Identification of amino acid networks governing catalysis in the closed complex of class I terpene synthases

Patrick Schrepfer[a], Alexander Buettner[a], Christian Goerner[a], Michael Hertel[a], Jeaphianne van Rijn[b], Frank Wallrapp[c], Wolfgang Eisenreich[d], Volker Sieber[e], Robert Kourist[f], and Thomas Brück[a,1]

[a]Division for Industrial Biocatalysis, Technische Universität München, 85748 Garching, Germany; [b]Department of Theoretical Chemistry, Max-Planck-Institut für Kohlenforschung, 45470 Mülheim an der Ruhr, Germany; [c]Chair of Bioinformatics and Computational Biology, Technische Universität München, 85748 Garching, Germany; [d]Chair of Biochemistry, Technische Universität München, 85748 Garching, Germany; [e]Chair of Chemistry of Biogenic Resources, Technische Universität München, 94315 Straubing, Germany; and [f]Junior Research Group for Microbial Biotechnology, Ruhr-Universität Bochum, 44780 Bochum, Germany

Edited by David H. Sherman, University of Michigan, Ann Arbor, MI, and accepted by the Editorial Board January 5, 2016 (received for review October 4, 2015)

Class I terpene synthases generate the structural core of bioactive terpenoids. Deciphering structure–function relationships in the reactive closed complex and targeted engineering is hampered by highly dynamic carbocation rearrangements during catalysis. Available crystal structures, however, represent the open, catalytically inactive form or harbor nonproductive substrate analogs. Here, we present a catalytically relevant, closed conformation of taxadiene synthase (TXS), the model class I terpene synthase, which simulates the initial catalytic time point. In silico modeling of subsequent catalytic steps allowed unprecedented insights into the dynamic reaction cascades and promiscuity mechanisms of class I terpene synthases. This generally applicable methodology enables the active-site localization of carbocations and demonstrates the presence of an active-site base motif and its dominating role during catalysis. It additionally allowed in silico-designed targeted protein engineering that unlocked the path to alternate monocyclic and bicyclic synthons representing the basis of a myriad of bioactive terpenoids.

computational biology | closed complex modeling | protein engineering | terpene synthases | terpene synthase catalysis

Terpene synthases transform aliphatic allylic diphosphate precursors to complex macrocycles, which represent the structural core of numerous bioactive terpenoids (1). Despite their low primary sequence identity, terpene synthases share highly conserved tertiary and quaternary structural features, which are dominated by α-helical barrel folds (1, 2). Taxadiene synthase (TXS/*Taxus brevifolia*) catalyzes the cyclization of the universal diterpene precursor (*E*,*E*,*E*)-geranylgeranyl diphosphate (GGPP, C20) to taxa-4,11-diene (taxadiene, T) (Fig. 1), the first committed biosynthetic step toward the clinically important tumor therapeutic Taxol (3–6). TXS, a class I diterpene synthase, uses a trinuclear Mg^{2+} ion cluster coordinated by two conserved binding motifs [DDXXD and (N,D)DXX(S,T)XXXE] to initiate catalysis. The Mg^{2+} ion cluster facilitates orientation of the GGPP pyrophosphate moiety (PP_i) in the active site, followed by active-site closure and GGPP ionization (7, 8). Although native GGPP-derived carbocation intermediates could hitherto not be trapped experimentally, mechanistic studies using GGPP analogs and quantum chemical calculations [quantum mechanics (QM)] allowed delineation of several potential carbocation intermediates (Fig. 1) (9–15). However, the nature of the deprotonating base directing formation of T remains elusive. Monocyclic and bicyclic carbocations postulated to form during taxadiene formation, structurally resemble the macrocyclic core of other bioactive diterpenoids (16–20). The molecular mechanism of TXS, which can be considered as template for all class I terpene synthases, may therefore improve a general understanding of terpene synthase mechanisms. TXS and other class I terpene synthases produce side-products [e.g., taxa-4(20),11-diene, T1] (Fig. 1). However, structural data and mechanistic considerations explaining this promiscuity on the molecular level are very limited (8, 21–25). The role of the protein scaffold has been widely thought to chaperone the cyclization cascade by merely conformational control (22, 26). Very recently, a hypothesis pointing to a possible involvement of electrostatic effects during the carbocation cascade of other terpene synthases has been proposed. These effects are thought to be mediated by the PP_i anion coproduct that may be retained in the active site (21, 24, 26–32).

