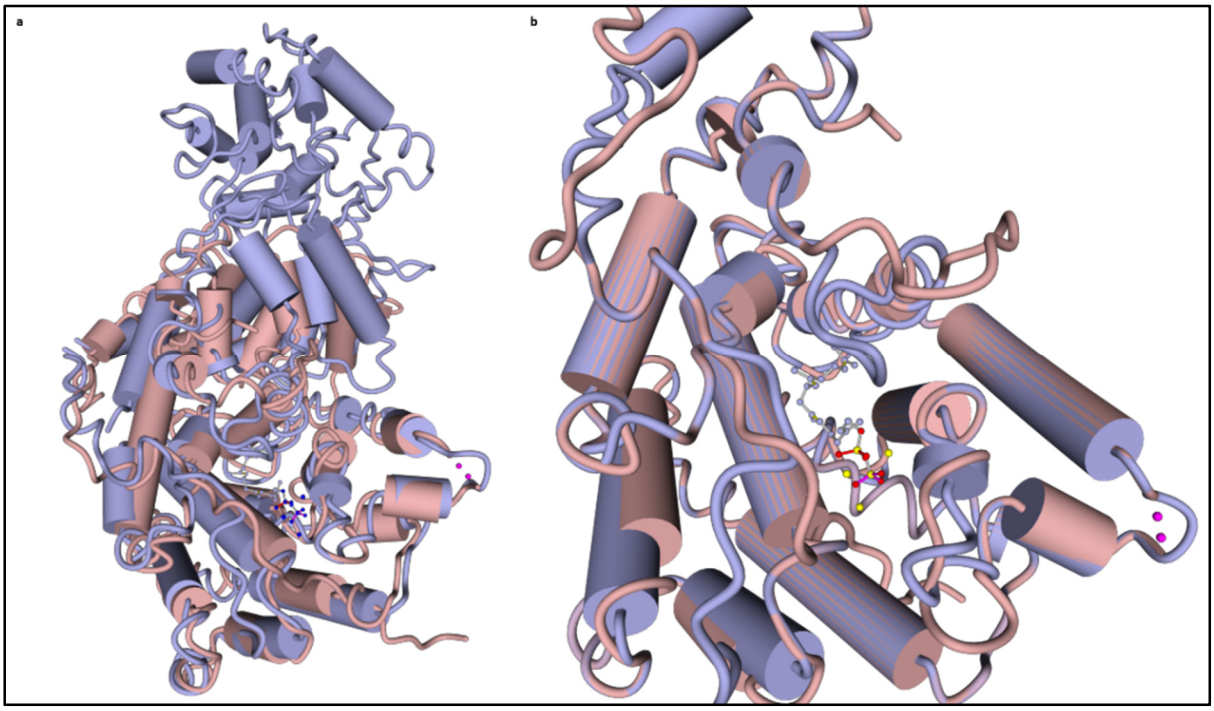


Fig. S1f: Structural superposition of TXS with the open conformation of CBTS from *N. tabacum*

(a) Structural superposition of TXS harboring **GGPP** (AA 80-862) with the open conformation homology model of CBTS (amino acids 50-598) from *N. tabacum*. **(b)** Structural superposition of the class I domain of TXS (AA80-130 + AA537-862) with the class I domain of CBTS (AA 50-105 + AA272-598).

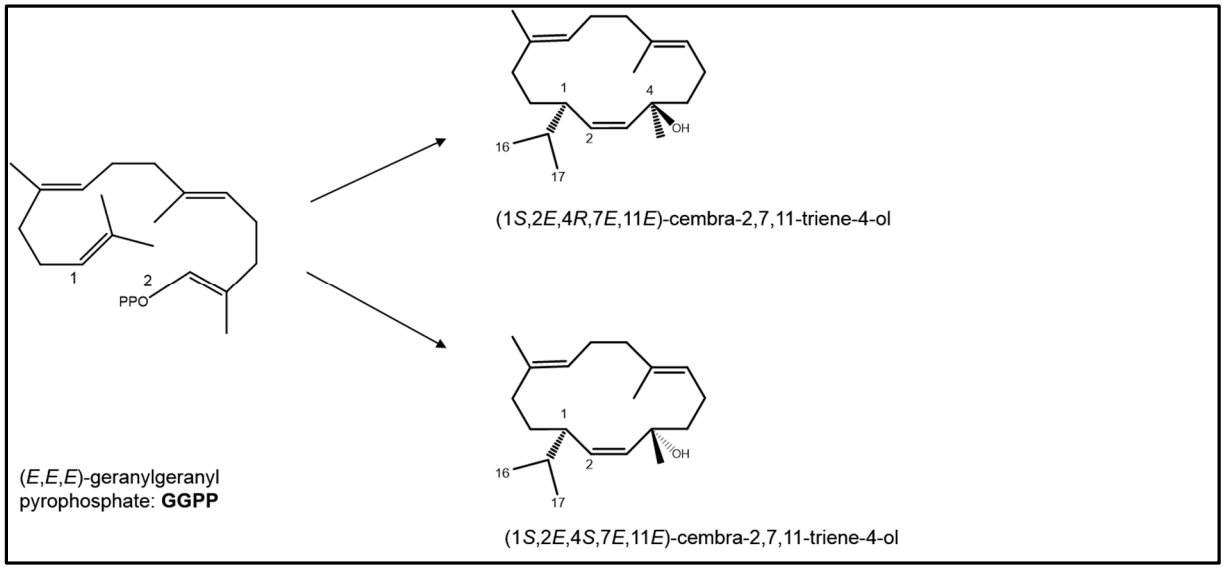


Fig. S1g: Reaction catalysed by CBTS

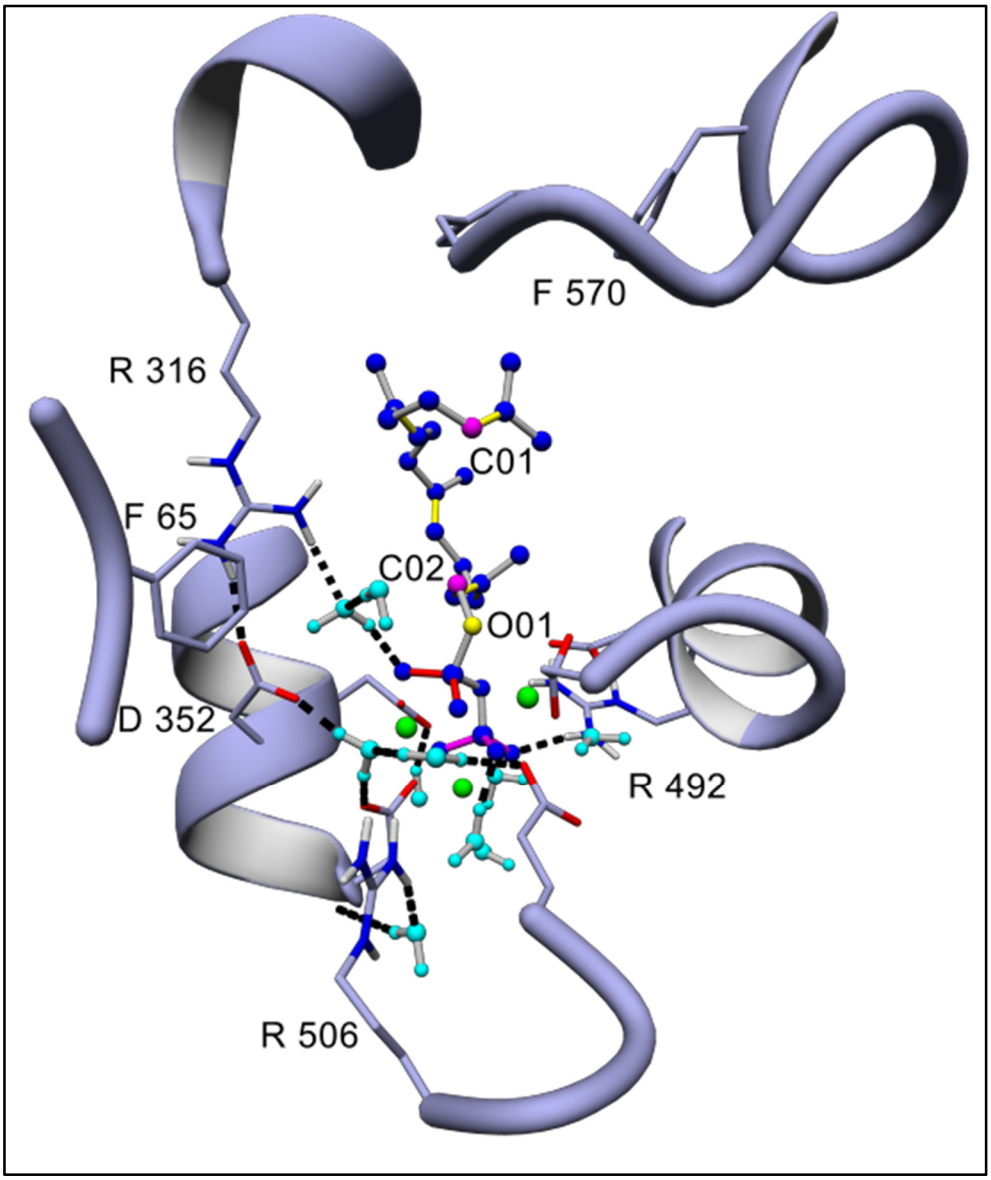


Fig. S1h: Closed conformation homology model of CBTS harboring GGPP

Closed conformation homology model of CBTS harboring **GGPP** and the H-bond donor amino acid network (R492, R506 and R570). The tyrosine residues facilitating binding of remaining bulk solvent water in TXS (Y89 and Y835) are natively replaced by F65 and F570, respectively. **GGPP** is coloured in blue, magnesium ions in green, remaining water in cyan. O01 of PP_i is coloured in yellow, C01 and C02 supposed to bind in order to form a monocyclic carbocation are shown in magenta.

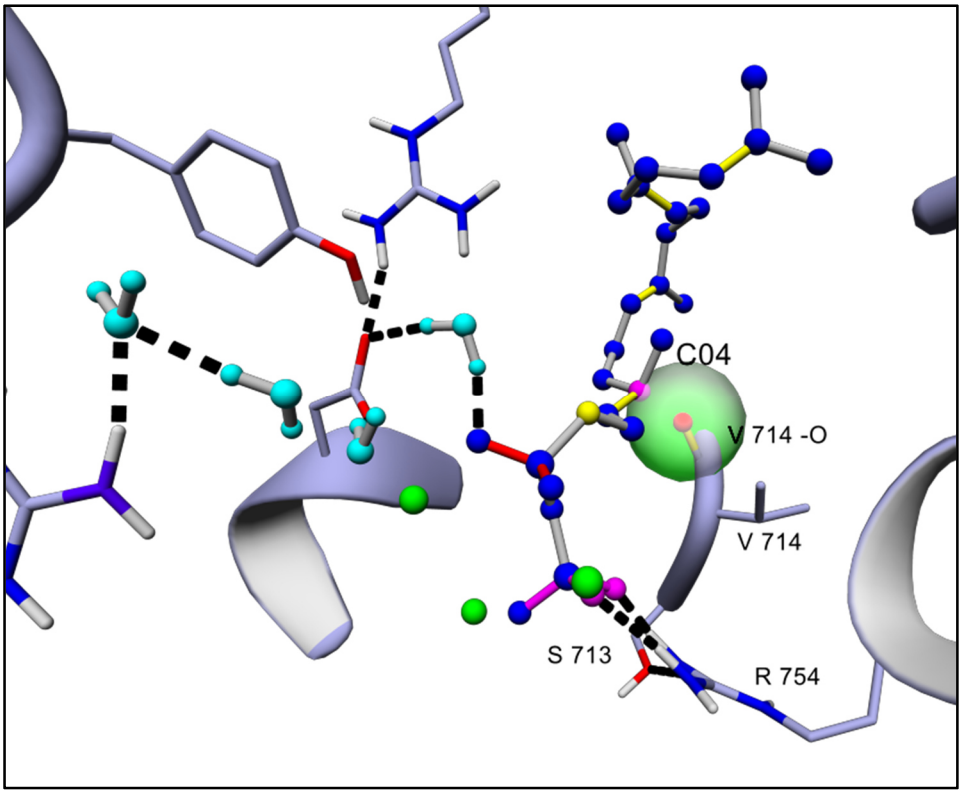


Fig. S2: Induced fit mechanism in TXS

Induced fit model of TXS. The model represents the time point immediately after active site closure. Upon substrate binding the PP_i sensor R754 has formed hydrogen bonds (black) to the O-atoms of **GGPP** (magenta) and to the linker S713. Interaction of the linker with the sensor leads to the effector V714 turning inwards. The V714 carbonyl atom points then directly towards the Δ3,4 of the substrate (magenta), which is thought to provoke abstraction of PP_i to form the geranylgeranyl cation and initiate the cyclization cascade. O01 of **GGPP** is shown in yellow, magnesium ion are shown in green.

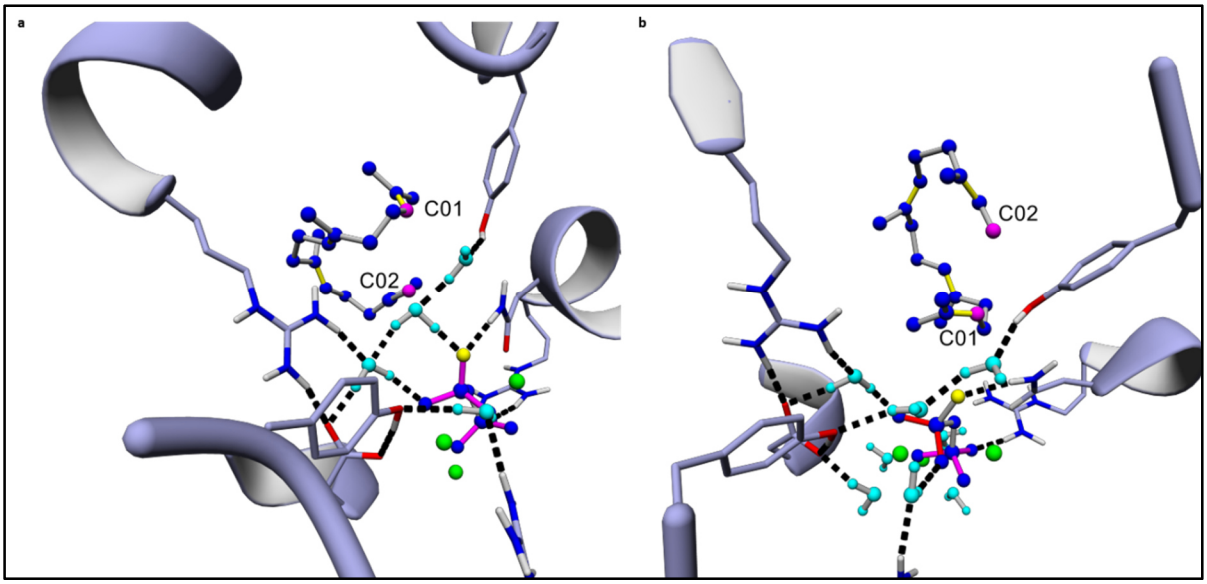


Fig. S3a: Molecular docking of QM-derived cation A into TXS

Molecular docking of QM-derived cation A into the active site of TXS by AutodockVina of YASARA structure. Cluster analysis of 999 docking runs shows two different conformation cluster of **cation A** in the active site. **(a)** Conformation A exhibits a binding energy of 7.206 kcal mol^{-1}, **(b)** conformation B exhibits a binding energy of 7.678 kcal mol^{-1}. Despite a higher binding energy of conformation B, the spatial positioning of conformation B in the active site does not correspond to the spatial positioning of productive **GGPP** in the modelled closed complex (**Fig. 2b**) and is therefore not in productive sequence **GGPP**->**cation A** during initiation of catalysis. Conformation A of **cation A** directly arises from productive **GGPP** after abstraction of PP_i.

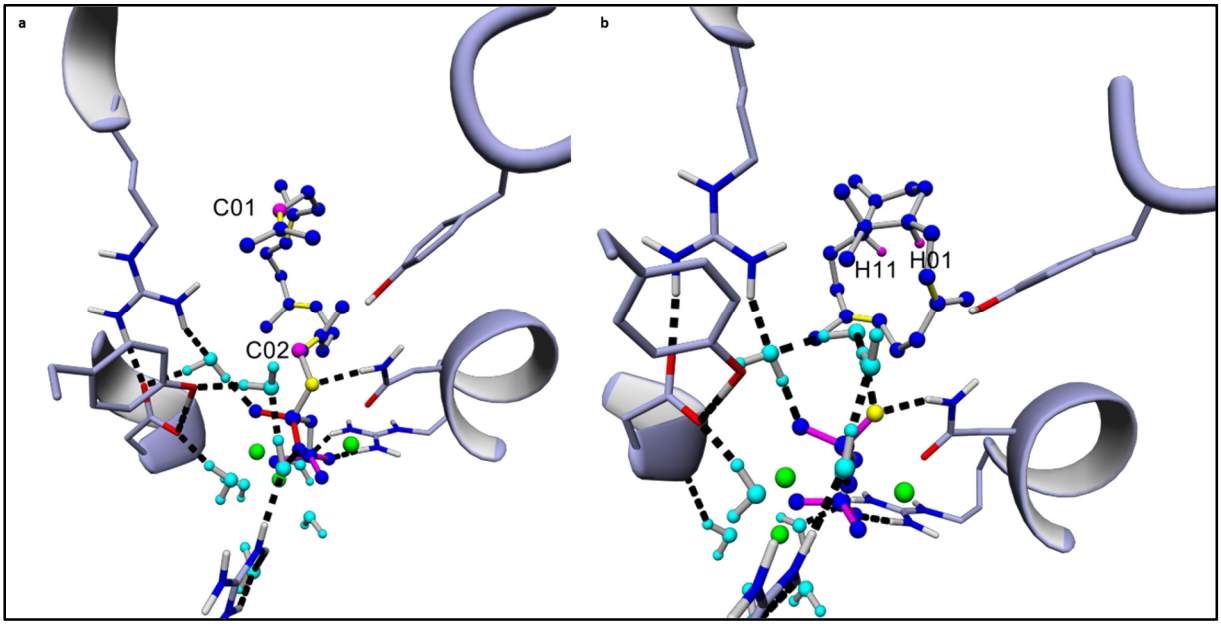


Fig. S3b: TXS harboring the GGPP conformation derived from the crystal structure

(a) TXS containing a **GGPP** in the unproductive conformation derived from the **2F-GGPP** conformation of the open complex crystal structure (PDB: 3P5R). **(b)** Manual bond formation and energy minimization to **cation C`**. This verticillen-12-yl cation` is characterized by an inverted trans-bridgehead stereochemistry of H11 and H1 compared to the verticillen-12-yl cation derived from a productive **GGPP** conformation (**Fig. 4a**), which disturbs the sequence of the hydride shifts. It exhibits moreover an overall inverted positioning in the active and is therefore not productive for the cyclization to **cation F** or -**D**.

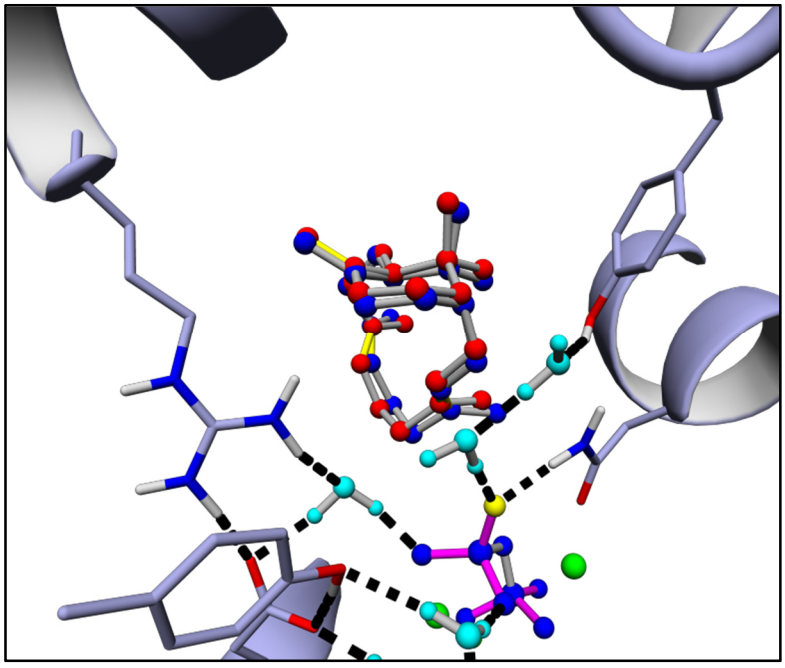


Fig. S3c: Manually built cation C and molecular docking of QM-derived cation C in TXS

Molecular docking of QM-derived **cation C** (red) into TXS and structural superposition with manually built **cation C** (blue).

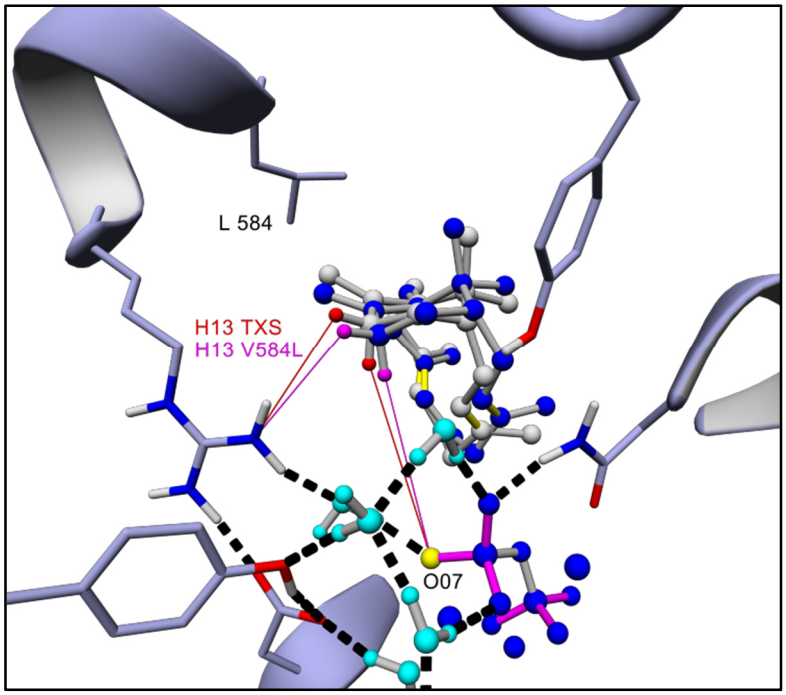


Fig. S3d: Structural superposition of TXS cation C with TXS-V584L cation C

Superposition of TXS-V584L harboring manually built **cation C** (blue) with manually built **cation C** of TXS (gray). C13 hydrogens abstracted during R580-PP_i-assisted deprotonation resulting in formation of verticillia-3,4-7,8-12,13-triene are shown (C13 hydrogens of TXS in red, C13 hydrogens of TXS-V584L in magenta). Distance between H13equatorial (H13e) of TXS-V584L and NH2 of R580 is 3.09 Å (magenta line), distance between H13e of TXS and NH2 of R580 is 3.87 Å (red line). Distance between H13axial (H13a) of TXS-V584L and O07 of PP_i is 4.22 Å, distance between H13a of TXS and O07 of PP_i is 4.83 Å.

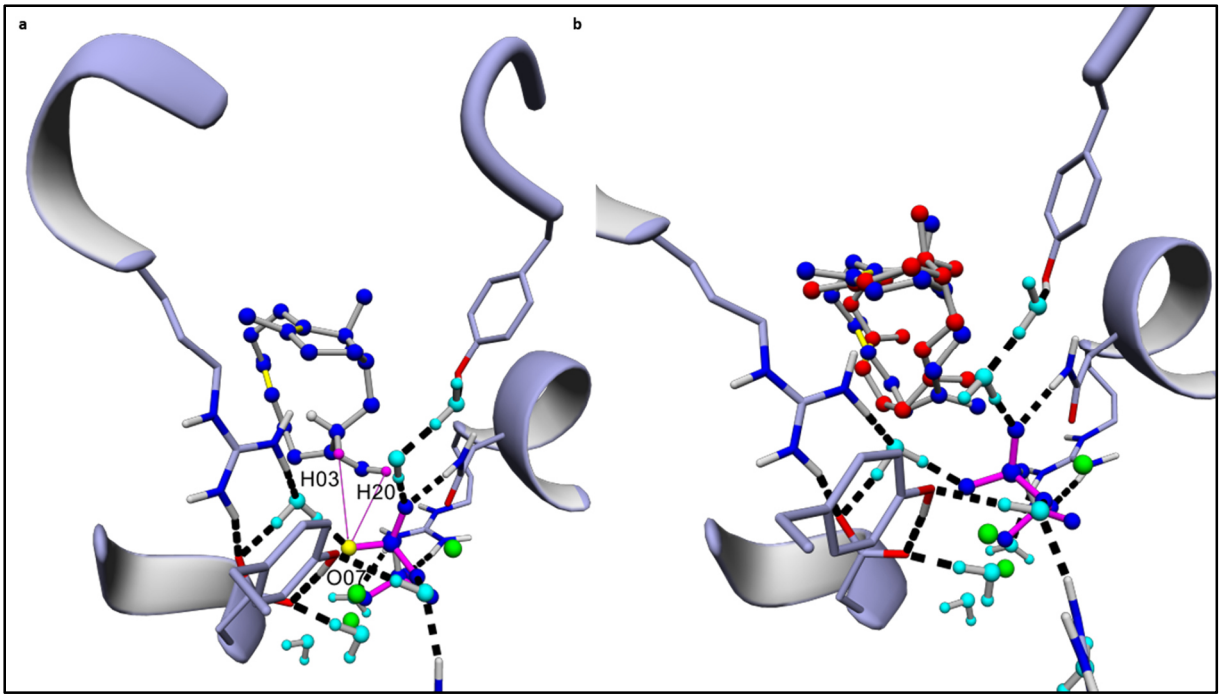


Fig. S3e: Manually built cation F and molecular docking of QM-derived cation F in TXS

(a) TXS containing manually built **cation F**. Distance (magenta line) between H03 and O07 of PP_i is 3.19 Å, distance (magenta line) between H20 (closest) and O07 of PP_i is 3.69 Å. Abstraction of C03 in this **cation F** by the R580-PP_i motif results in formation of **V2**, abstraction of one of the C20 hydrogens results in formation of **V1**. The distance between R580-PP_i and the positively charged C-atom (C03) is 4.06 Å. **(b)** Molecular docking of QM-derived **cation F** (red) into TXS and structural superposition with manually built **cation F** (blue).

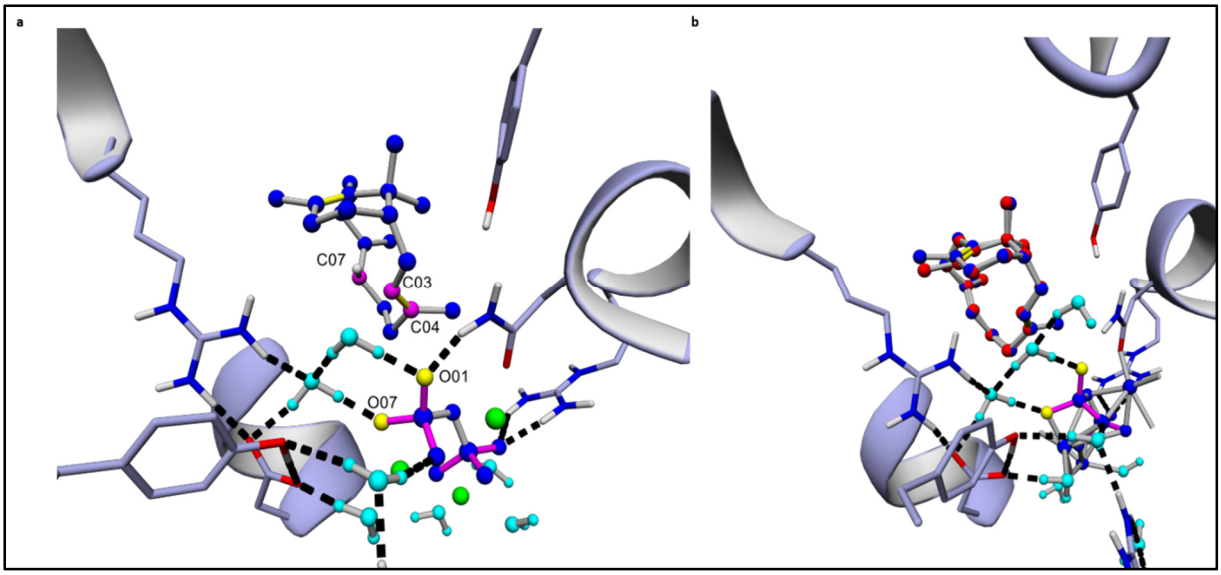


Fig. S3f: Manually built cation D1 and molecular docking of QM-derived cation D1 in TXS

(a) TXS containing manually built **cation D1**. Distance between Δ3,4 and O01 of PP_i is 3.62 Å. **(b)** Molecular docking of QM-derived **cation D1** (red) into TXS and structural superposition with manually built **cation D1** (blue).

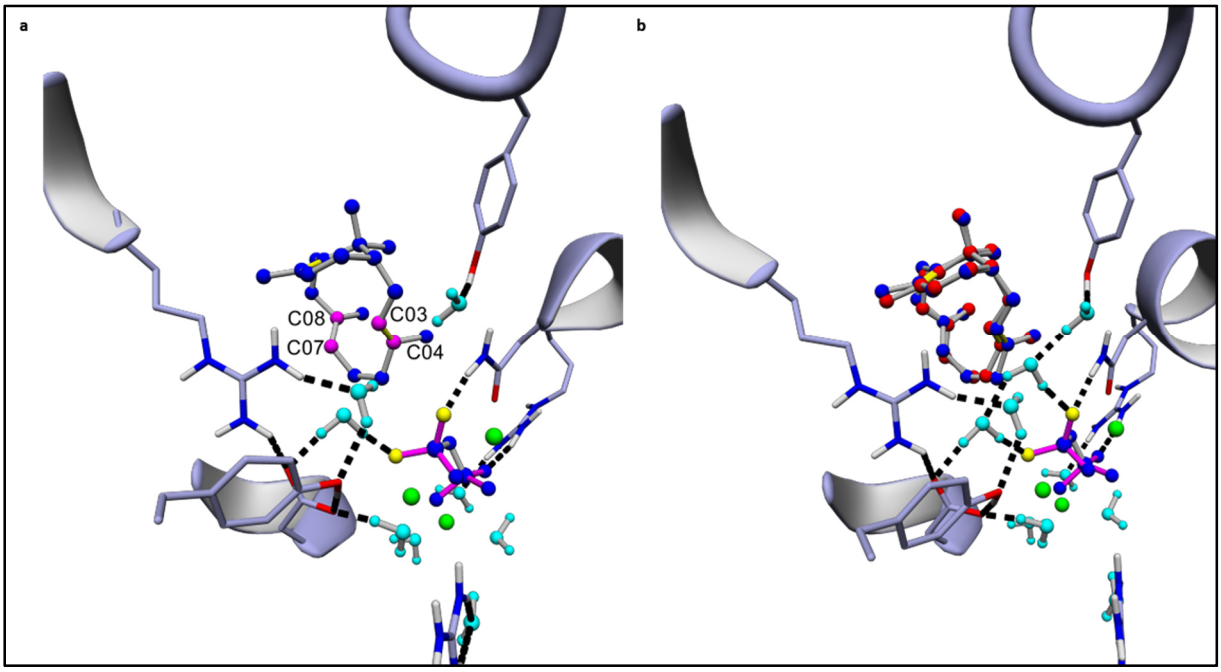


Fig. S3g: Manually built cation D2 and molecular docking of QM-derived cation D2 in TXS

(a) TXS containing manually built **cation D2**. **(b)** Molecular docking of QM-derived **cation D2** (red) into TXS and structural superposition with manually built **cation D2** (blue).

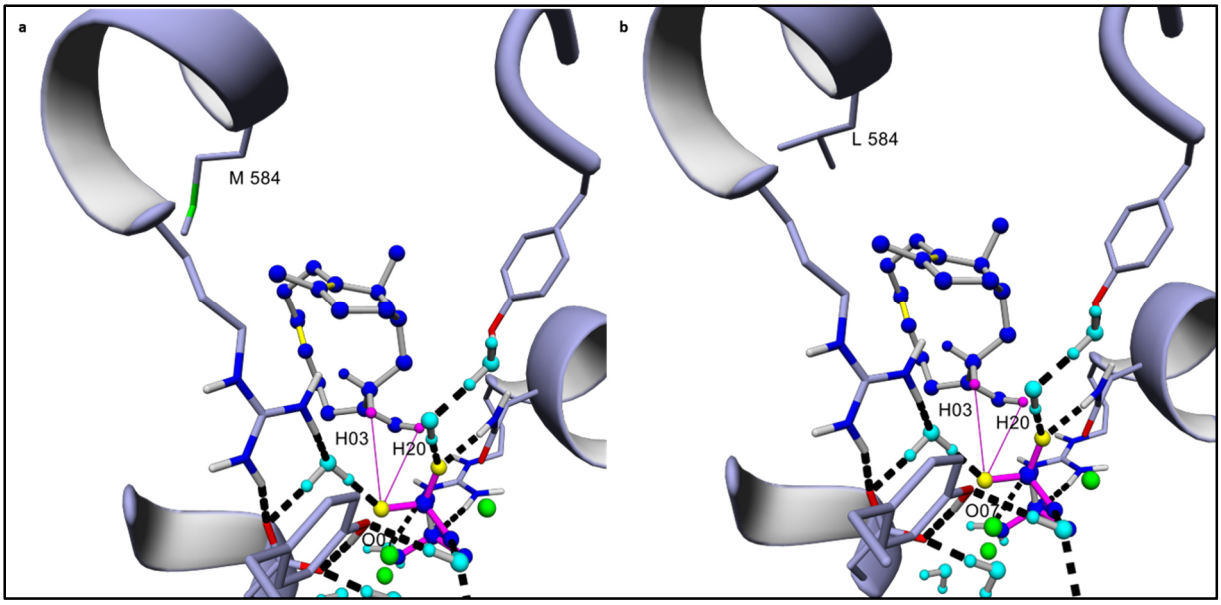


Fig. S3h: TXS-V584M/L harboring cation F

(a) TXS-V584M harboring manually built **cation F**. **(b)** TXS-V584L harboring manually built **cation F**. Magenta lines indicate the distance of H03 and H20 (closest) to the R580-PP_i bi-functional motif.

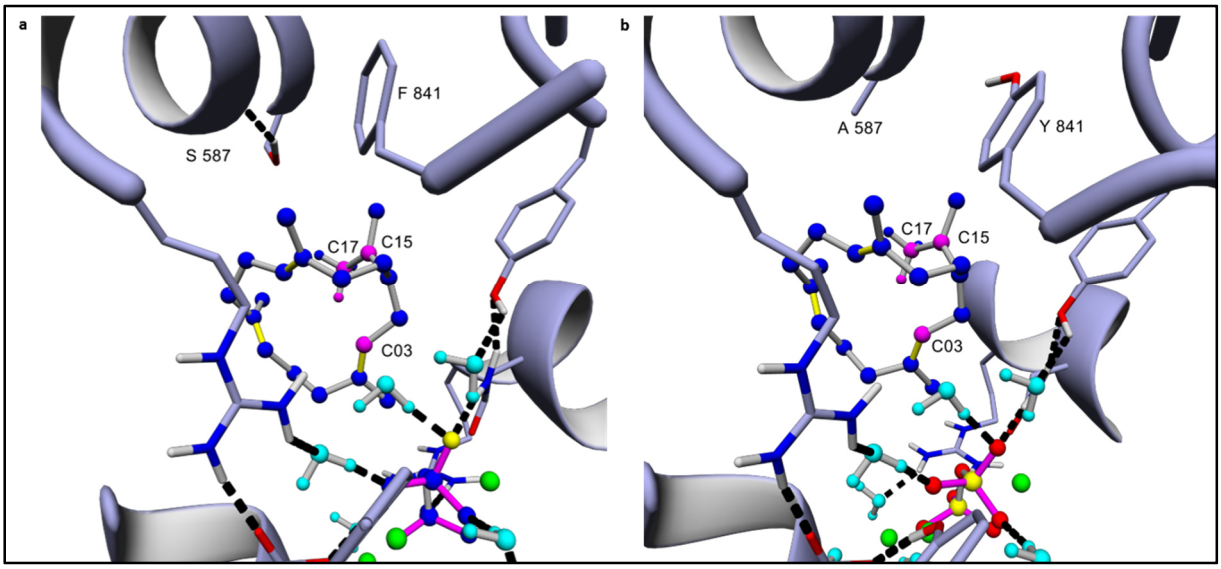


Fig. S3i: TXS-Y841F and TXS-S587A harboring cation B

(a) TXS-Y841F harboring manually built **cation B**. Distance between "liberated" and thus cyclization interfering S587-OH and cationic C15 is 4.162 Å. **(b)** TXS-S587A harboring manually built **cation B**. Distance between "liberated" and thus cyclization interfering Y841-OH and cationic C15 is 4.862 Å. Formation of **CM** in Y841F and S587A mutants can be explained by electrostatic stabilization of the proximal positive charge at C15 by S587 or Y841. An interfered transition of **cation B** to -**C** would enable, governed by R580-PP_i, a not native hydride-shift of

a single C17 hydrogen to Δ3,4 in **cation B**. This, in turn, brings the positive charge into close proximity to R580-PP_i with consequent deprotonation of C3.

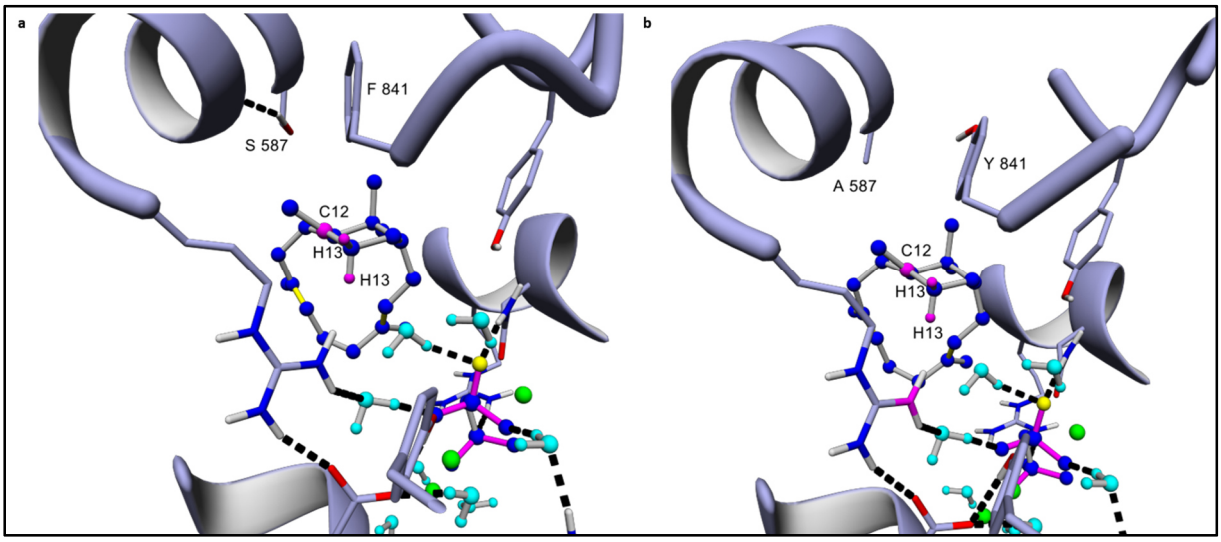


Fig. S3j: TXS-Y841F and TXS-S587A harboring cation C

(a) TXS-Y841F harboring manually built **cation C**. Distance between "liberated" and thus cyclization interfering S587-OH and cationic C12 is 3.859 Å. **(b)** TXS-S587A harboring manually built **cation C**. Distance between "liberated" and thus cyclization interfering Y841-OH and cationic C12 is 5.805 Å. Deprotonation of C13 resulting in formation of **V** seems to be facilitated by electrostatic stabilization of cationic C12 by "liberated" S587 or Y841, respectively. This stabilization leads consequently to an interfered and slowed transition to **cation F**, accompanied by simultaneous cation steering by R580-PP_i towards itself. The bi-functional motif seems then accordingly to be the active site base

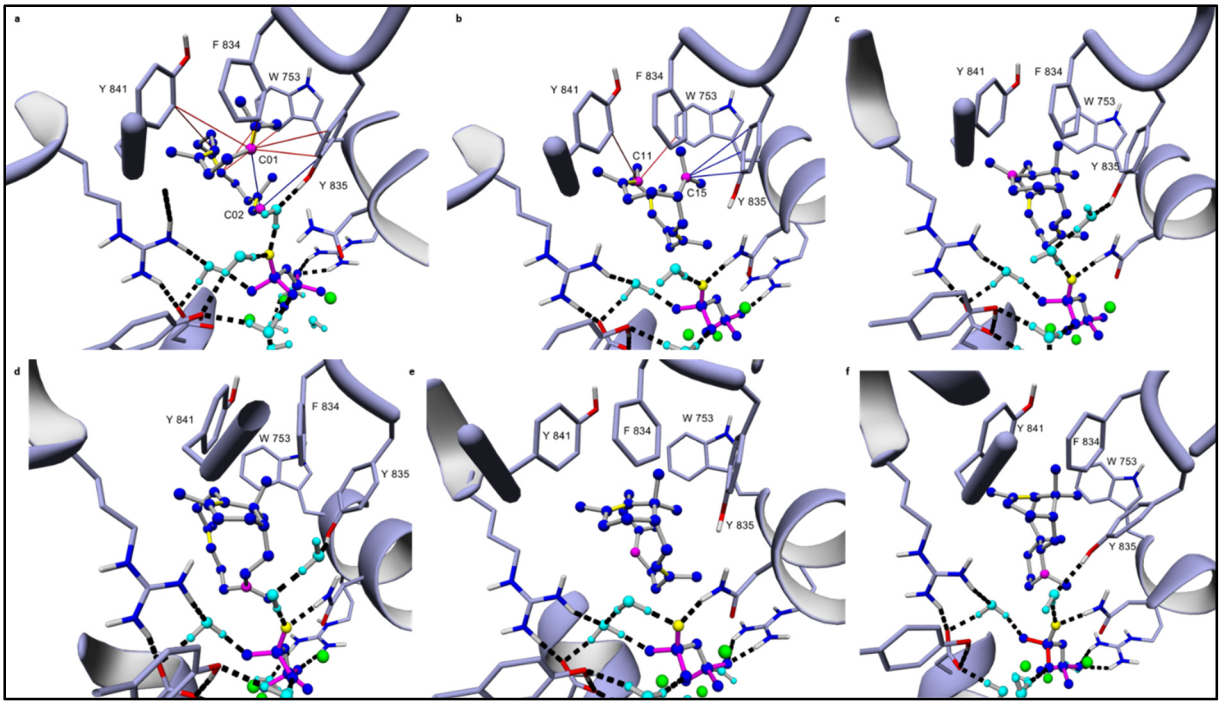


Fig. S4: Amino acid-assisted stabilization of cations

(a-f) Amino acid-assisted π-π (red) and cation-π interactions (blue) acting in stabilization of cations in the active site. **(a)** TXS harboring manually built **cation A**, **(b)** TXS harboring manually built **cation B**, **(c)** TXS harboring manually built **cation C**, **(d)** TXS harboring manually built **cation F**, **(e)** TXS harboring manually built **cation D1** (**-D2** is also characterized by no amino acid-assisted π-π and cation-π interactions) and **(f)** TXS harboring manually built **cation E**.