A crystal structure of the open, catalytically inactive TXS conformation containing an unproductive fluorinated substrate analog (2-F-GGPP) in the active site is available [Protein Data Bank (PDB) ID code 3P5R] (8, 15). This dataset, however, only provides limited information on the dynamic carbocation processes involved in the cyclization cascade. At present, none of the reported mechanistic or structural studies on class I terpene synthases do consider detailed structure–function relationships of the catalytically active enzyme conformation. This situation arises as the use of substrate analogs and crystal structures of open conformations do not reflect the concerted, dynamic events inside the closed, catalytically active enzyme complex. In fact, the closed enzyme complex containing a native GGPP substrate has

Significance

Class I terpene synthases are essential in biosynthesis of bioactive terpenoids (e.g., Taxol). Identification of structure–function correlations is hampered by highly dynamic carbocation-driven reactions and the limited availability of catalytically relevant crystal structures. We provide the closed, catalytically active conformations of taxadiene synthase (TXS) and various taxonomically unrelated class I terpene synthases in complex with their catalytically relevant carbocationic intermediates. Our methodology allows direct prediction and validation of universal structural features that govern the highly dynamic catalytic processes of class I terpene synthases. Our data enabled delineation of a discrete reaction pathway for TXS and allowed targeted enzyme engineering to generate alternate macrocyclic core structures, which can act as synthons for bioactive lead structures.

Author contributions: P.S. and T.B. conceived the study; P.S., F.W., W.E., V.S., R.K., and T.B. designed the experimental setup; P.S., A.B., M.H., and R.K. performed research; P.S., C.G., J.v.R., F.W., W.E., and R.K. analyzed data; P.S., W.E., V.S., R.K., and T.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. D.H.S. is a guest editor invited by the Editorial Board.

[1]To whom correspondence should be addressed. Email: brueck@tum.de.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1519680113/-/DCSupplemental.