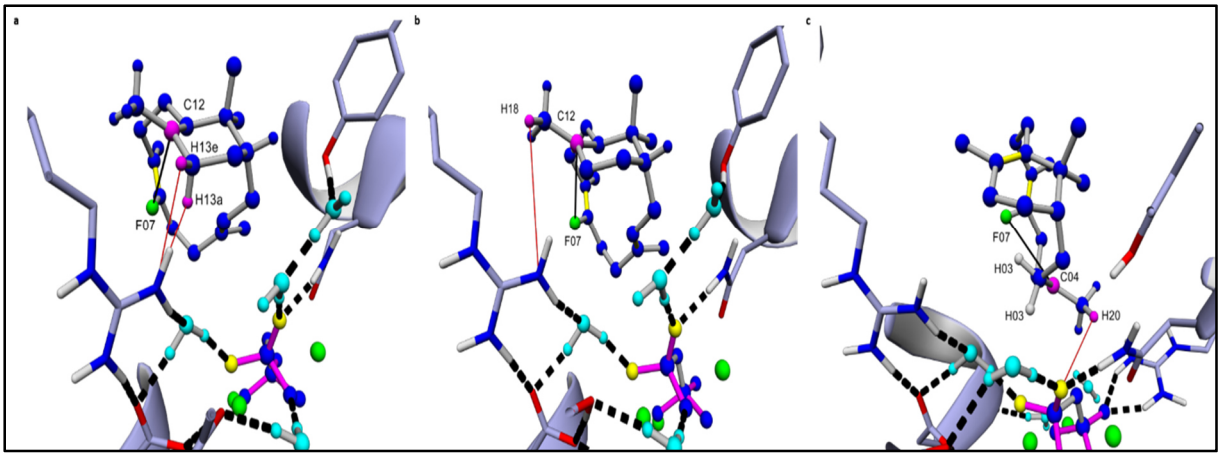


Fig. S5a: TXS•cation C and -F complexes harboring 7-fluoro-GGPP-derived cations C and F

(a-b) TXS harboring the manually built **7-fluoro-GGPP**-derived **cation C**. **(a)** The **7-fluoro**-ligand is shown in green, positive charged C12 and closed H18 presumably abstracted by R580-PP_i, resulting in formation of *exo*-7-fluoroverticillene are shown in magenta. Distance between C12 and F07 is 3.012 Å, distance between closest H18 and R580-PP_i is 4.306 Å. **(b)** The **7-fluoro**-ligand is shown in green, positive charged C12 and H13axial (H13a) as well as H13equatorial (H13e) presumably abstracted by R580-PP_i, resulting in formation of *endo*-7-fluoroverticillene are shown in magenta. Distance between C12 and F07 is 3.012 Å, distance between H13a and R580-PP_i is 3.158 Å, distance between H13e and R580-PP_i is 3.290 Å. **(c)** TXS harboring manually built **cation F** and **7-fluoro-GGPP** as substrate. The **7-fluoro**-ligand is shown in green, positive charged C04 and closest H20 presumably abstracted by R580-PP_i, resulting in formation of 7-fluoroverticillia-4(20),7,11-triene are shown in magenta. H03 atoms are shown in gray. Distance between C04 and F07 is 3.334 Å, distance between closest H20 and R580-PP_i is 3.411 Å.

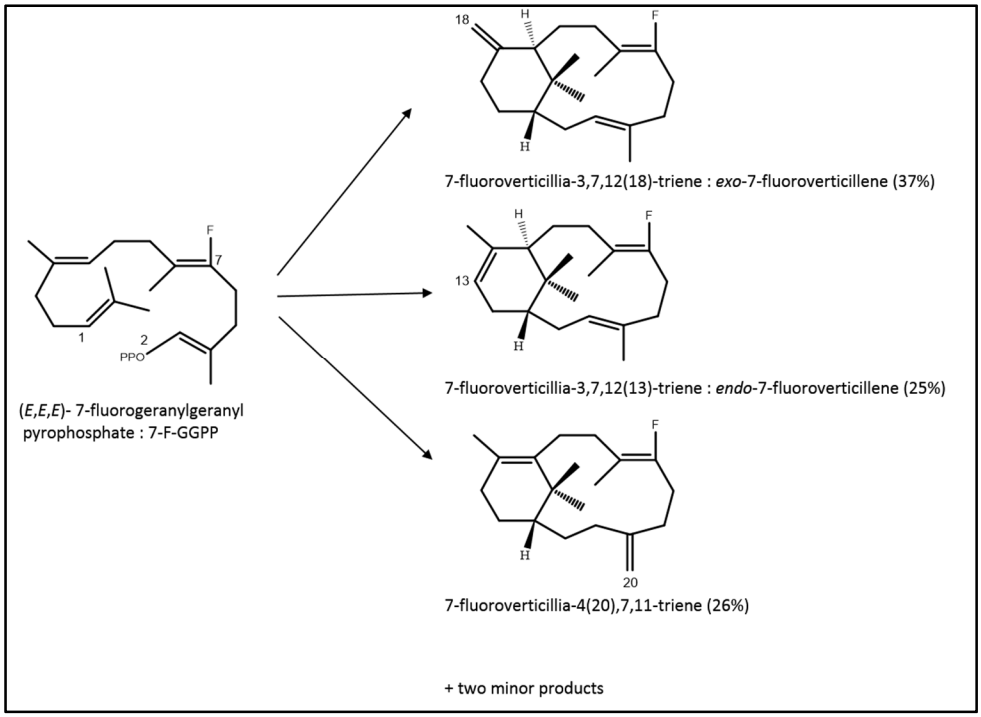


Fig. S5b: Product distribution of TXS incubated with 7-fluoro-GGPP

The TXS•**cation C** and -**F** complexes harboring the manually built **7**-**fluoro**-**GGPP**-derived **cations C** and -**F** (**Fig. S5a-c**) are able to give an indication for the product distributions observed during incubation with **7-fluoro-GGPP**.[3] These experiments conducted by Jin and coworkers, show a 37% outcome of *exo*-7-fluoroverticillene. Our models (**Fig. S5a-b**) show closer distances between H13a/e, abstracted in formation of *endo*-7-fluoroverticillene, and R580-PP_i, compared to the distance between closest H18 (to R580-PP_i), abstracted in formation of *exo*-7-fluoroverticillene and R580-PP_i. This would suggest a favored *endo*-7-fluoroverticillene formation. However, the models further indicate that the **7-fluoro** ligand is able to counteract H13a/e abstraction by a transient electrostatic attraction of H13a/e. This attraction is reduced in respect to closest H18, resulting in the observed product distribution derived from deprotonation of **cation C**. The distance between F07 and H13a is 3.247 Å and the distance between F07 and H13e is 4.358 Å. In contrast, distance between closest H18 and F07 is 4.859 Å. The electrostatic attraction of F07 with H03 protons in **cation F** (**Fig. S5c**) is moreover able to explain the absence of a 7-fluoroverticillia-3,7,11(12)-triene as another major product in their observed product distribution, despite close proximity of H03 protons to R580-PP_i, corresponding to the observed formation of **V2** in **Fig. S3e**.

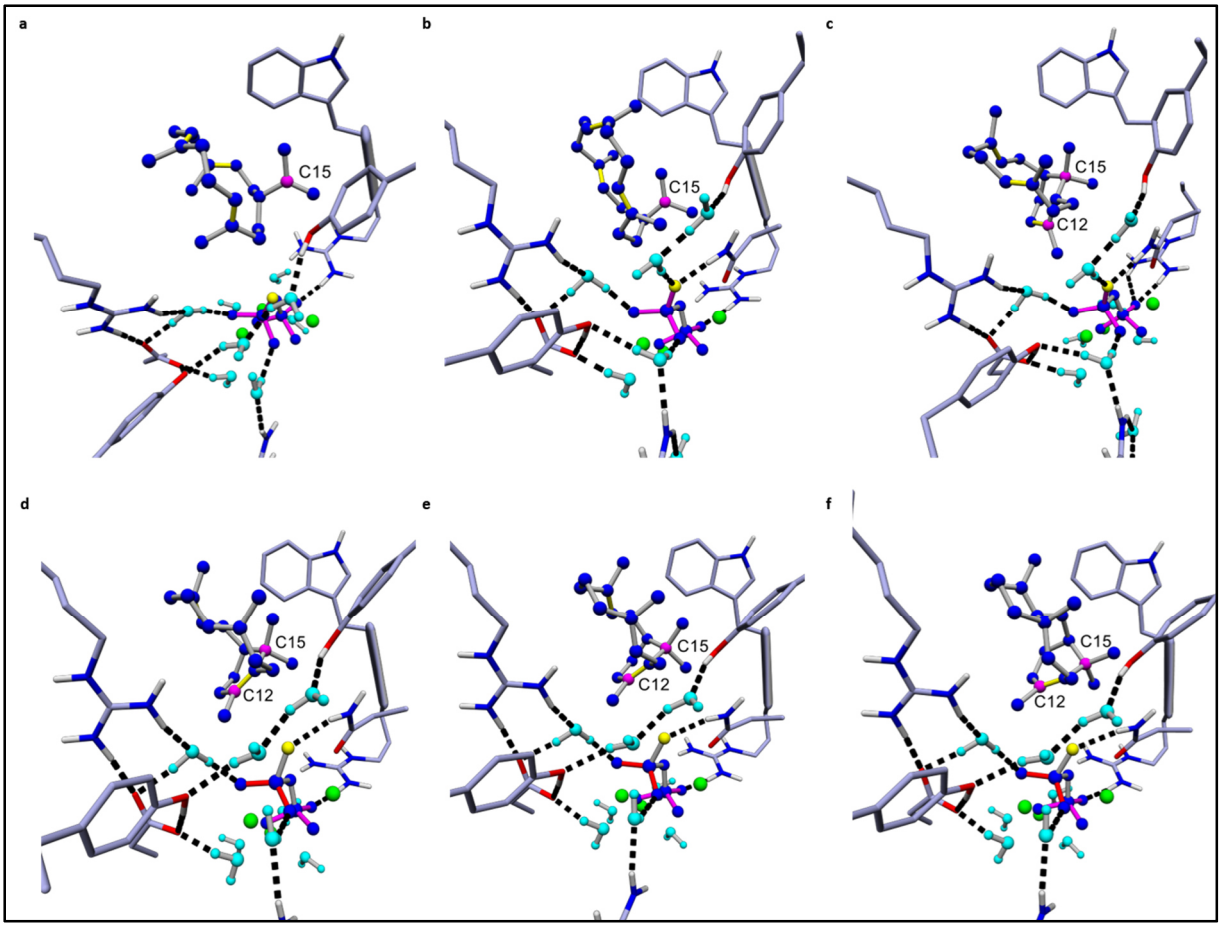


Fig. S6: Representation of unproductive docking cluster

Representation of unproductive docking cluster in accordance to **Table S3**. **(a)** Representation of the conformational cluster of **cation B**, that is structurally not derived from productive **cation A** due to its localization in the active site (**Fig. 2c, d** and **Fig. S3a**), obtained upon docking of QM-derived **cation B** into its corresponding transient enzyme complex. **(b)** Representation of the conformational cluster of **TS B-C**, that is, due to its localization it the active site, structurally not derived from productive **cation A**, obtained upon docking of QM-derived **TS B-C** into its corresponding transient enzyme complex. **(c)** Representation of the conformational cluster of **cation C**, that is, due to its localization it the active site, structurally not derived from productive **TS B-C**, obtained upon docking of QM-derived **cation C** into its corresponding transient enzyme complex. **(d)** Representation of the conformational cluster of **cation D1**, that is, due to its localization it the active site, structurally not derived from productive **cation C**, obtained upon docking of QM-derived **cation D1** into its corresponding transient enzyme complex. **(e)** Representation of the conformational cluster of **cation D2**, that is, due to its localization it the active site, structurally not derived from productive **cation D1**, obtained upon docking of QM-derived **cation D2** into its corresponding transient enzyme complex. **(f)** Representation of the conformational cluster of **cation E**, that is, due to its localization it the active site, structurally not derived from productive **cation D2**, obtained upon docking of QM-derived **cation E** into its corresponding transient enzyme complex.

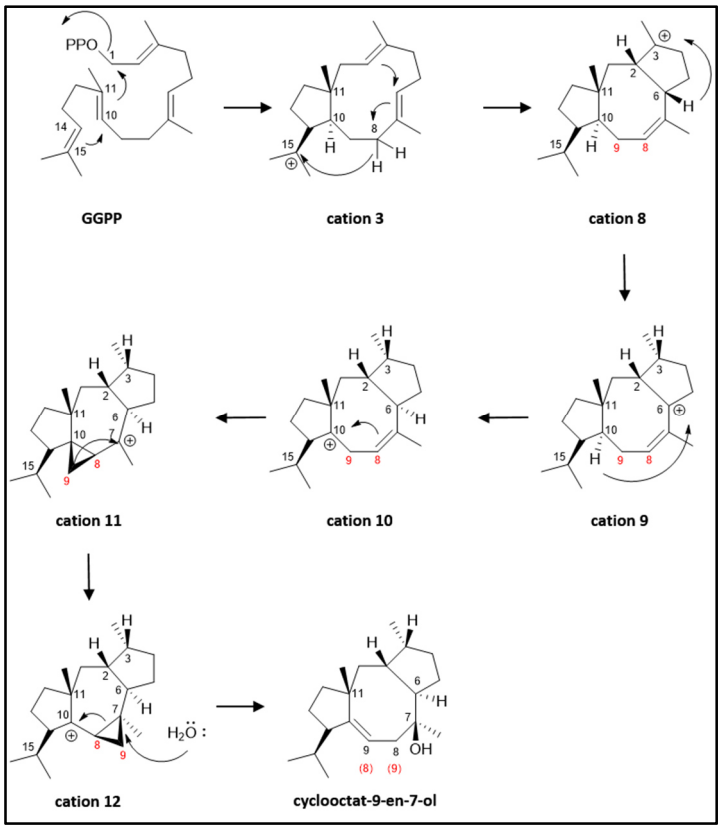


Fig. S7a: Proposed mechanism for the formation of cyclooctat-9-en-7-ol

Proposed mechanism for the formation of cyclooctat-9-en-7-ol from **GGPP** by CotB2 derived by NMR-spectroscopic studies conducted by Meguro et al.[4]

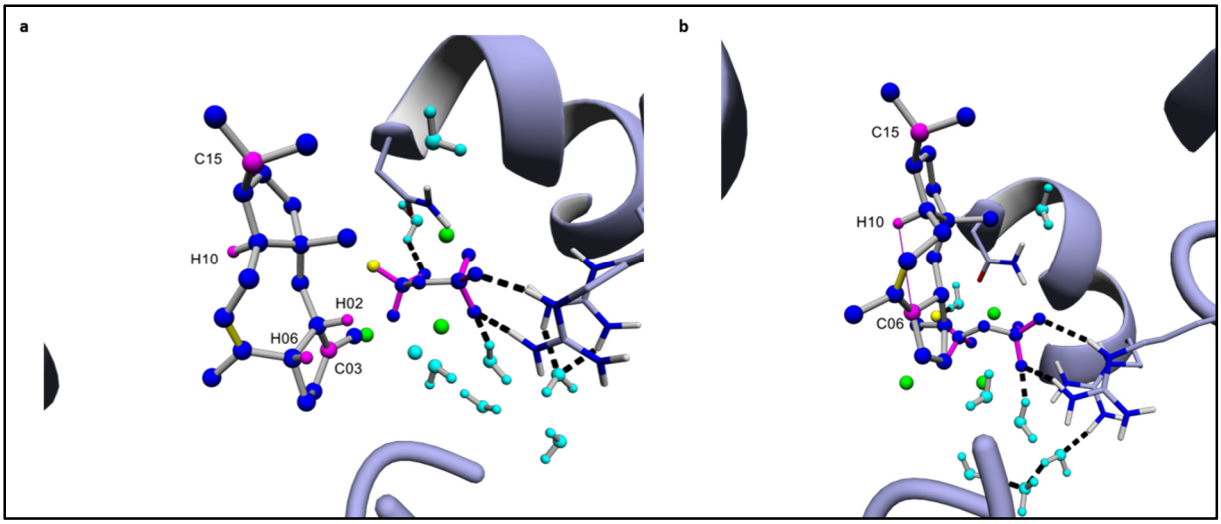
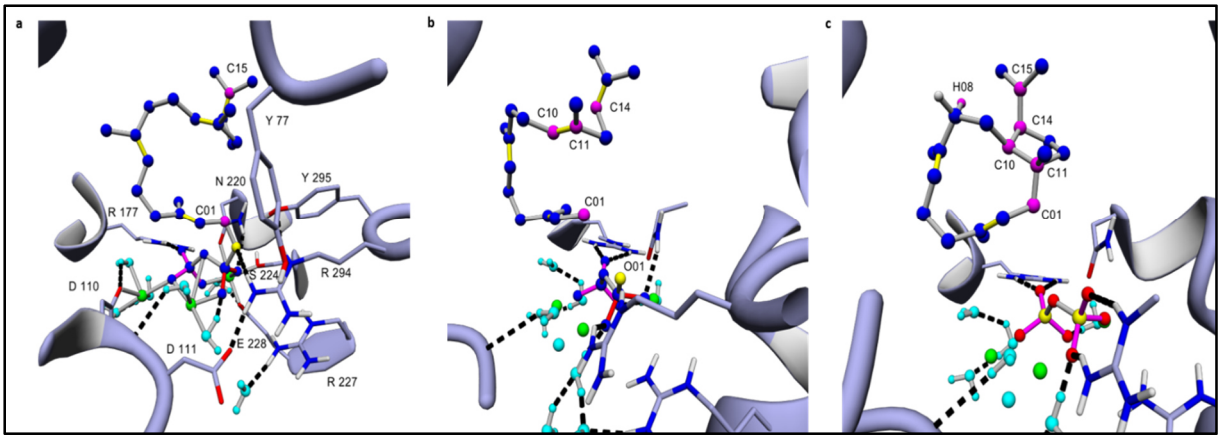


Fig. S7b: CotB2 harboring the proposed cation 8 and -9

(a) CotB2 harboring the manually built proposed **cation 8** (**Fig. S7a**), derived from the CotB2•**cation B** complex (**Fig. S22a**). Note that the here established **cation B** corresponds to the proposed **cation 3.**[5] **(b)** CotB2 harboring the manually built proposed **cation 9** (**Fig. S7a**), derived from the CotB2•**cation 8** complex (**Fig. S7b**).[5] Manual bonding followed by energy minimization that simulate the proposed subsequent concerted steps in transition of **cation B**->-**8**, the 1,5-hydride shift (H8->C15), the double bond rearrangement of Δ6,7->Δ7,8 as well as the attack of Δ2,3 on C06 in the proposed **cation 3** lead to the proposed **cation 8** in the active site of CotB2 (**Fig. S7a**). However, the proposed subsequent hydride shifts of C06->cationic C03 (**cation8**) and the second 1,5-hydride shift of H10->cationic H06 (**cation 9** in **Fig. S7a**) seems rather not plausible in these respective complexes as the cationic empty p-orbitals in these cations do not point towards H06 or H10, respectively. This would thus lead to a cis-H02-H06-stereochemistry instead of the experimentally verified trans-stereochemistry (**Fig. S7a**).

Therefore, we propose the existence of an additional intermediate, **cation C**, with cationic C06 (**Fig. S7d**, **Scheme 2d**). We further propose that the 1,5-hydride shift (H8->C15) and the double bond rearrangement (Δ6,7->Δ7,8) take place during the transition of **cation B**->-**C** (**Fig. S7c**, **Scheme1c**). This, in turn, is followed by the attack of Δ2,3 on C06 during the transition of **cation C**->-**D**, accompanied by the simultaneous 1,4- (H03->H06) and the second 1,5-hydride shift (H10->H06) (**Fig. S7d**, **Scheme 2d**). Consequently, the CotB2 cyclization cascade proceeds *via* **GGPP**->**cation A**->-**B**->-**C**->-**D**->-**E**->-**F**->**cyclooctat-9-en-7-ol**.



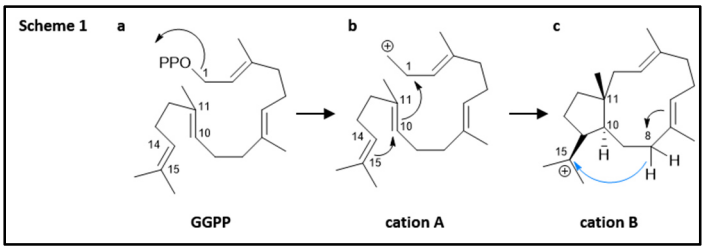
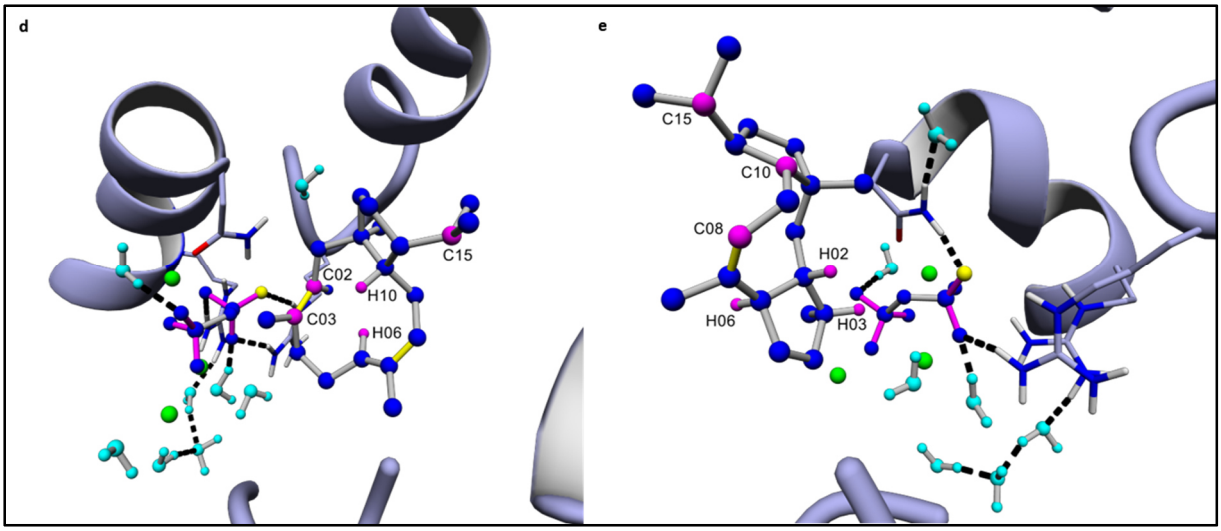


Fig. S7c: CotB2 harboring GGPP and cations A and B

(a) Closed conformation of CotB2, hereafter designated as CotB2, harboring productive **GGPP** (**scheme 1a**) (C01 and C15 of **GGPP** are shown in magenta, magnesium ions are shown in green, O01 of PP_i is shown in yellow, remaining bulk water solvent is shown in cyan and coordinating bonds are shown as gray lines). **(b)** CotB2 harboring manually built **cation A** (**scheme 1b**). Carbon atoms involved in cyclization to **cation B** are shown in magenta. Note, that numbering of cations herein differ from the shown experimental observed catalytic mechanism (**Fig. S7a**). **(c)** CotB2 harboring manually built **cation B** (**scheme 1c**). This complex is able to explain the structural basis for the experimentally observed 1,5-hydride shift of C8 to C15 (**blue line in scheme 1**), as distance between H8 (magenta) and cationic C15 (magenta) is 2.469Å.



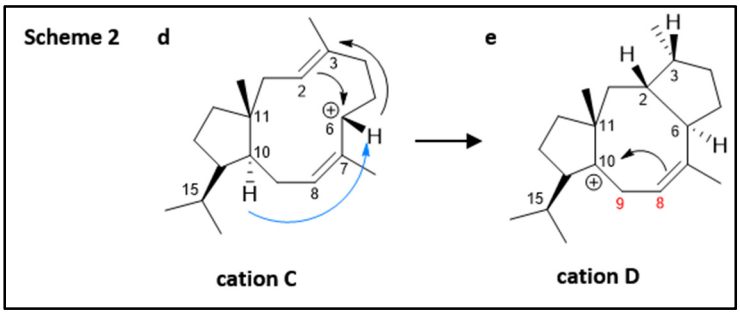
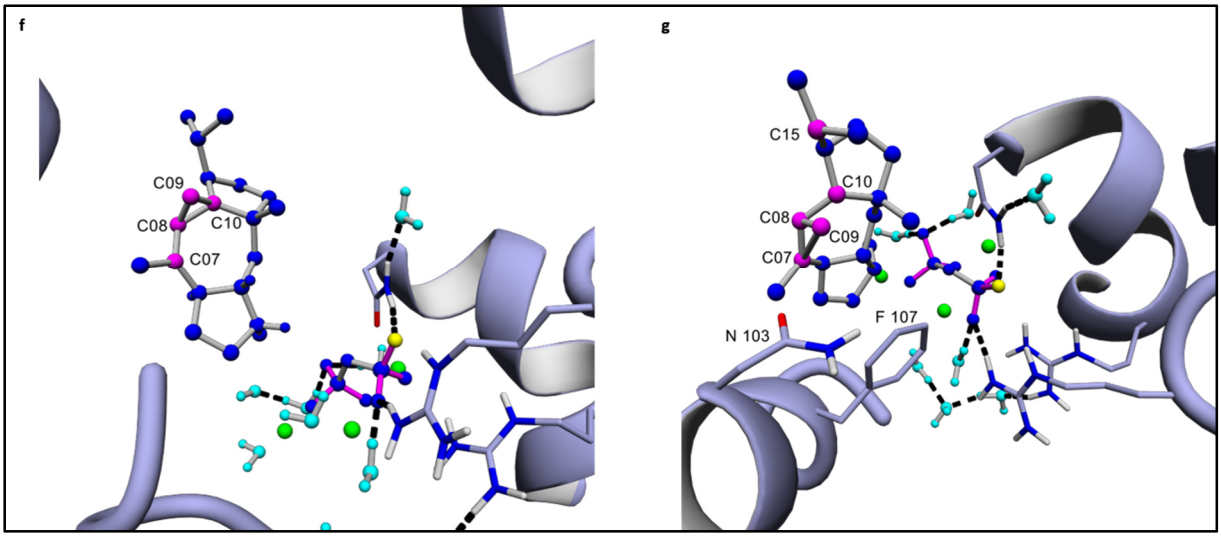


Fig. S7d: CotB2 harboring cations C and D

(d) CotB2 harboring manually built **cation C** (**scheme 2d**), carbon atoms involved in cyclization to **cation D** are shown in magenta. The complex model shows the Δ2,3 double bond directly pointing towards C06 and H06 and cationic C06 directly point towards H10. This indicates that the attack of Δ2,3 on C06 is synchronized with the 1,4-hydride shift H06->C03 while the second 1,5-hydride shift H10->C06 is asynchronously concerted to this event. This leads to the experimentally observed trans-stereochemistry of H02/H06 (**Fig. S7a**). The existence of a delocalized 3-center 2-electron bonding array between C03, C02 and H06 would be able to explain the synchronized first step. **(e)** CotB2 harboring manually built **cation D** (**scheme 2e**). This complex shows the Δ7,8 π-orbital overlapping with the empty p-orbital of cation C10 leading to the experimentally observed carbon rearrangement of C8 and C9, as shown in **Fig. S7a**, **Fig. S7e** and **scheme 3f-g**.



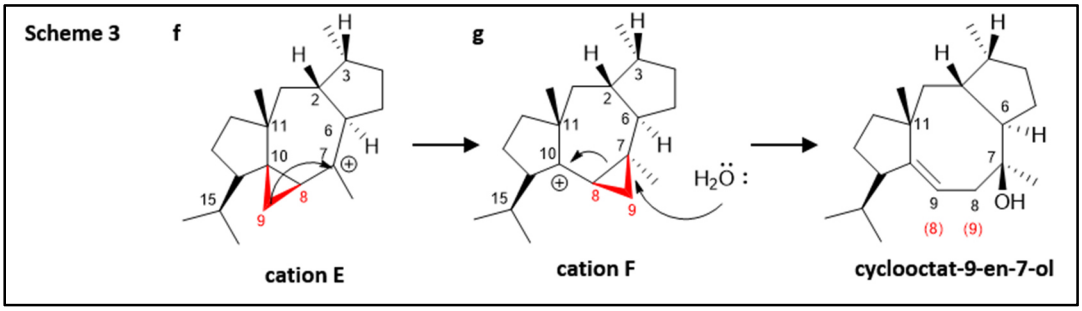


Fig. S7e: CotB2 harborig cations E and F

(a) CotB2 harboring manually built **cation E** (**scheme 3f**), carbon atoms involved in the carbon rearrangement of C8 and C9 are shown in magenta. **(b)** CotB2 harboring manually built **cation F** (**scheme 2g**). Amino acid residues in close proximity to C07 that is attacked by an active site bound water molecule, leading to formation of **cyclooctat-9-en-7-ol** (**scheme 3**) are shown.

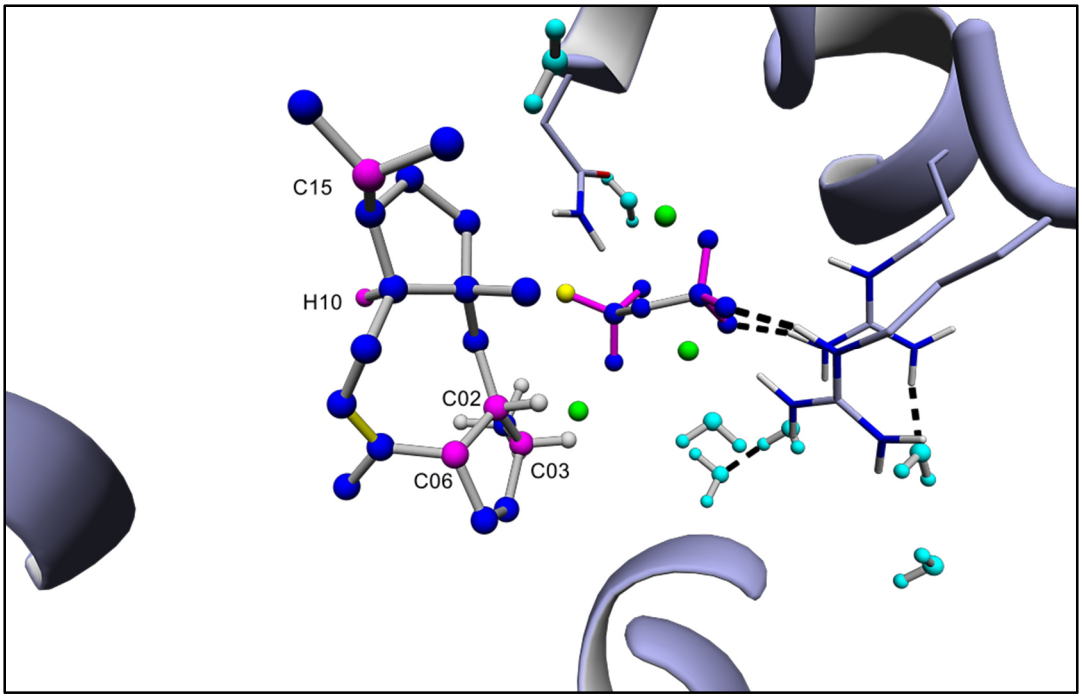


Fig. S7f: CotB2 harboring the proposed cation E upon QM gas phase calculations [6]

During the finalization of this manuscript, a QM gas phase study of the CotB2 mechanism based on the NMR-spectroscopic studies was published [6] (Note that the numbering there is different from the numbering in here and also different to the numbering in Meguro et al). The authors corroborate our hypothesis that additional carbocationic intermediates must exist, in addition to the intermediates shown in **Fig. S7a**. Especially during the **cation B**-> -**C** transition and **cation C**-> -**D** transition that included the H10->H06 hydride shift Hong and Tantillo [6] propose a differing cyclization cascade, however in the absence of the protein environment. The QM structure of cation E shown in their study demonstrated a distance of 3.81 Å between H10 and cationic H06 within our CotB2 complex after following their proposed cyclization mechanism (**Fig. S7f**). A structural rearrangement of **cation E** is therefore needed for their proposed **cation E**-> -**F** transition that our model is not able to show. We cannot rule out their shown pathway, thus, in line with them, additional stereo labelling experiments or QM/MM studies including the protein environment have to be conducted to specify the additional carbocation intermediates.

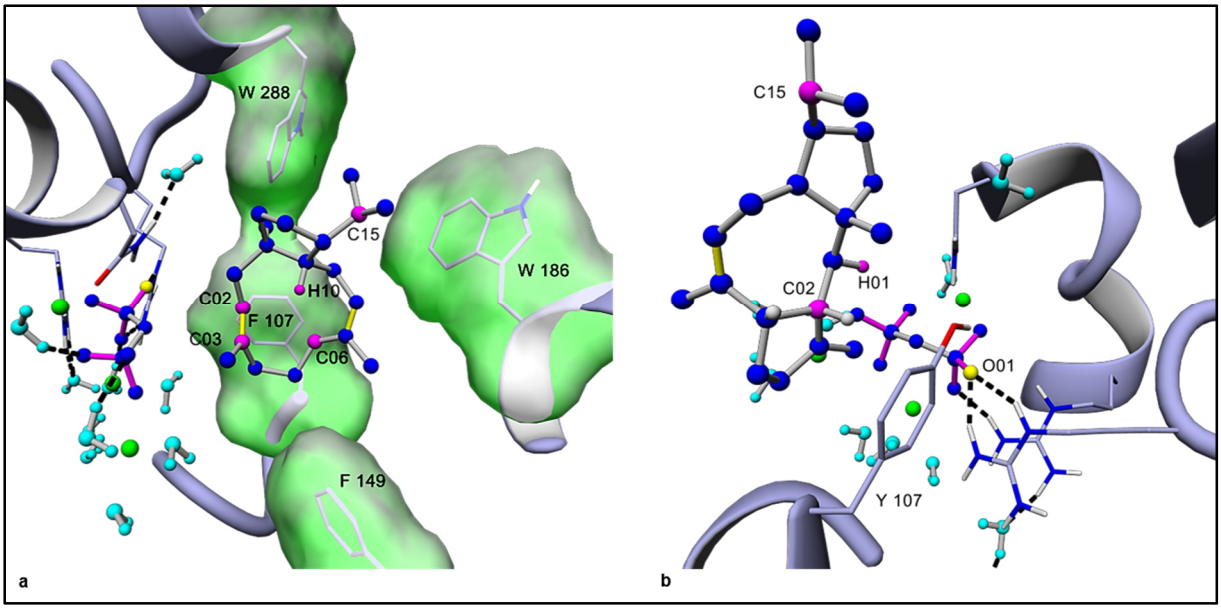


Fig. S7g: Proposed molecular basis for the recently generated CotB2 mutants 1 [7]

(a) Closed conformation of CotB2 harboring cation C (**Fig. S7d, scheme 2**). Formation of **cation C** seems to be forced by the space spanned by aromatic residues (F107, F149, W186 and W288). Transition to **cation D** (**Fig. S7d, scheme 2**) involves a H10-> H06 hydride shift based on the proximity of C06 and C10. Substitution of F149 to leucine allows more sterical freedom upon formation of **cation C** that prevents the H10-> H06 hydride shift resulting in formation of cyclooctat-7-en-3-ol instead of cyclooctat-9-en-7-ol based on final hydroxylation at C03 in **cation D**.[7] **(b)** Closed conformation of CotB2-F107Y harboring cation C. This complex suggests that the newly introduced hydroxyl group of Y107 interacts with O01 of the R294-PP_i bi-functional motif upon formation of **cation C**. Deprotonation of the hydroxyl group of Y107 by R294-PP_i allows direct proton abstraction of H01, due its proximal positioning, accompanied with the proposed H02->H03 hydride shift[7] resulting in the formation of cyclooctat-1,7-diene.

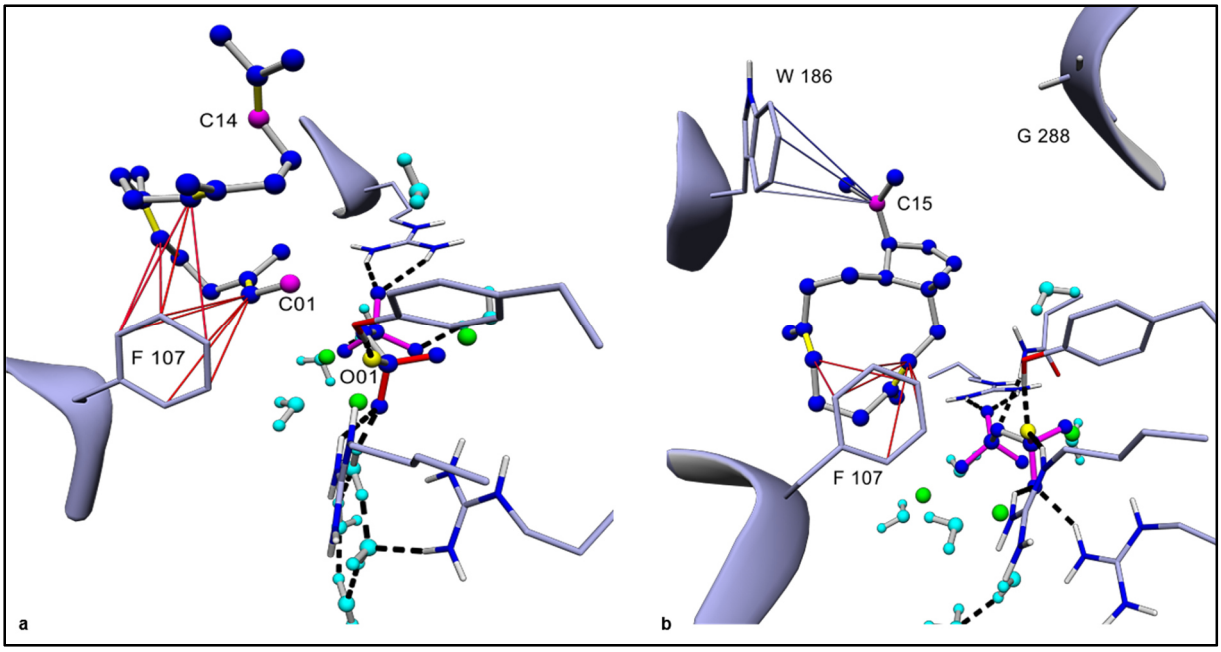


Fig. S7h: Proposed molecular basis for the recently generated CotB2 mutants 2 [7,8]

(a) Closed conformation of CotB2 harboring **cation A** (**Fig. S7c, scheme 1)**. The model demonstrates the importance of F107 in π-π-stabilization (red lines) of the first cationic intermediate. Substitution of F107 to alanine abolishes this stabilization resulting in a higher degree of initial **GGPP** misfolding and cation tumbling that leads to the formation of the monocyclic cembrene A.[7,8] **(b)** Closed conformation of CotB2 harboring **cation B** (**Fig. S7c, scheme 1**). The model indicates that the substitution of W288 to glycine promotes the formation of strong cation-π interactions between cationic C15 and W186 in **cation B** in constrast to the wild type. This stabilization is able to explain the formation of the bicyclic dolabella-3,7,18-triene detected upon mutagenesis.[8]

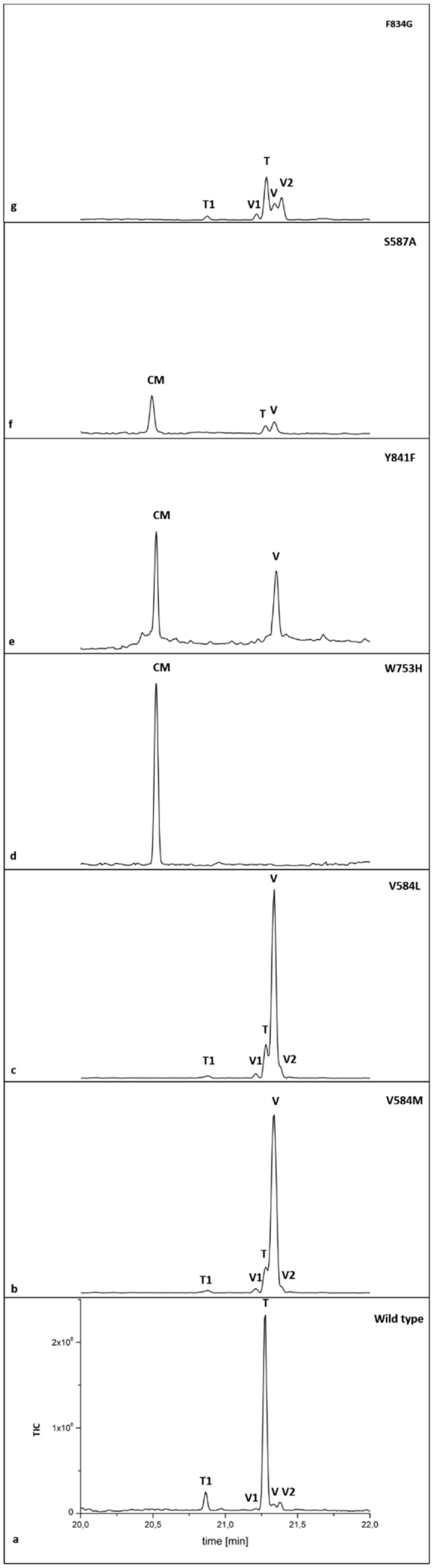

F834G
T
V2
V
T1
V1
g
S587A
CM
T
V
f
Y841F
CM
V
e
W753H
CM
d
V584L
V
T
T1
V1
V2
c
V584M
V
T
T1
V1
V2
b
Wild type
T
2x10⁸
1x10⁸
TIC
T1
V1
V
V2
0
20,0
20,5
21,0
21,5
22,0
time [min]
a

Fig. S8a: GC-chromatograms of wild type and mutant TXS variants

GC-chromatograms of (**a**) Wild type TXS, (**b**) TXS-V584M, (**c**) TXS**-**V584L, (**d**) TXS-W753H, (**e**) TXS-Y841F, (**f**) TXS-S587A and (**g**) TXS-F834G. **T**: taxa-4,11-diene, **T1**: taxa-4(20),11-diene, **CM**: cembrene A, **V**: verticilla-3,7,12(13)-triene, **V2**: verticilla-3,7,11(12)-triene, **V1**: verticilla-4(20),7,11-triene (% activity in comparison to WT corresponding to **Table 1 and Table S4**).