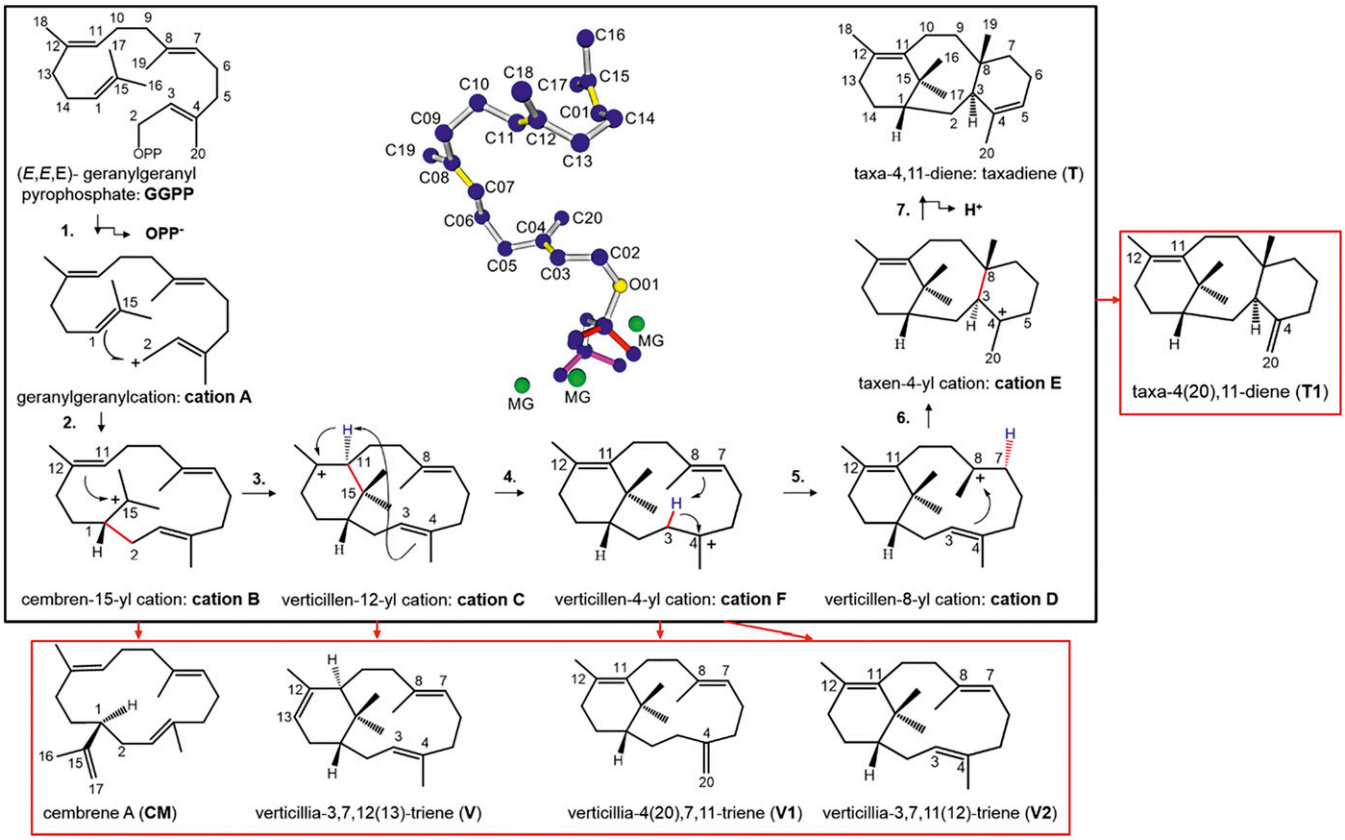

Fig. 1. Cascade pathway, carbocation rearrangements, and deprotonated intermediates. The black box depicts the postulated carbocationic intermediates during the cascade pathway from GGPP→**T** and the numbering of the carbocation cascade steps. The red boxes depict the found deprotonated intermediates, whereby the red arrows indicate the cationic intermediate from which the deprotonated intermediate is derived. The structure in the middle depicts the conformation of productive GGPP (blue, double bonds in yellow) and the localization of the trinuclear Mg^{2+} cluster (green) in the active site of the closed TXS conformation. Note that numbering of cationic and deprotonated intermediates differs from previous reports (9, 11, 12, 15).

not been considered to decipher native structure–function correlations and catalytic mechanisms of terpene synthases.

In this study, we constructed a closed, active TXS conformation harboring native GGPP in its productive conformation. To elucidate the catalytically relevant complexes, reflecting the carbocation cascade in the active site as well as structure–function relationships, we further applied a molecular mechanic (MM)-based modeling approach combined with site-directed mutagenesis-guided experimental verification. The data provide new, unprecedented insights into the control of the cyclization mechanism and the product promiscuity of class I terpene synthases. Application of in silico-guided protein engineering enabled selective attenuation of the carbocation cascade at several intermediate steps. The methodology allowed experimental confirmation of native GGPP and carbocation structures and delineation of a unified reaction pathway. The targeted quenching gave rise to monocyclic and bicyclic diterpene macrocycles that represent the hydrocarbon basis of a myriad of known bioactive compounds. Beyond the TXS model, we have examined the here-modeled closed conformation of the bacterial diterpene synthase cyclooctat-9-en-7-ol synthase (CotB2) in detail and examined several other class I terpene synthases. Therefore, the applied methodology combined with the mechanistic insights into TXS structure–function relationships enables the identification of universal structural motifs, productive GGPP folding patterns, localizations of carbocationic intermediates in the active site, prime candidates for the active-site bases, and mechanisms that cause product formation and promiscuity in any other class I terpene synthase.

Results

TXS•GGPP Complex: Evolutionarily Conserved Motifs Initiate the Carbocation Cascade. The reported crystal structure of TXS, representing the open, catalytically inactive enzyme, contains 2-F-GGPP in an unproductive orientation leading to subsequent intermediates in which several stereocenters are inverted (Fig. 2*A*) (15). Using molecular mechanics, we constructed a model of the closed TXS•GGPP complex. This complex is considered to be the catalytically active conformation of the enzyme and represents the initial step of the catalytic carbocation cascade (Fig. 2*B* and *SI Appendix*, Fig. S1*A*). The model omits the 80-residue N-terminal transit sequence. To provide for correct substrate folding, we manually docked a geometry-optimized GGPP structure, derived from gas phase QM calculations of cation A into the TXS active site. This ensures that T can be produced in the correct configuration. Examination of the closed TXS•GGPP complex indicated an extended hydrogen bond network of bound GGPP compared with the open conformation. In addition to the reported hydrogen bonds formed by R754, R768, and Y835 (15), PP_i accepts water-mediated hydrogen (H) bonds extending from R580 and N-terminal Y89 in the closed complex (Fig. 2*B*). Water- and amino acid-mediated bonding of productive GGPP resembles that of the closed conformation of bornyl diphosphate synthase (BPPS), a monoterpene synthase (*SI Appendix*, Fig. S1 *B–D*) (26). In closed complexes of TXS and BBPS, PP_i accepts water-mediated H bonds from a network of H-bond donors composed of Y89/60 (TXS/BPPS), R580/316, R768/507, R754/493, and Y835/572 and water molecules. Structural alignments of TXS with prokaryotic and eukaryotic monoterpene, sesquiterpene, and