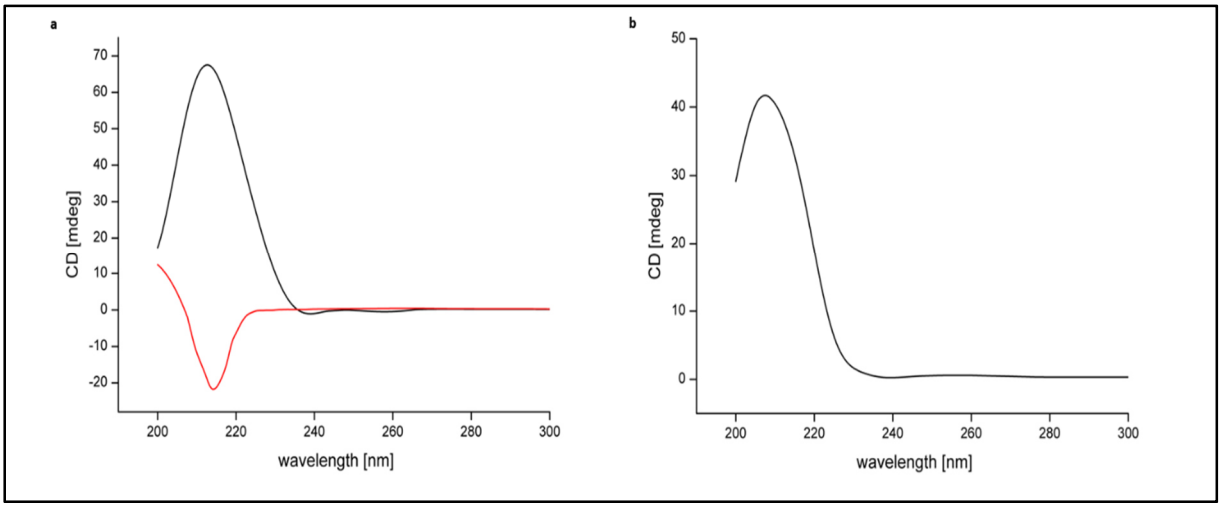


Fig. S8b: CD-spectra of CM, (+)-cembrene and V

CD-spectra of **(a)** (+)-cembrene (black line) and **CM** (red line), **(b)** CD-spectrum of **V** (black line).

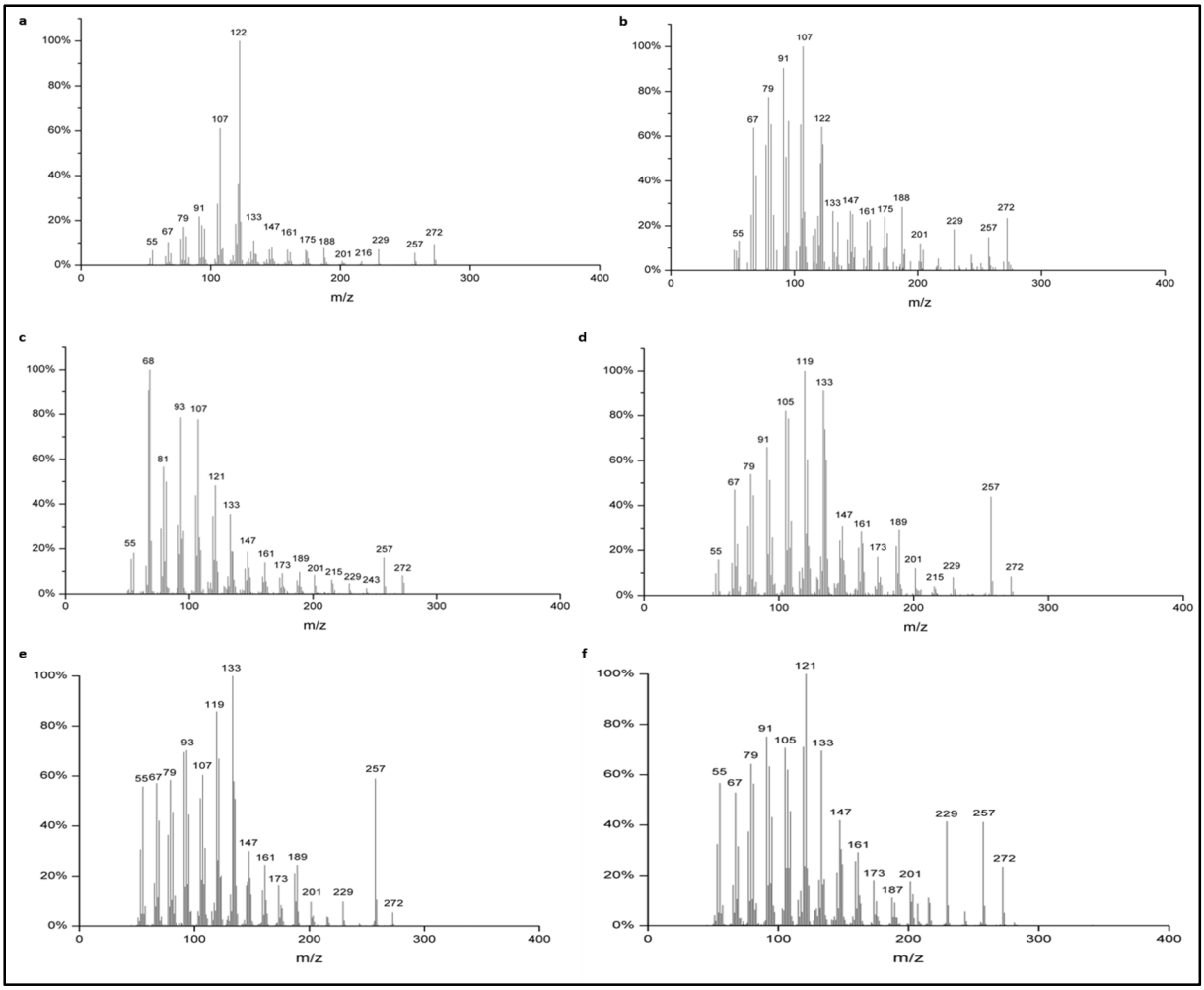


Fig. S8c: MS-Spectra of deprotonated intermediates CM, V, V1, V2, T and T1

Mass spectra of deprotonated intermediates according to **Fig. 1**. **(a) T**: taxa-4,11-diene (Retention time (RT) in gas chromatogram (**Fig. S8a**): 21,28 min), **(b) T1**: taxa-4(20),11-diene (RT: 20,87 min), **(c) CM**: cembrene A (RT: 20,56 min), **(d) V**: verticilla-3,7,12(13)-triene (RT: 21,33 min), **(e) V1**: verticilla-4(20),7,11-triene (RT: 21,20 min), **(f) V2**: verticilla-3,7,11(12)-triene (RT: 21,37).

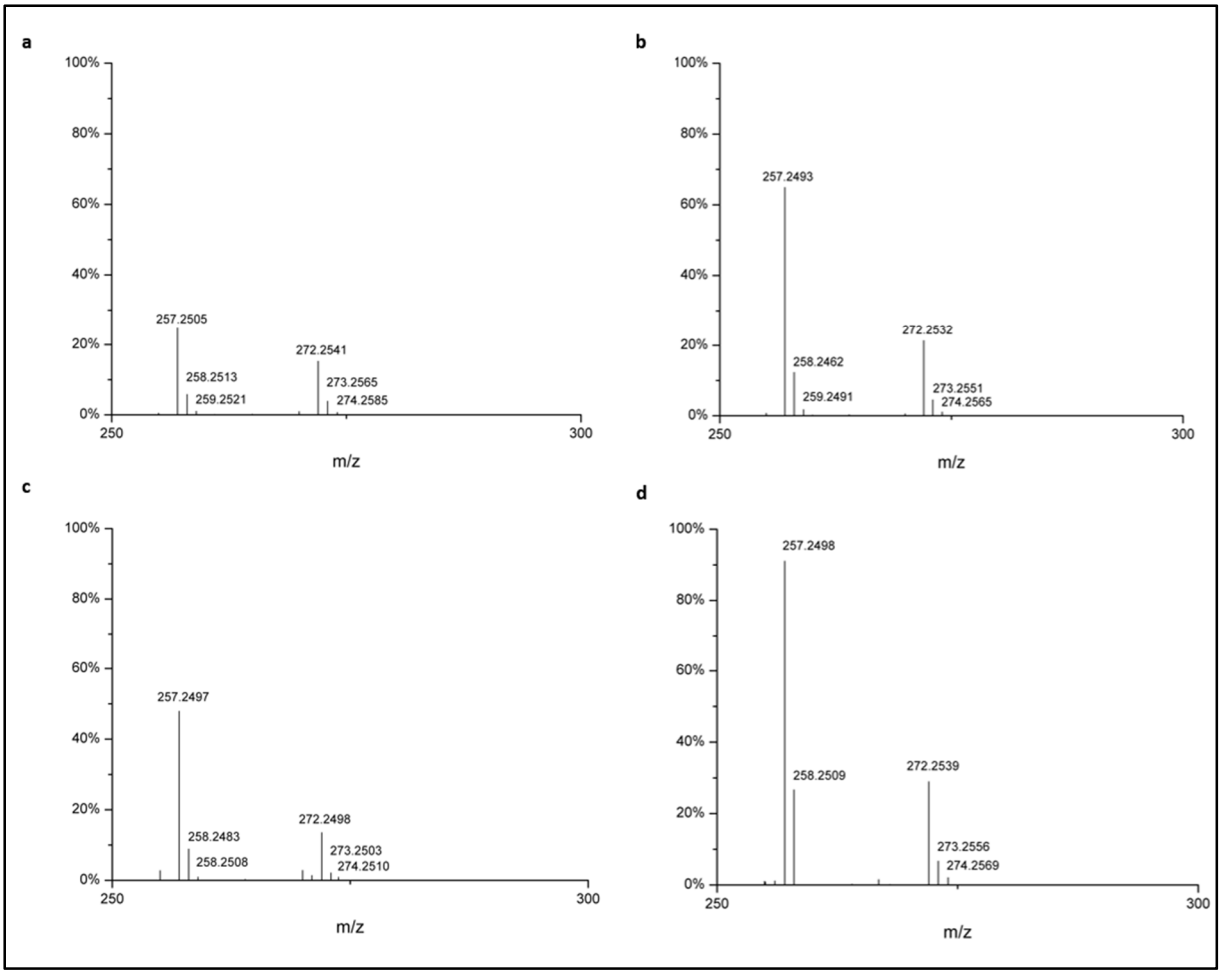


Fig. S8d: HR-MS-Spectra of deprotonated intermediates CM, V, V1 and V2

High resolution mass spectra of deprotonated intermediates according to **Fig. 1. (a) CM**: cembrene A, (m/z $C_{20}H_{32}$ calculated: 272.2504), **(b) V**: verticilla-3,7,12(13)-triene, (m/z $C_{20}H_{32}$ calculated: 272.2504), **(c) V2**: verticilla-3,7,11(12)-triene, (m/z $C_{20}H_{32}$ calculated: 272.2504), **(d) V1**: verticilla-4(20),7,11-triene, (m/z $C_{20}H_{32}$ calculated: 272.2504).

NMR spectral data for (-)-(*R)*-cembrene A			
W753H			(R)-cembrene A [9]
#	***δc* (ppm)**	***δ*$_H$ (ppm), *J* (Hz)**	***δc* (ppm)**
1	46.1	2.04 (m)	45.98
2	32.55	1.99 (m)	32.43
3	121.9	5.06 (t, J = 6.1)	121.87
4	134.0	n/a	133.91
5	39.1	2.03 (m)	38.94
6	25.0	2.17 (m)	24.89
7	124.1	5.19 (t, J = 7.2)	124.07
8	134.9	n/a	134.79
9	39.5	2.07 (m)	39.41
10	23.8	2.10 (m)	23.76
11	126	4.98 (t, J = 6.2)	125.90
12	133.5	n/a	133.43
13	34.0	2.13 (m)	33.99
14	28.2	1.96 (m)	28.22
15	149.4	n/a	149.29
16	19.4	1.66 (s)	19.31
17	110.2	4.65 (s), 4.71 (s)	110.10
18	18.1	1.56 (s)	17.99
19	15.4	1.59 (s)	15.25
20	15.6	1.57 (s)	15.48

Fig. S8e: NMR spectral data of CM

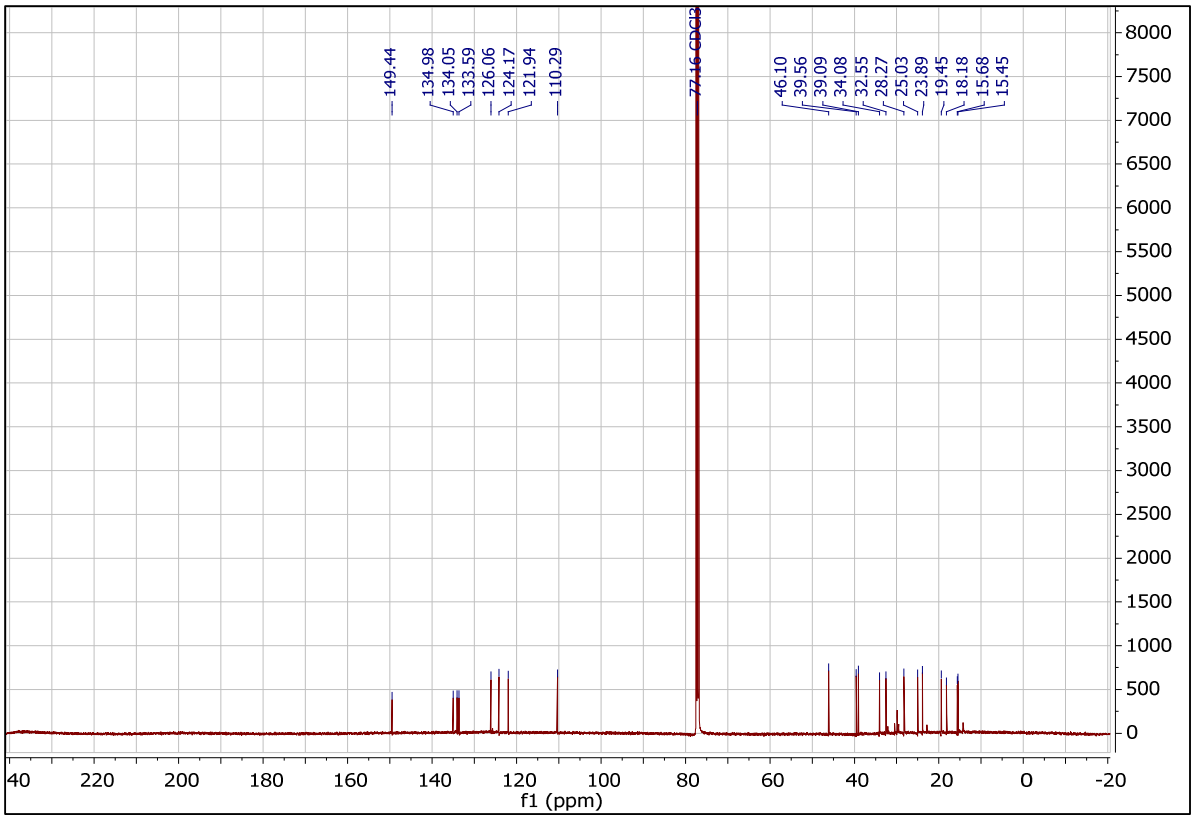


Fig. S8f: 13-C Spectrum of CM

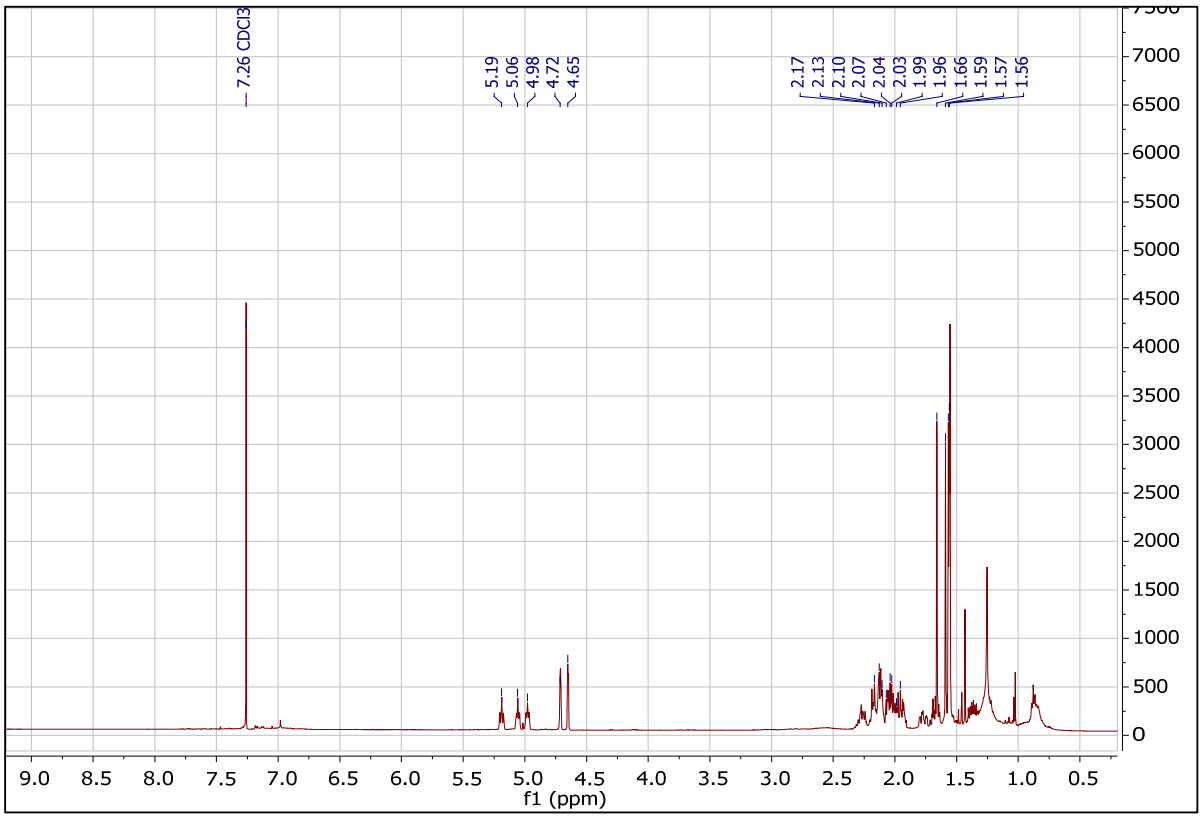


Fig. S8g: 1-H Spectrum of CM

(a) NMR spectral data for (+)-(*R*, *R)*-verticilla-3,7,12(13)-triene				
V584M/L				
#	***δc* (ppm)**	***δ*$_H$ (ppm), *J* (Hz)**	**HMBC**	**NOESY**
1	42.5	1.51	16, 17	
2	34.3	1.84 , 2.64	1	1, 17, 20,
3	124.8	5.29 (br d, J = 11.9)	2, 20	5, 7, 11, 14, 18
4	132.8	n/a	20	
5	41.1	2.02 , 2.17	6, 20	
6	26.8	1.99 , 2.44	5	5
7	130.0	4.83 (br d, J = 12.0)	6, 19	3, 5, 6, 9, 11, 18
8	133.1	n/a	19	
9	39.7	1.96 , 2.16	10, 11	
10	21.6	1.37 , 1.67	9, 16	
11	38.2	2.97	10, 16, 17	18, 3, 7
12	136.0	n/a	10, 14	
13	121.8	5.39 (br s)	14	14, 18
14	30.9	1.87, 2.47	1	2
15	35.8	n/a	1, 16, 17	
16	27.3	0.74 (br s)	17	
17	23.8	0.81 (br s)	16	
18	23.1	1.76 (br s)	10, 11	
19	15.9	1.50 (br s)	9	6, 2
20	15.4	1.58 (br s)	2	6

(b)

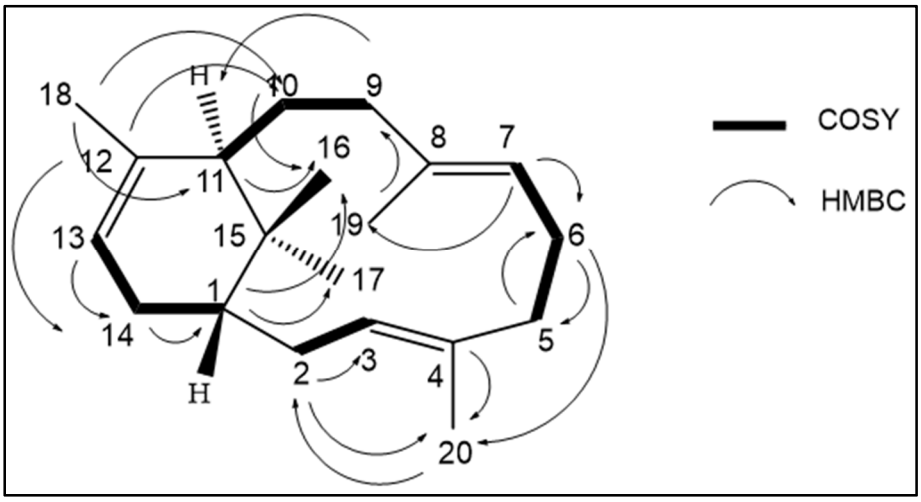


Fig. S9a: NMR Spectral data of V

(a) NMR Spectra table of **V**. **(b)** Structure of **V** showing selected key correlations in the COSY and HMBC NMR spectra.

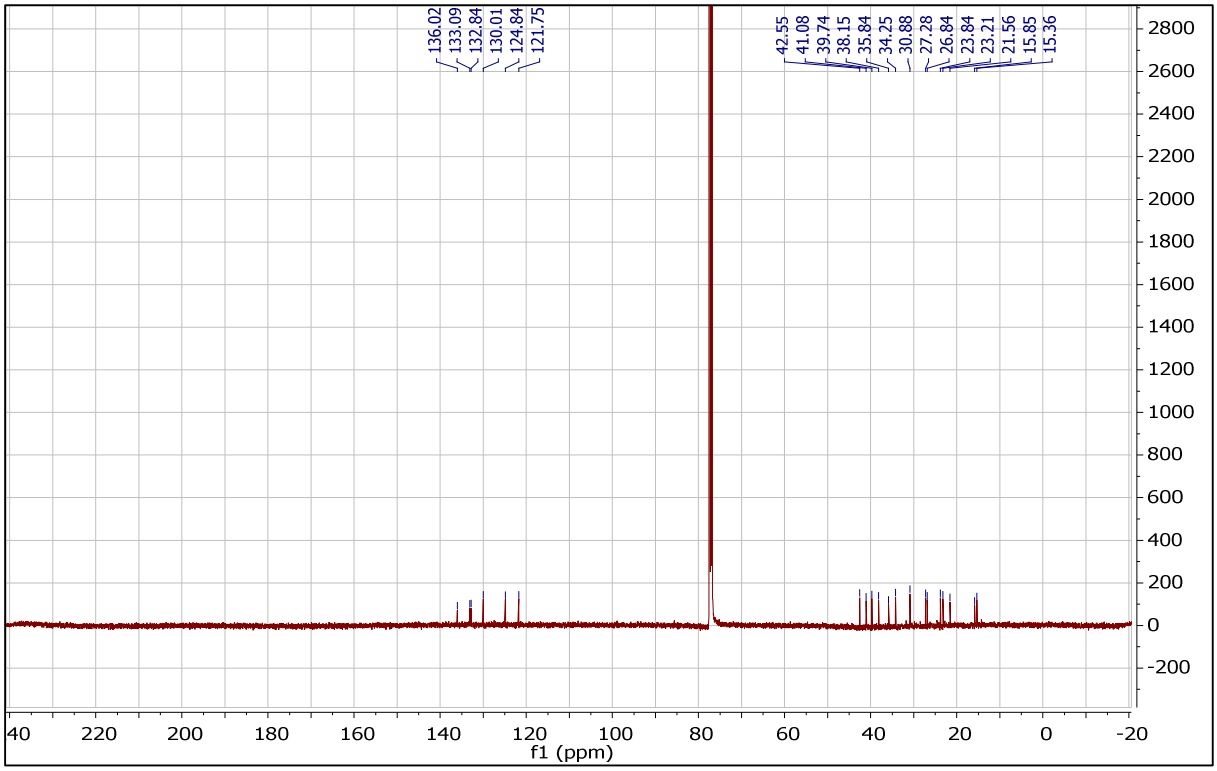


Fig. S9b: 13-C Spectrum of V

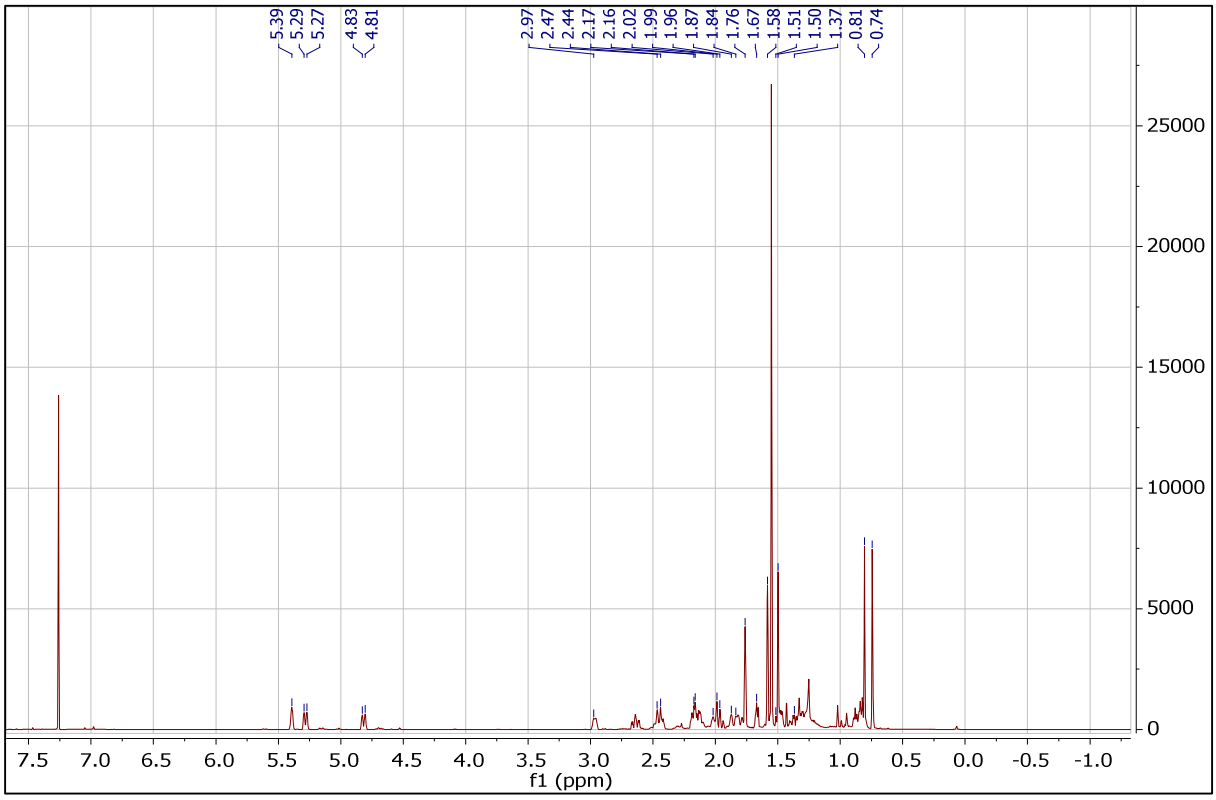


Fig. S9c: 1H-Spectrum of V

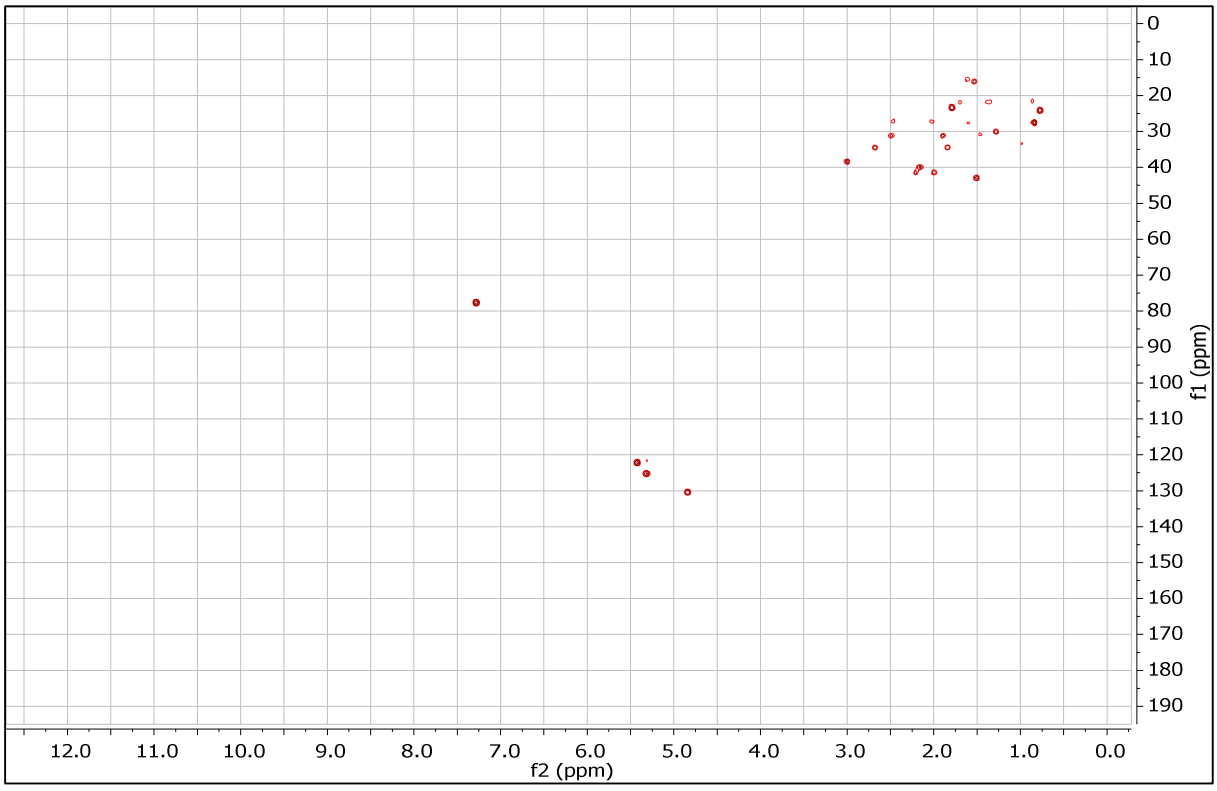


Fig. S9d: HSQC Spectrum of V

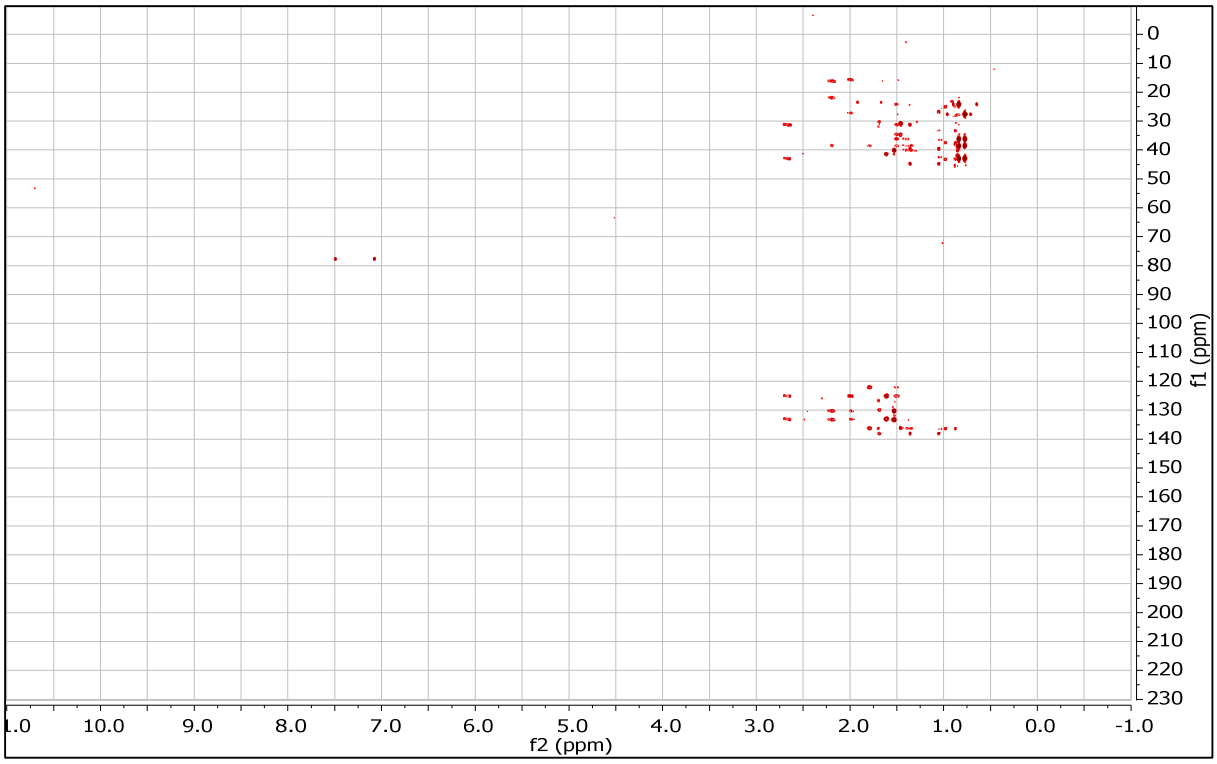


Fig. S9e: HMBC Spectrum of V

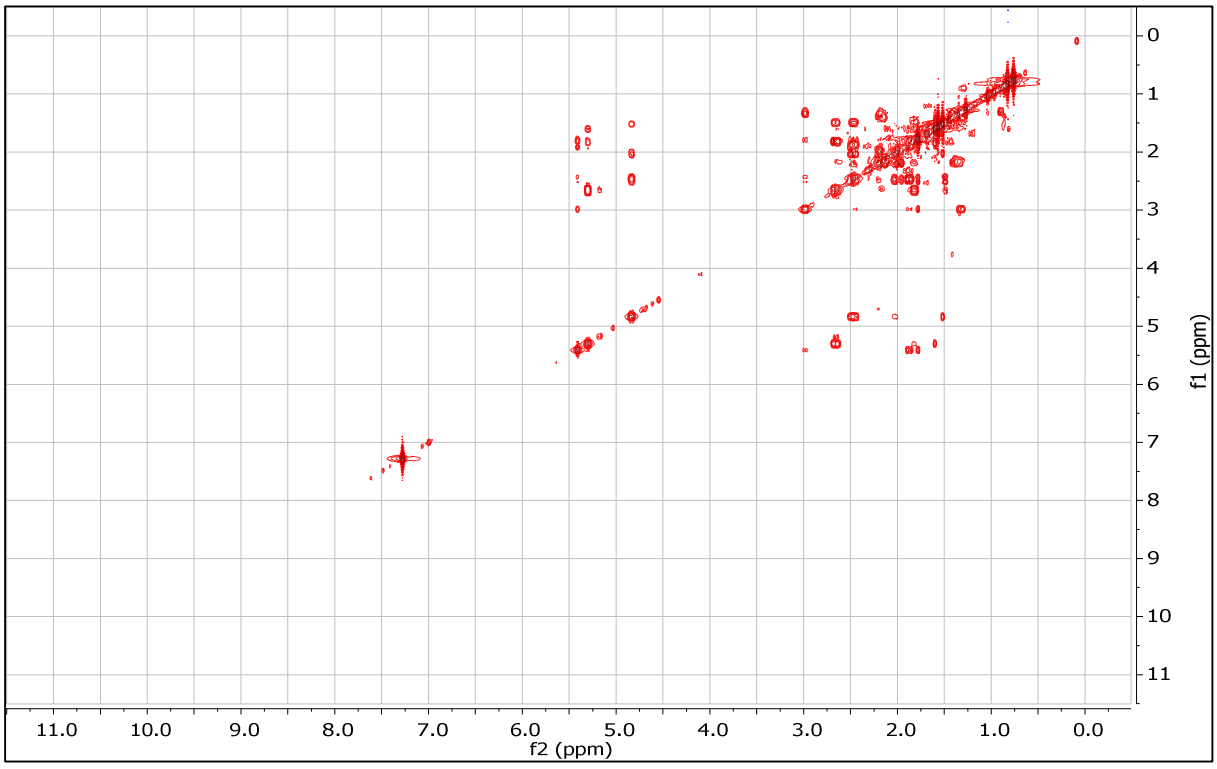


Fig. S9f: COSY Spectrum of V

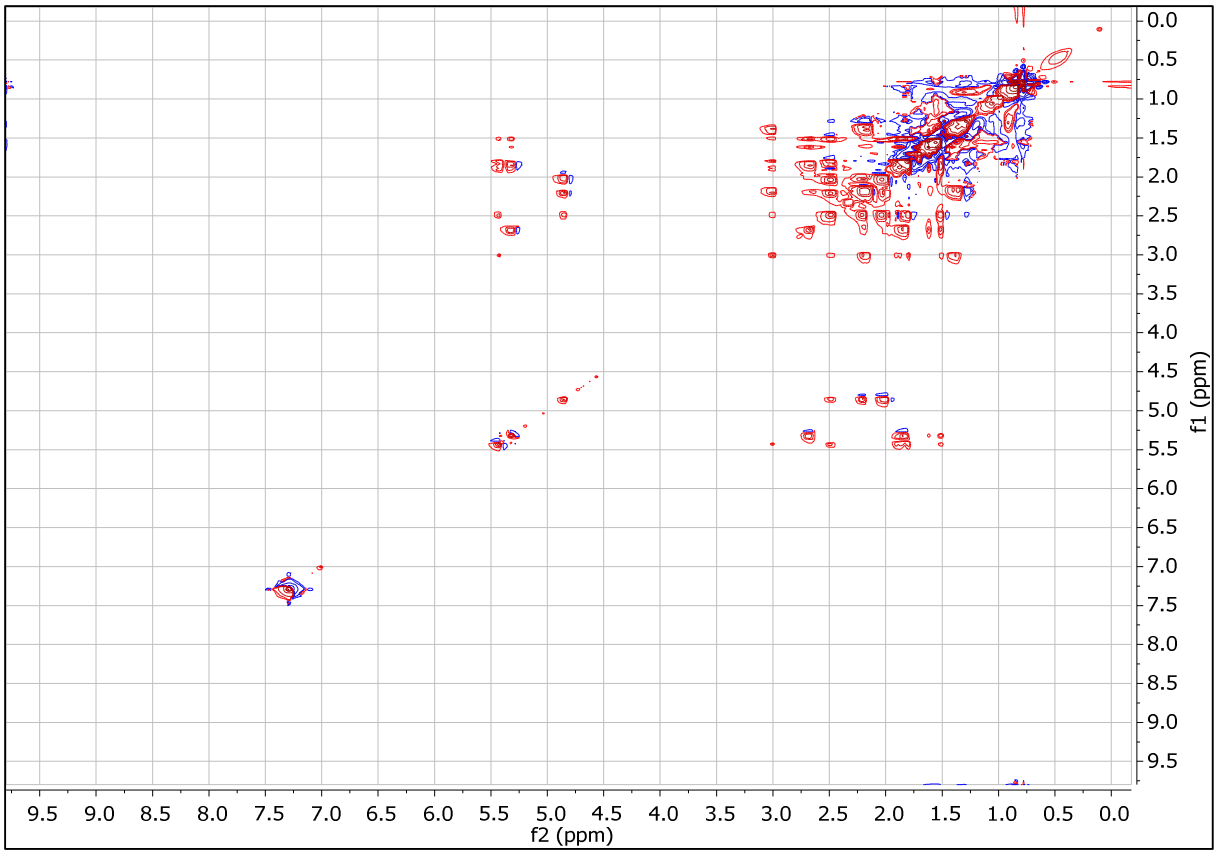


Fig. S9g: TOCSY Spectrum of V

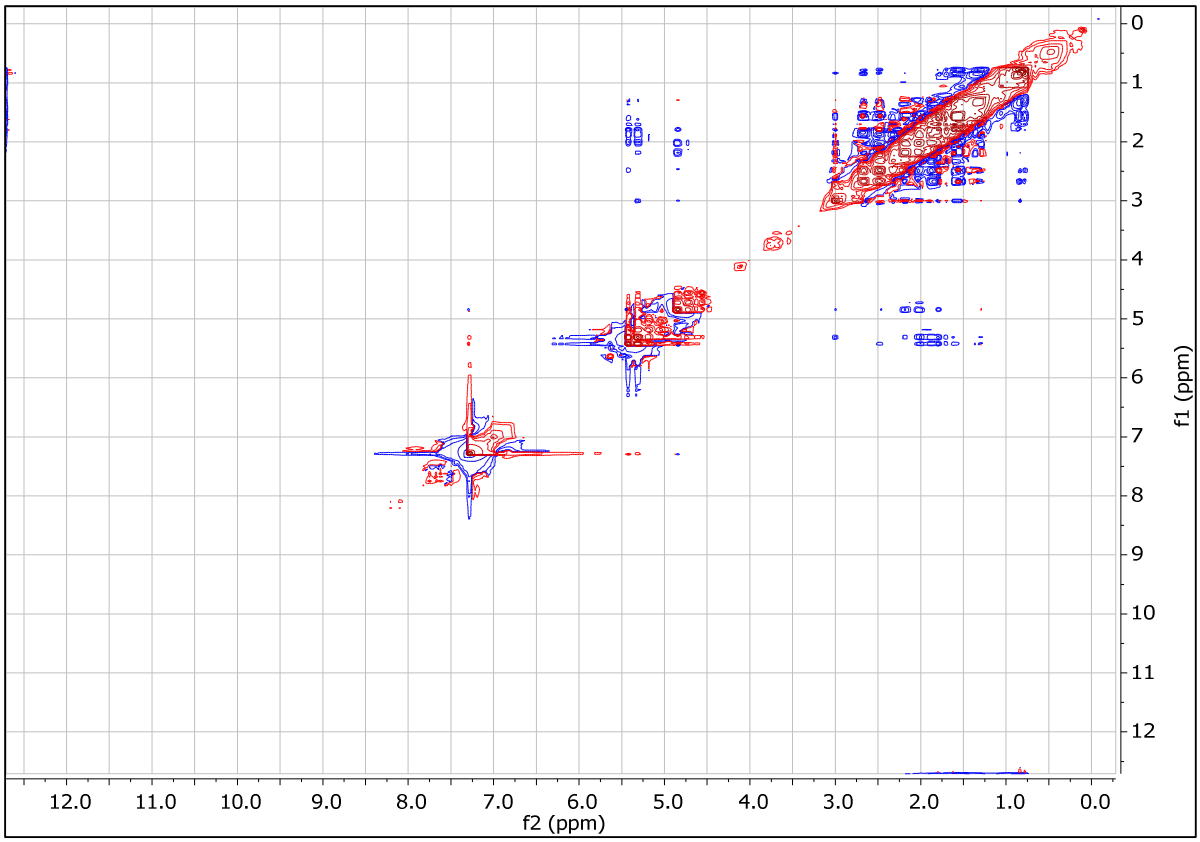


Fig. S9h: NOESY Spectrum of V

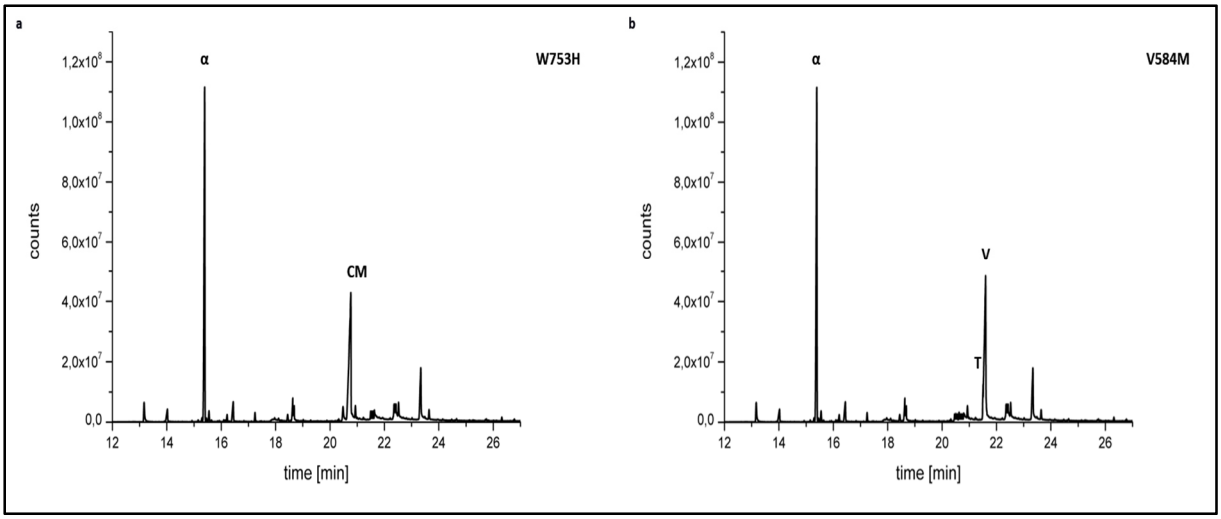


Fig. S10a: GC-FID Spectra of batch bioprocess supernatants of TXS-W753H and TXS-V584M

(a) GC-FID spectrum of the extracted supernatant (500ml) of the batch bioprocess of TXS-W753H. (α: alpha humulene, **CM**: cembrene A). **(b)** GC-FID spectrum of the extracted supernatant (500ml) of the batch bioprocess of TXS-V584M. (α: alpha humulene, **V**: verticilla-3,7,12(13)-triene, **T**: taxa-4,11-diene).