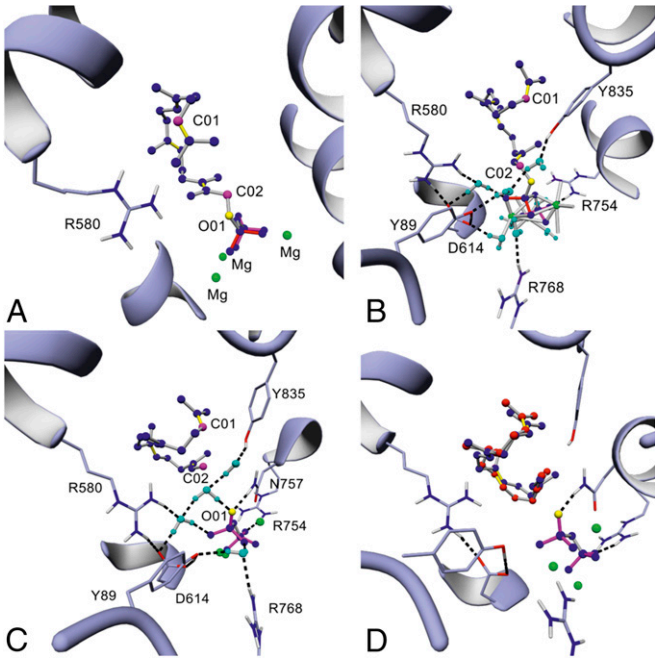

Fig. 2. Closed- and open-conformation TXS model, productive cation A complex. (*A*) TXS harboring the cocrystallized unproductive 2-F-GGPP (8) (PDB ID code 3RP5) (C01, C02: magenta; O01 of PP_i: yellow; Mg^{2+}: green). (*B*) TXS harboring manually built productive GGPP and the closed-conformation specific H-bond donor network: Y89, R580, R754, R768, and Y835. Additionally, D614 and N757 are labeled. Coordinated waters: cyan; H-bonds: black; Mg^{2+}-coordinating dative bonds: gray lines. (*C*) TXS harboring docked QM-derived productive cation A. (*D*) Superposition of TXS harboring docked cation A (blue) with TXS harboring manually built cation A (red).

diterpene synthases demonstrated a common H-bond network (*SI Appendix*, Fig. S1 *D* and *E*) (27, 29, 33–41). This suggests that a conserved basic amino acid-rich triad sets and holds PP_i in place during the complete catalytic cascade. Based on these data, we suggest to expand the proposed PP_i-triggered active-site closure and charge stabilization mechanism onto the whole family of terpene synthases. Most notably, the R580-PP_i motif resulting from active-site closure is a promising candidate for the active site base in TXS (see below). These data are consistent with the observation in the *Streptomyces pristinaespiralis*-derived selinadiene synthase, where single arginine substitutions of the corresponding network lead to inactive enzyme variants (37). Furthermore, in TXS, a conserved amino acid-mediated H-bond network including both Y residues seems to be mandatory to bind bulk water remaining in substrate vicinity after active-site closure, and thus prevent premature carbocation quenching (Fig. 2*B*). This notion is supported by analyzing the supplementary modeled closed complex of cembratriene-ol synthase (CBTS), a class I diterpene synthase from *Nicotiana tabacum*, which produces two epimeric alcohols and contains the three-arginine network, but lacks both tyrosine residues (*SI Appendix*, Fig. S1 *F* and *G*) (41). Additional analysis also substantiates the recently suggested induced-fit mechanism leading to ionization via an effector triad comprising the PP_i sensor R754 on helix H, the linker S713, and the effector V714-O in the G1/2 helix-break motif (*SI Appendix*, Fig. S2) (37). Here, we experimentally verified the necessity of the H-bond donor amino acid network as well as the linker and effector residues via single-disruptive amino acids (aliphatic, aromatic, and charged residues) substitutions, resulting in catalytically inactive or significantly impaired TXS variants (*SI Appendix*, Table S4).