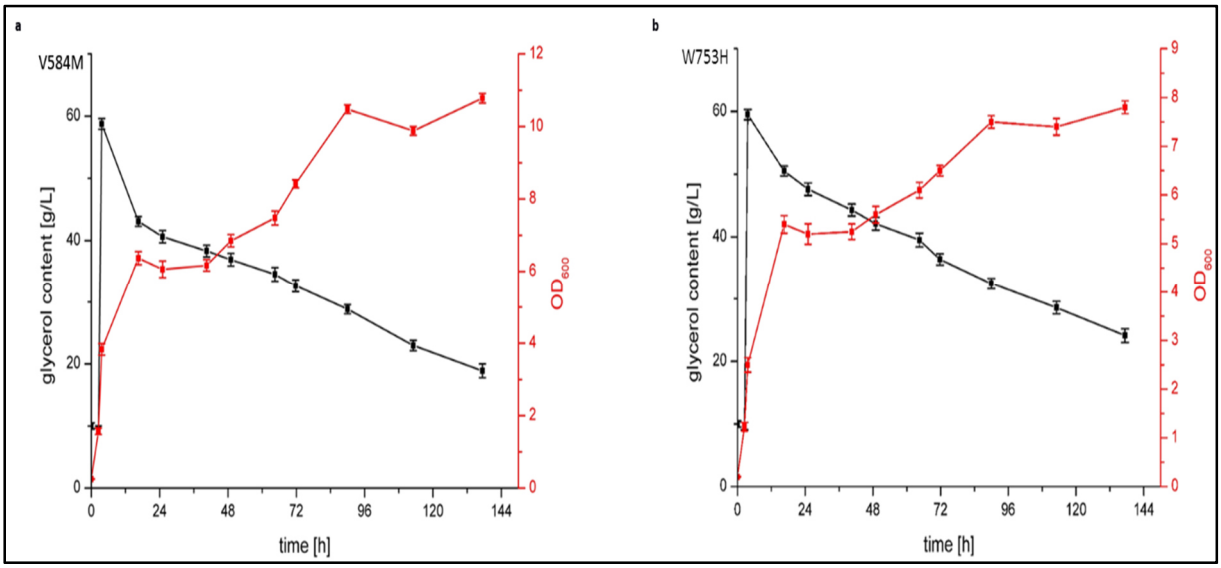


Fig. S10b: Batch bioprocess characteristica of TXS-W753H and TXS-V584M

(a) Time course of glycerol content (black line) and OD_{600} during the 30L batch fermentation process of TXS-V584M-*in* vivo production system with glycerol as sole carbon source. **(b)** Time course of glycerol content (black line) and OD_{600} during the 30L batch fermentation process of TXS-W753H-*in* vivo production system with glycerol as sole carbon source.

	Predicted free energy of binding for the cationic ligand with lowest dissociation constant in the respective cluster [kcal mol^{-1}]	Dissociation constant for the cationic ligand with highest predicted free energy of binding in the respective cluster [pM]	Free energy of binding spread of cationic ligands in the respective clusters [kcal mol^{-1}] (average and standard deviation)
Cation A			
Cluster			
1	8.4720	616512.1875	8.472 +- 0.0
2	**7.8910**	**1643691.5000**	**7.891 +- 0.0**
Cation B			
Cluster			
1	10.2850	28906.4688	9.3118 +- 0.5918
2	**8.6750**	**437663.8750**	**8.1593 +- 0.4584**
TS B-C			
Cluster			
1	10.0050	46370.0703	6.9122 +- 2.4141
2	**7.9690**	**1440938.8750**	**6.7142 +- 1.6325**
Cation C			
Cluster			
1	9.3070	150618.5781	7.5646 +- 1.2115
2	**8.4770**	**611331.2500**	**7.3673 +- 0.8611**
Cation F			
Cluster			
1	**10.0840**	**40581.6758**	**7.1243 +- 2.0671**
Cation D1			
Cluster			
1	9.1340	201692.2500	7.4407 +- 1.1357
2	**8.1720**	**1022926.8750**	**6.8544 +- 1.5005**
Cation D2			
Cluster			
1	9.1340	201692.2500	7.0101 +- 1.6486
2	**9.0110**	**248227.2344**	**6.7997 +- 1.7352**
Cation E			
Cluster			
1	**11.3550**	**4749.7925**	**6.7744 +- 2.4352**
2	7.9450	1500506.1250	6.7014 +- 1.0455

Table S3: Cluster analyses of docked QM-carbocations into the transient enzyme intermediate complexes

Cluster analyses of **QM-cations A-E** and -**TS B-C** docked into their corresponding transient enzyme intermediate complexes according to methods part 3 and 4. **Cations A**, -**B**, -**C**, -**D1**, -**D2**, -**E** and **TS B**-C exhibit two distinct conformational clusters in their corresponding transient enzyme intermediate complexes. **Cation F** exhibits only one distinct conformation. Two conformations belonged to different clusters, if the ligand RMSD was larger than 2 Å. Only cluster that are distinctly derived from the conformation of unionized **GGPP** (**cation A**) or the respective former cascade cation (**cation B**-**E** and **TS B**-**C**), depicted in red were considered for validation and structural

superposition of the respective manually formed cations. These cationic structures are shown in **Fig. 5** (**cation E**), **Fig. 2c, d** and **Fig. S3a** (**cation A**), **Fig. 3a, b** (**cation B** and **TS B**-**C**), **Fig. 4a and Fig. S3c** (**cation C**), **S3e** (**cation F**), **S3f** (**cation D1**) and **S3g** (**cation D2**). The conformational cluster that is not structurally derived from unionized **GGPP** or the respective former cascade cation, depicted in black are shown in **Fig. S3a** (**cation A**) and **S6** (**cation B**-**E** and **TS B**-**C**). Note that single conformations of **cations F** and -**E** in their respective considered cluster exhibit the overall highest predicted free energies of binding (second column), indicating the most favorable interaction with their respective receptor. This in line with the assumption that **cations C**-**E** are not in equilibrium state as suggested by QM gas phase calculations.[10] Instead, it indicates that **cation C**->-**F** transition is an exothermic event and **cation E** is indeed the global energy minimum of the reaction cascade in relation to **cation A**. Moreover, cluster analyses demonstrate that according to the predicted spread in free energies of binding (fourth column) the average predicted free energies of binding of the cationic ligands decrease from **cation B**->-**E** and that cation **A** and **B** exhibit the highest average free energies of binding as well as the lowest standard deviations. This indicates that a high percentage of conformations in the cluster of **cation A** and **B** exhibit a more favorable interaction with the respective receptor than **cations C**-**E**. This, in turn, is in line with our observation that **cations C**->-**E** suffer from a reduced binding in TXS, resulting in cation tumbling and imprecise barrier crossings that lead to the observed side products derived from these cationic intermediates.

Target	Mutation	Activity [a]	T [b]	T1	CM	V	V1	V2
TXS		100,0	93,2	4,7	N.D.	0,8	0,2	1,1
Y89	Y89F	0,0						
	Y89A	0,0						
	Y89E	2,8	100,0	N.D.	N.D.	N.D.	N.D.	N.D.
R580	R580H	0,0						
	R580A	0,0						
	R580E	0,0						
R754	R754H	0,0						
	R754A	0,0						
	R754E	0,0						
R768	R768H	0,0						
	R768A	0,0						
	R768E	0,0						
V584	V584M	92,3	13,8	0,6	N.D.	83,4	0,8	1,4
	V584K	89,1	29,8	0,6	N.D.	67,4	0,8	1,4
	V584L	92,1	13,8	0,6	N.D.	83,4	0,8	1,4
	V584N	0,0						
	V584S	0,0						
	V584P	0,0						
	V584R	0,0						
S587	S587D	0,0						
	S587Y	0,0						
	S587K	0,0						
	S587L	0,0						
	S587A	21,8	8,9	N.D.	58,9	32,2	N.D.	N.D.
	S587G	0,0						
F602	F602W	9,0	100,0	N.D.	N.D.	N.D.	N.D	N.D.
	F602I	22,0	93,4	4,9	N.D.	N.D.	N.D.	1,7
	F602A	1,0	100,0	N.D.	N.D.	N.D.	N.D.	N.D.
V610	V610H	0,0						
	V610S	0,0						
	V610F	0,0						
	V610A	0,0						
S713	S713T	97,4	92,7	5,1	N.D.	1,0	0,4	0,8
	S713A	0,0						
	S713L	0,0						
V714	V714T	1,2	100,0	N.D.	N.D.	N.D.	N.D.	N.D.
	V714A	8,7	94,3	5,7	N.D.	N.D.	N.D.	N.D.
	V714I	10,4	94,1	5,9	N.D.	N.D.	N.D.	N.D.
	V714G	0,0						
	V714P	0,0						

G715	G715A	0,0						
	G715S	0,0						
S713	S713T	97,4	N.D.	N.D.	100,0	N.D.	N.D.	N.D.
	S713A	0,0						
	S713L	0,0						
W753	W753H	51,3	N.D.	N.D.	100,0	N.D.	N.D.	N.D.
	W753L	0,0						
	W753V	0,0						
	W753E	0,0						
	W753C	0,0						
	W753H/C830A	48,5	N.D.	N.D.	100,0	N.D.	N.D.	N.D.
	W753A	0,0	93,5	4,6	N.D.	0,7	0,2	1,0
C830	C830A	88,5	93,1	4,8	N.D.	0,8	0,2	1,1
	C830S	92,4	79,3	7,3	N.D.	1,1	2,1	10,2
F834	F834A	25,6	38,3	6,5	N.D.	8,8	21,9	24,5
	F834G	28,3	87,4	6,1	N.D.	0,8	2,5	3,2
	F834Y	93,1						
	F834H	0,0						
Y835	Y835F	3,5	100,0	N.D.	N.D.	N.D.	N.D.	N.D.
	Y835W	0,0						
	Y835A	0,0						
Y841	Y841A	0,0						
	Y841F	41,3	N.D.	N.D.	56,4	43,6	N.D.	N.D.
	Y841T	0,0						

Table S4: Characterization of TXS and mutants

(a) Enzyme activities of mutants (in %) in comparison to TXS (= 100%). The activities were determined by peak integration of the enzyme products in relation to the peak area of the internal standard alpha-humulene (methods section 14). **(b)** Compound numbers **CM-T1** refer to compound numbers in **Fig. 1, Table 1** and **Supplementary Fig. 23**. The tabulated values represent the product distributions in % for TXS and each mutant (sum = 100%). (N.D.: Not detectable)

TXS Mutant	Sequence 5`-> 3`
Fw_R580H	CTTTACCCGTCATCATGTTGCCGAAGTG
Rv_R580H	CACTTCGGCAACATGATGACGGGTAAAG
Fw_R580A	GATATTAACTTTACCCGTCATGCCGTTGCCGAAGTGTATTTTAG
Rv_R580A	CTAAAATACACTTCGGCAACGGCATGACGGGTAAAGTTAATATC
Fw_R580E	GATATTAACTTTACCCGTCATGAGGTTGCCGAAGTGTATTTTAGC
Rv_R580E	GCTAAAATACACTTCGGCAACCTCATGACGGGTAAAGTTAATATC
Fw_R754H	GTAAGCCTGAGCTGGCATCTGACCAATGATAC
Rv_R754H	GTATCATTGGTCAGATGCCAGCTCAGGCTTAC
Fw_R754A	GGTAAGCCTGAGCTGGGCGCTGACCAATGATACC
Rv_R754A	GGTATCATTGGTCAGCGCCCAGCTCAGGCTTACC
Fw_R754E	GAGCTTGTATCCTTGAGCTGGGAACTAACAAACGACACCAAAAC
Rv_R754E	GTTTTGGTGTCGTTTGTTAGTTCCCAGCTCAAGGATACAAGCTC
Fw_R768H	CAGGCTGAAAAGGCTCATGGACAACAAGCCTCAG
Rv_R768H	CTGAGGCTTGTTGTCCATGAGCCTTTTCAGCCTG
Fw_R768A	CAGGCTGAAAAGGCTGCGGGACAACAAGCCTC
Rv_R768A	GAGGCTTGTTGTCCCGCAGCCTTTTCAGCCTG
Fw_R768E	CAGGCTGAAAAGGCTGAAGGACAACAAGCCTC
Rv_R768E	GAGGCTTGTTGTCCTTCAGCCTTTTCAGCCTG
Fw_V584M	GACACCGAGTGGCGGAGATGTATTTTTCATCAGCTAC
Rv_V584M	GTAGCTGATGAAAAATACATCTCCGCCACTCGGTGTC
Fw_V584K	CGACACCGAGTGGCGGAGAAATATTTTTCATCAGCTAC
Rv_V584K	GTAGCTGATGAAAAATATTTCTCCGCCACTCGGTGTCG
Fw_V584L	CATCGTGTTGCCGAACTGTATTTTAGCAGTG
Rv_V584L	CACTGCTAAAATACAGTTCGGCAACACGATG
Fw_V584N	GACACCGAGTGGCGGAGAATTATTTTTCATCAGCTAC
Rv_V584N	GTAGCTGATGAAAAATAATTCTCCGCCACTCGGTGTC
Fw_V584S	GACACCGAGTGGCGGAGAGCTATTTTTCATCAGCTAC
Rv_V584S	GTAGCTGATGAAAAATAGCTCTCCGCCACTCGGTGTC
Fw_V584P	GACACCGAGTGGCGGAGCCGTATTTTTCATCAGCTAC
Rv_V584P	GTAGCTGATGAAAAATACGGCTCCGCCACTCGGTGTC
Fw_V584H	CGTCATCGTGTTGCCGAACACTATTTTAGCAGTGCAAC
Rv_V584H	GTTGCACTGCTAAAATAGTGTTCGGCAACACGATGACG
Fw_V584R	GTCATCGTGTTGCCGAAAGGTATTTTAGCAGTGCAAC
Rv_V584R	GTTGCACTGCTAAAATACCTTTCGGCAACACGATGAC
Fw_S587D	GTGGCGGAGGTTTATTTTGATTCAGCTACATTTGAACCCG
Rv_S587D	CGGGTTCAAATGTAGCTGAATCAAAATAAACCTCCGCCAC
Fw_S587Y	GGCGGAGGTTTATTTTTATTCAGCTACATTTGAACCC
Rv_S587Y	GGGTTCAAATGTAGCTGAATAAAAATAAACCTCCGCC
Fw_S587K	GTGGCGGAGGTTTATTTTAAATCAGCTACATTTGAACCC
Rv_S587K	GGGTTCAAATGTAGCTGATTTAAAATAAACCTCCGCCAC
Fw_S587L	GTGGCGGAGGTTTATTTTCTGTCAGCTACATTTGAACCC

Rv_S587L	GGGTTCAAATGTAGCTGACAGAAAATAAACCTCCGCCAC
Fw_S587A	GTTGCCGAAGTGTATTTTGCCAGTGCAACCTTTGAACCG
Rv_S587A	CGGTTCAAAGGTTGCACTGGCAAAATACACTTCGGCAAC
Fw_S587G	GTTGCCGAAGTGTATTTTGGCAGTGCAACCTTTGAAC
Rv_S587G	GTTCAAAGGTTGCACTGCCAAAATACACTTCGGCAAC
Fw_F602W	CTGCAGGTCCTGTGGGATGATATGGCC
Rv_F602W	GGCCATATCATCCCACAGGACCTGCAG
Fw_F602I	GTCTGCAGGTCCTGATCGATGATATGGCC
Rv_F602I	GGCCATATCATCGATCAGGACCTGCAGAC
Fw_F602A	GTCTGCAGGTCCTGGCAGATGATATGGCCG
Rv_F602A	CGGCCATATCATCTGCCAGGACCTGCAGAC
Fw_V610H	CAAAATTGGTTGTCTGCAGCACCTGTTCGATGATATGGC
Rv_V610H	GCCATATCATCGAACAGGTGCTGCAGACAACCAATTTTG
Fw_V610S	CAAAATTGGTTGTCTGCAGAGCCTGTTCGATGATATGGC
Rv_V610S	GCCATATCATCGAACAGGCTCTGCAGACAACCAATTTTG
Fw_V610F	CAAAATTGGTTGTCTGCAGTTCCTGTTCGATGATATGG
Rv_V610F	CCATATCATCGAACAGGAACTGCAGACAACCAATTTTG
Fw_V610A	CAAAATTGGTTGTCTGCAGGCCCTGTTCGATGATATGGCC
Rv_V610A	GGCCATATCATATCATCGAACAGGGCCAGCAGACAACCAATTTTG
Fw_V714A	CTTATGCTATATCAGCGGGCCTTGGACCGTG
Rv_V714A	CACGGTCCAAGGCCCGCTGATATAGCATAAG
Fw_V714I	GACTTATGCTATATCAATTGGCCTTGGACCGTGTAC
Rv_V714I	GTACACGGTCCAAGGCCAATTGATATAGCATAAGTC
Fw_V714T	GAAAACCTATGCAATTAGCACCGGTCTGGGTCCGTGTACC
Rv_V714T	GGTACACGGACCCAGACCGGTGCTAATTGCATAGGTTTTC
Fw_V714G	GAAAACCTATGCAATTAGCGGCGGTCTGGGTCCGTGTACCC
Rv_V714G	GGGTACACGGACCCAGACCGCCGCTAATTGCATAGGTTTTC
Fw_V714P	GACTTATGCTATATCACCGGGCCTTGGACCGTGTAC
Rv_V714P	GTACACGGTCCAAGGCCCGGTGATATAGCATAAGTC
Fw_G715A	CTTATGCTATATCAGTAGCGCTTGGACCGTGTACCCTAC
Rv_G715A	GTAGGGTACACGGTCCAAGCGCTACTGATATAGCATAAG
Fw_G715S	CTTATGCTATATCAGTAAGCCTTGGACCGTGTAC
Rv_G715S	GTACACGGTCCAAGGCTTACTGATATAGCATAAG
Fw_S713T	CTTAAAGACTTATGCTATAACAGTAGGCCTTGGACCGTG
Rv_S713T	CACGGTCCAAGGCCTACTGTTATAGCATAAGTCTTTAAG
Fw_S713A	CTTAAAGACTTATGCTATAGCGGTAGGCCTTGGACCGTGTAC
Rv_S713A	GTACACGGTCCAAGGCCTACCGCTATAGCATAAGTCTTTAAG
Fw_S713L	CTTAAAGACTTATGCTATACTGGTAGGCCTTGGACCGTGTAC
Rv_S713L	GTACACGGTCCAAGGCCTACCAGTATAGCATAAGTCTTTAAG
Fw_W753H	GAACTGGTAAGCCTGAGCCACCGTCTGACCAATGATACC
Rv_W753H	GGTATCATTGGTCAGACGGTGGCTCAGGCTTACCAGTTC
Fw_W753L	GGTAAGCCTGAGCTTGCGTCTGACCAATG
Rv_W753L	CATTGGTCAGACGCAAGCTCAGGCTTACC

Fw_W753V	CTGGTAAGCCTGAGCGTGCGTCTGACCAATG
Rv_W753V	CATTGGTCAGACGCACGCTCAGGCTTACCAG
Fw_W753E	GAGCTTGTATCCTTGAGCGAACGACTAACAAACGACAC
Rv_W753E	GTGTCGTTTGTTAGTCGTTCGCTCAAGGATACAAGCTC
Fw_W753C	GTATCCTTGAGCTGCCGACTAACAAACGAC
Rv_W753C	GTCGTTTGTTAGTCGGCAGCTCAAGGATAC
Fw_W753A	GAACTGGTAAGCCTGAGCGCACGTCTGACCAATGATAC
Rv_W753A	GTATCATTGGTCAGACGTGCGCTCAGGCTTACCAGTTC
Fw_C830A	CCTTTATCTTTAATCTGCGTCTGGCCGTGCAGATCTTCTATAAATTC
Rv_C830A	GAATTTATAGAAGATCTGCACGGCCAGACGCAGATTAAAGATAAAGG
Fw_C830S	CTTTATCTTTAATCTGCGTCTGAGCGTGCAGATCTTCTATAAATTC
Rv_C830S	GAATTTATAGAAGATCTGCACGCTCAGACGCAGATTAAAGATAAAG
Fw_F834A	CTTAGATTGTGTGTCCAAATCGCCTACAAGTTTATAGATGGGTAC
Rv_F834A	GTACCCATCTATAAACTTGTAGGCGATTTGGACACACAATCTAAG
Fw_F834G	CTTAGATTGTGTGTCCAAATCGGCTACAAGTTTATAGATGGGTAC
Rv_F834G	GTACCCATCTATAAACTTGTAGCCGATTTGGACACACAATCTAAG
Fw_F834Y	GATTGTGTGTCCAAATCTACTACAAGTTTATAGATGG
Rv_F834Y	CCATCTATAAACTTGTAGTAGATTTGGACACACAATC
Fw_F834H	CGTCTGTGCGTGCAGATCCACTATAAATTCATTGATG
Rv_F834H	CATCAATGAATTTATAGTGGATCTGCACGCACAGACG
Fw_Y835F	GTGTGTCCAAATCTTTTTTAAGTTTATAGATGGGTAC
Rv_Y835F	GTACCCATCTATAAACTTAAAAAAGATTTGGACACAC
Fw_Y835W	GTGTGTCCAAATCTTTTGGAAGTTTATAGATGGGTAC
Rv_Y835W	GTACCCATCTATAAACTTCCAAAAGATTTGGACACAC
Fw_Y835A	GATTGTGTGTCCAAATCTTTGCGAAGTTTATAGATGGGTACG
Rv_Y835A	CGTACCCATCTATAAACTTCGCAAAGATTTGGACACACAATC
Fw_Y841A	CTATAAATTCATTGATGGTGCCGGCATTGCCAACGAAGAG
Rv_Y841A	CTCTTCGTTGGCAATGCCGGCACCATCAATTTATAG
Fw_Y841F	CTATAAATTCATTGATGGTTTCGGCATTGCCAACGAAGAG
Rv_Y841F	CTCTTCGTTGGCAATGCCGAAACCATCAATGAATTTATAG
Fw_Y841T	CTATAAATTCATTGATGGTACCGGCATTGCCAACGAAGAG
Rv_Y841T	CTCTTCGTTGGCAATGCCGGTACCATCAATGAATTTATAG
Fw_Y89F	GTCTGAGCGCAAATTTTCATGGTGATCTGTG
Rv_Y89F	CACAGATCACCATGAAAATTTGCGCTCAGAC
Fw_Y89A	ACTCTCCGCCAATGCGCATGGCGATCTGTG
Rv_Y89A	CACAGATCGCCATGCGCATTGGCGGAGAGT
Fw_Y89E	CGACTCTCCGCCAATGAGCATGGCGATCTGTG
Rv_Y89E	CACAGATCGCCATGCTCATTGGCGGAGAGTCG

Table S5: Mutagenesis Primer

CCATGGATGAGTTTTGATATTGCCAAATACCCGACCCTGGCACTGGTCGACTCCACCCAGGAGTT
ACGACTGTTGCCGAAAGAGAGTTTACCGAAACTCTGCGACGAACTGCGCCGCTATTTACTCGACA
GCGTGAGCCGTTCCAGCGGGCACTTCGCCTCCGGGCTGGGCACGGTCGAACTGACCGTGGCGC
TGCACTATGTCTACAACACCCCGTTTGACCAATTGATTTGGGATGTGGGGCATCAGGCTTATCCG
CATAAAATTTTGACCGGACGCCGCGACAAAATCGGCACCATCCGTCAGAAAGGCGGCCTGCACC
CGTTCCCGTGGCGCGGCGAAAGCGAATATGACGTATTAAGCGTCGGGCATTCATCAACCTCCAT
CAGTGCCGGAATTGGTATTGCGGTTGCTGCCGAGAAAGAAGGCAAAAATCGCCGCACCGTCTGT
GTCATTGGCGATGGCGCGATTACCGCTGGCATGGCGTTTGAAGCGATGAATCACGCGGGCGATA
TCCGTCCTGATATGCTGGTGGTCCTCAACGACAATGAAATGTCGATTTCCGAAAATGTCGGCGCG
CTCAATAACCATCTGGCACAGCTGCTTTCCGGTAAGCTTTACTCTTCGCTGCGCGAAGGCGGGAA
AAAAGTTTTCTCTGGCGTTCCGCCAATTAAAGAGCTGCTCAAACGTACCGAAGAACATATTAAAGG
CATGGTAGTGCCTGGCACGTTGTTTGAAGAGCTGGGCTTTAACTACATCGGCCCGGTTGACGGT
CACGATGTGCTGGGGCTTATCACCACGCTGAAGAACATGCGCGACCTGAAAGGCCCGCAGTTCC
TGCATATCATGACCAAAAAAGGTCGTGGTTATGAACCGGCAGAAAAAGACCCCATCACTTTCCAC
GCCGTGCCTAAATTTGATCCCTCCAGCGGTTGTTTGCCGAAAAGTAGCGGCGGTTTGCCGAGCT
ATTCAAAAATCTTTGGCGACTGGTTGTGCGAAACGGCAGCGAAAGACAACAAGCTGATGGCGATT
ACTCCGGCGATGCGTGAAGGTTCCGGCATGGTCGAGTTTTCACGTAAATTCCCGGATCGTTACTT
CGACGTGGCAATCGCCGAGCAACACGCGGTGACCTTTGCCGCCGGTCTGGCGATTGGTGGGTA
CAAACCCATTGTCGCGATTTACTCCACTTTCCTGCAACGCGCCTATGATCAGGTGCTGCATGACG
TGGCGATTCAAAAGCTCCCGGTCCTGTTCGCCATCGACCGCGCGGGCATTGTTGGTGCTGACGG
TCAAACCCATCAGGGCGCTTTTGACCTCTCTTACCTGCGCTGTATACCGGAAATGGTCATTATGA
CCCCGAGCGATGAAAACGAATGTCGCCAGATGCTCTATACCGGCTATCACTATAACGACGGCCC
GTCCGCGGTGCGCTACCCGCGCGGTAACGCGGTTGGCGTGGAACTGACGCCGCTGGAAAAACT
GCCAATTGGCAAAGGCATTGTGAAGCGTCGTGGCGAGAAACTGGCGATCCTTAACTTTGGTACG
CTGATGCCAGACGCGGCGAAAGTCGCTGAATCGCTGAACGCTACGCTGGTCGATATGCGTTTTG
TGAAACCGCTTGATGAAGCGTTAATTCTGGAAATGGCCGCCAGCCATGAAGCGCTGGTCACCGT
AGAAGAAAACGCCATTATGGGCGGCGCAGGCAGCGGCGTGAACGAAGTGCTAATGGCCCATCG
TAAACCAGTACCCGTGCTGAACATTGGCCTGCCTGACTTCTTTATTCCACAAGGAACTCAGGAAG
AAATGCGCGCCGAACTCGGCCTCGATGCCGCCGGTATGGAAGCCAAAATCAAGGCCTGGCTGG
CATAAGAATTC

Gene of 1-deoxy-D-xylulose 5-phosphate synthase (*dxs*)

CATATGAAGCAACTCACCATTCTGGGCTCGACCGGCTCGATTGGTTGCAGCACGCTGGACGTGG
TGCGCCATAATCCCGAACACTTCCGCGTAGTTGCGCTGGTGGCAGGCAAAAATGTCACTCGCAT
GGTAGAACAGTGCCTGGAATTCTCTCCCCGCTATGCCGTAATGGACGATGAAGCGAGTGCGAAA
CTTCTTAAAACGATGCTACAGCAACAGGGTAGCCGCACCGAAGTCTTAAGTGGGCAACAAGCCG
CTTGCGATATGGCAGCGCTTGAGGATGTTGATCAGGTGATGGCAGCCATTGTTGGCGCTGCTGG
GCTGTTACCTACGCTTGCTGCGATCCGCGCGGGTAAAACCATTTTGCTGGCCAATAAAGAATCAC
TGGTTACCTGCGGACGTCTGTTTATGGACGCCGTAAAGCAGAGCAAAGCGCAATTGTTACCGGT
CGATAGCGAACATAACGCCATTTTTCAGAGTTTACCGCAACCTATCCAGCATAATCTGGGATACG
CTGACCTTGAGCAAAATGGCGTGGTGTCCATTTTACTTACCGGGTCTGGTGGCCCTTTCCGTGAG
ACGCCATTGCGCGATTTGGCAACAATGACGCCGGATCAAGCCTGCCGTCATCCGAACTGGTCGA
TGGGGCGTAAAATTTCTGTCGATTCGGCTACCATGATGAACAAAGGTCTGGAATACATTGAAGCG
CGTTGGCTGTTTAACGCCAGCGCCAGCCAGATGGAAGTGCTGATTCACCCGCAGTCAGTGATTC
ACTCAATGGTGCGCTATCAGGACGGCAGTGTTCTGGCGCAGCTGGGGGAACCGGATATGCGTAC
GCCAATTGCCCACACCATGGCATGGCCGAATCGCGTGAACTCTGGCGTGAAGCCGCTCGATTTT
TGCAAACTAAGTGCGTTGACATTTGCCGCACCGGATTATGATCGTTATCCATGCCTGAAACTGGC
GATGGAGGCGTTCGAACAAGGCCAGGCAGCGACGACAGCATTGAATGCCGCAAACGAAATCACC
GTTGCTGCTTTTCTTGCGCAACAAATCCGCTTTACGGATATCGCTGCGTTGAATTTATCCGTACTG
GAAAAAATGGATATGCGCGAACCACAATGTGTGGACGATGTGTTATCTGTTGATGCGAACGCGCG
TGAAGTCGCCAGAAAAGAGGTGATGCGTCTCGCAAGCTGACTCGAG

Gene of 1-deoxy-D-xylulose 5-phosphate reductoisomerase (*dxr*)

CCATGGATGGCAACCACTCATTTGGATGTTTGCGCCGTGGTTCCGGCGGCCGGATTTGGCCGTC
GAATGCAAACGGAATGTCCTAAGCAATATCTCTCAATCGGTAATCAAACCATTCTTGAACACTCGG
TGCATGCGCTGCTGGCGCATCCCCGGGTGAAACGTGTCGTCATTGCCATAAGTCCTGGCGATAG
CCGTTTTGCACAACTTCCTCTGGCGAATCATCCGCAAATCACCGTTGTAGATGGCGGTGATGAGC
GTGCCGATTCCGTGCTGGCAGGTTTGAAAGCCGCTGGCGACGCGCAGTGGGTATTGGTGCATG
ACGCCGCTCGTCCTTGTCTGCATCAGGATGACCTCGCGCGATTGTTGGCGTTGAGCGAAACCAG
CCGCACGGGAGGGATCCTAGCCGCACCAGTGCGCGATACGATGAAACGTGCCGAACCGGGCAA
AAATGCCATTGCTCATACCGTTGATCGCAACGGCTTATGGCACGCGCTGACGCCGCAATTTTTCC
CTCGTGAGCTGTTACATGACTGTCTGACGCGCGCTCTAAATGAAGGCGCGACTATTACCGACGAA
GCCTCGGCGCTGGAATATTGCGGATTCCATCCTCAGTTGGTCGAAGGCCGTGCGGATAACATTA
AAGTCACGCGCCCGGAAGATTTGGCACTGGCCGAGTTTTACCTCACCCGAACCATCCATCAGGA
GAATACATAAGCAGGAGCAGGAGCAGAAGGAGGAGCAGGAATGCGAATTGGACACGGTTTTGAC
GTACATGCCTTTGGCGGTGAAGGCCCAATTATCATTGGTGGCGTACGCATTCCTTACGAAAAAGG
ATTGCTGGCGCATTCTGATGGCGACGTGGCGCTCCATGCGTTGACCGATGCATTGCTTGGCGCG
GCGGCGCTGGGGGATATCGGCAAGCTGTTCCCGGATACCGATCCGGCATTTAAAGGTGCCGAC
AGCCGCGAGCTGCTACGCGAAGCCTGGCGTCGTATTCAGGCGAAGGGTTATACCCTGGGCAAC
GTCGATGTCACTATCATCGCTCAGGCACCGAAGATGTTGCCGCACATTCCACAAATGCGCGTATT
TATTGCCGAAGATCTCGGCTGCCATATGGATGATGTTAACGTGAAAGCCACTACTACGGAAAAAC
TTGGATTTACCGGACGTGGGGAAGGGATTGCCTGTGAAGCGGTGGCGCTACTCATTAAGGCAAC
AAAATGAGAATTC

Gene of bi-cistronic *ispD/ispF*

Bi-cistronic operon of 2-C-methyl-D-erythriol 4-phosphate cytidyltransferase synthase (*ispD*) 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (*ispF*)

CATATGCAAACGGAACACGTCATTTTATTGAATGCACAGGGAGTTCCCACGGGTACGCTGGAAAA
GTATGCCGCACACACGGCAGACACCCGCTTACATCTCGCGTTCTCCAGTTGGCTGTTTAATGCCA
AAGGACAATTATTAGTTACCCGCCGCGCACTGAGCAAAAAAGCATGGCCTGGCGTGTGGACTAA
CTCGGTTTGTGGGCACCCACAACTGGGAGAAAGCAACGAAGACGCAGTGATCCGCCGTTGCCGT
TATGAGCTTGGCGTGGAAATTACGCCTCCTGAATCTATCTATCCTGACTTTCGCTACCGCGCCAC
CGATCCGAGTGGCATTGTGGAAAATGAAGTGTGTCCGGTATTTGCCGCACGCACCACTAGTGCG
TTACAGATCAATGATGATGAAGTGATGGATTATCAATGGTGTGATTTAGCAGATGTATTACACGGT
ATTGATGCCACGCCGTGGGCGTTCAGTCCGTGGATGGTGATGCAGGCGACAAATCGCGAAGCCA
GAAAACGATTATCTGCATTTACCCAGCTTAAATAACTCGAG

Gene of Isopentenyl-diphosphate delta isomerase (*idi*)

CATATGGCAATGAGCAGCAGCACCGGCACCAGCAAAGTTGTTAGCGAAACCAGCAGTACCATTG
TTGATGATATTCCGCGTCTGAGCGCAAATTATCATGGTGATCTGTGGCATCATAATGTGATTCAGA
CCCTGGAAACCCCGTTTCGTGAAAGCAGCACCTATCAAGAACGTGCAGATGAACTGGTTGTGAAA
ATCAAAGATATGTTTAACGCACTGGGTGATGGTGATATTAGCCCGAGCGCCTATGATACCGCATG
GGTTGCACGTCTGGCAACCATTAGCAGTGATGGTAGCGAAAAACCGCGTTTTCCGCAGGCACTG
AATTGGGTTTTTAACAATCAGCTGCAGGATGGTAGTTGGGGTATTGAAAGCCATTTTAGCCTGTGT
GATCGTCTGCTGAATACCACCAATAGCGTTATTGCACTGAGCGTTTGGAAAACCGGTCATAGCCA
GGTTCAGCAGGGTGCAGAATTTATTGCAGAAAATCTGCGCCTGCTGAATGAAGAAGATGAGCTGA
GTCCGGATTTTCAGATTATCTTTCCGGCACTGCTGCAGAAAGCAAAAGCACTGGGTATTAATCTG
CCGTATGATCTGCCGTTTATCAAATATCTGAGCACCACCCGTGAAGCACGTCTGACCGATGTTAG
CGCAGCAGCAGATAATATTCCGGCAAATATGCTGAATGCACTGGAAGGTCTGGAAGAAGTTATTG
ACTGGAACAAAATTATGCGCTTCCAGAGCAAAGATGGTAGCTTTCTGAGTAGTCCGGCAAGCACC
GCATGTGTTCTGATGAATACCGGTGATGAAAAATGCTTTACCTTCCTGAATAACCTGCTGGATAAA
TTTGGTGGTTGTGTTCCGTGTATGTATAGCATTGATCTGCTGGAACGTCTGAGCCTGGTTGATAAT
ATTGAACATCTGGGTATTGGTCGCCACTTCAAACAAGAAATTAAAGGTGCACTGGATTACGTGTAT
CGTCATTGGAGCGAACGTGGTATTGGTTGGGGTCGTGATAGCCTGGTTCCGGATCTGAATACAA
CCGCACTGGGCCTGCGTACCCTGCGTATGCATGGTTATAATGTTAGCTCAGATGTGCTGAACAAC
TTTAAAGATGAAAACGGTCGCTTTTTTAGCAGCGCAGGTCAGACCCATGTTGAACTGCGTAGCGT
TGTTAACCTGTTTCGTGCAAGCGATCTGGCATTTCCGGATGAACGTGCAATGGATGATGCACGTA

AATTTGCAGAACCGTATCTGCGTGAAGCCCTGGCCACCAAAATTAGCACCAATACAAAACTGTTTA
AAGAAATCGAATATGTGGTCGAGTATCCGTGGCACATGAGCATTCCTCGTCTGGAAGCACGTAGC
TATATTGATAGCTATGATGATAACTATGTGTGGCAGCGTAAAACCCTGTATCGTATGCCGAGCCTG
AGCAATAGCAAATGTCTGGAACTGGCAAAACTGGATTTTAACATTGTTCAGAGCCTGCACCAAGA
AGAACTGAAACTGCTGACCCGTTGGTGGAAAGAAAGCGGTATGGCAGATATTAACTTTACCCGTC
ATCGTGTTGCCGAAGTGTATTTTAGCAGTGCAACCTTTGAACCGGAATATAGCGCAACCCGTATT
GCCTTTACCAAAATTGGTTGTCTGCAGGTCCTGTTCGATGATATGGCCGATATTTTTGCAACCCTG
GATGAACTGAAAAGTTTTACCGAAGGTGTTAAACGTTGGGATACCAGTCTGCTGCATGAAATCCC
GGAATGTATGCAGACCTGTTTTAAAGTGTGGTTTAAACTGATGGAAGAGGTGAATAACGATGTGG
TTAAAGTTCAGGGTCGCGATATGCTGGCCCATATTCGTAAACCGTGGGAACTGTATTTCAACTGC
TATGTTCAAGAACGCGAATGGCTGGAAGCCGGTTATATTCCGACCTTTGAAGAATATCTGAAAAC
CTATGCAATTAGCGTTGGTCTGGGTCCGTGTACCCTGCAGCCGATTCTGCTGATGGGTGAACTG
GTGAAAGATGATGTTGTTGAGAAAGTTCATTACCCGAGCAACATGTTTGAACTGGTAAGCCTGAG
CTGGCGTCTGACCAATGATACCAAAACCTATCAGGCAGAAAAAGCACGTGGTCAGCAGGCAAGC
GGTATTGCATGTTATATGAAAGACAATCCGGGTGCAACCGAAGAGGATGCAATCAAACATATTTG
TCGTGTTGTTGATCGTGCACTGAAAGAAGCCAGCTTTGAATATTTCAAACCGAGCAACGATATTCC
GATGGGCTGTAAATCCTTTATCTTTAATCTGCGTCTGTGCGTGCAGATCTTCTATAAATTCATTGAT
GGTTACGGCATTGCCAACGAAGAGATCAAAGATTATATCCGCAAAGTGTATATCGATCCGATTCA
GGTTTAACTCGAG

Gene of the M60 truncation of TXS (*txs*)

References

1 Williams, D. C. *et al.* Heterologous expression and characterization of a "Pseudomature" form of taxadiene synthase involved in paclitaxel (Taxol) biosynthesis and evaluation of a potential intermediate and inhibitors of the multistep diterpene cyclization reaction. *Archives of biochemistry and biophysics* **379**, 137-146, doi:10.1006/abbi.2000.1865 (2000).

2 Whittington, D. A. *et al.* Bornyl diphosphate synthase: structure and strategy for carbocation manipulation by a terpenoid cyclase. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 15375-15380, doi:10.1073/pnas.232591099 (2002).

3 Jin, Y., Williams, D. C., Croteau, R. & Coates, R. M. Taxadiene synthase-catalyzed cyclization of 6-fluorogeranylgeranyl diphosphate to 7-fluoroverticillenes. *Journal of the American Chemical Society* **127**, 7834-7842, doi:10.1021/ja050592r (2005).

4 Meguro, A. *et al.* An Unusual Terpene Cyclization Mechanism Involving a Carbon-Carbon Bond Rearrangement. *Angewandte Chemie*, doi:10.1002/anie.201411923 (2015).

5 Meguro, A. *et al.* An unusual terpene cyclization mechanism involving a carbon-carbon bond rearrangement. *Angewandte Chemie* **54**, 4353-4356, doi:10.1002/anie.201411923 (2015).

6 Hong, Y. J. & Tantillo, D. J. The energetic viability of an unexpected skeletal rearrangement in cyclooctatin biosynthesis. *Organic & biomolecular chemistry*, doi:10.1039/c5ob01785h (2015).

7 Görner, C. H., I. Schrepfer, P. Eisenreich, W. Brück, T. Targeted Engineering of Cyclooctat-9-en-7-ol Synthase: AStereospecific Access to Two New Non-naturalFusicoccane-Type Diterpenes. *ChemCatChem* **5**, 3289-3298, doi:10.1002/c ctc.201300285 (2013).

8 Janke, R., Gorner, C., Hirte, M., Bruck, T. & Loll, B. The first structure of a bacterial diterpene cyclase: CotB2. *Acta crystallographica. Section D, Biological crystallography* **70**, 1528-1537, doi:10.1107/S1399004714005513 (2014).

9 Meguro, A., Tomita, T., Nishiyama, M. & Kuzuyama, T. Identification and characterization of bacterial diterpene cyclases that synthesize the cembrane skeleton. *Chembiochem : a European journal of chemical biology* **14**, 316-321, doi:10.1002/cbic.201200651 (2013).

10 Hong, Y. J. & Tantillo, D. J. The taxadiene-forming carbocation cascade. *Journal of the American Chemical Society* **133**, 18249-18256, doi:10.1021/ja2055929 (2011).