Acyclic Cation A: PP_i Stabilization and Conformational Control Imposed by the Active Site. TXS initiates catalysis by Mg^{2+}-mediated GGPP ionization forming the transient, acyclic carbocation A and PP_i as a coproduct (Figs. 1 and 2 *C* and *D*). Initial molecular docking experiments indicate that cation A adopts a productive geometry in the closed TXS active site before cyclization to cation B. The docked QM-derived cation A exists in two distinct conformations. With respect to the TXS active-site localization of unionized GGPP, only one conformational cluster represents a catalytically relevant cation A structure (Fig. 2*C* and *SI Appendix*, Fig. S3*A* and Table S3). Interestingly, the manually built cation A spontaneously folds to a structure that is highly similar to that of the catalytically relevant cation A, strongly supporting the MM-based manual building approach (Fig. 2*D*). As a negative control, the cyclization mechanism of a GGPP was examined that adopts the cocrystallized 2-F-GGPP conformation (8, 15). Energy minimizations yielded a verticillen-12-yl cation with an inverted configuration of C11 that prevents the productive cyclization to T (*SI Appendix*, Fig. S3*B*). This strongly supports the model that terpene synthases mainly dictate reaction outcome by sterically restricting their substrate and discrete reaction intermediates to a subset of possible productive conformations (template model) (22, 24). Moreover, it demonstrates differing binding modes followed by nonnative cyclization characteristics for substrate analogs that hamper the correct prediction of important structure–function relationships in class I terpene synthases.

Monocyclic Cation B: Native Existence of a Monocyclic Intermediate in the Carbocation Cascade. The second step in the carbocation cascade is the conversion of the acyclic cation A to a monocyclic cation B (Fig. 1), resembling the general first cyclization step in class I terpene synthase catalysis. Using the TXS complex harboring the docked QM-derived cation A template, the C1–C2 bond was manually built followed by energy minimization. The resulting monocyclic cation B is structurally similar to the QM-derived transition state (TS) between cations B and C rather than to the QM-derived cation B (Fig. 3 *A* and *B*) (15).

The active site of the enzyme left enough space for rotation of the C1–C15 bond that is spanned between W753 and the aromatic network consisting of Y835, F834, and Y841. During binding of GGPP, the C15–C17 isopropyl group folds into this space as a preorganizational event. Docking of QM-derived TS B-C into the corresponding transient TXS complex further corroborates its fitting into the active site and specifies its conformation as distinctly derived from the catalytically relevant docked cation A structure (Fig. 3*B* and *SI Appendix*, Fig. S6 and Table S3).

Although it has previously been speculated that TXS mutants could form cembranoid-type isomers by deprotonation of cation B (11), use of the closed TXS conformation model harboring the monocyclic cation enabled identification of potential "hot-spot" amino acid residues for production of monocyclic cembranoids in contrast to an open conformation structure. Analysis of the TXS•cation B complex suggests that S587, C830, Y835, and Y841 represent polar groups in proximity to cationic C15 and C16/C17 methyl-hydrogens. However, none of them can act as a deprotonating base as the carbocation cascade does not stop at cation B in native TXS.

The experimental data demonstrate that Y835A/W substitutions result in inactive TXS variants. Additionally, C830S/A substitutions had no effect on product spectra and yield (*SI Appendix*, Table S4). The model further shows that W753 is directly adjacent to the C16/17 methyl groups of cation B (Fig. 3*A*). Therefore, we suggested that a substitution to histidine may act as a proton acceptor, resulting in cembranoid formation by premature deprotonation of cation B (Fig. 3*C*). Experimental in vitro and in vivo data demonstrate that substitution of TXS-W753H, indeed, converts GGPP to a monocyclic product with no

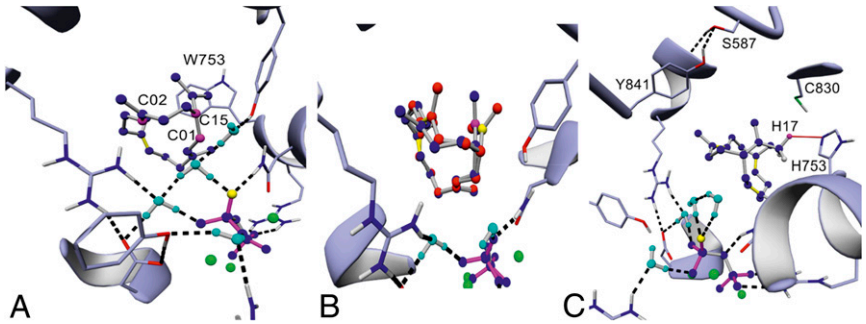

Fig. 3. Wild-type and mutant cation B complexes. (*A*) TXS harboring manually built cation B, which resembles the QM-derived TS B-C (C01, C02, C11, and C15: magenta). (*B*) Structural overlay of TXS harboring the manually built cation B (blue; C15: magenta) with docked QM-derived TS B-C (red; C15: yellow). (*C*) TXS-W753H harboring manually built cation B. Deprotonation of C17 (closest H17: magenta) results in formation of monocyclic CM. The shielded dipolic network (S587, Y841) that does not interfere with the cyclization cascade due to hydrogen bonding and spatial separation is further shown.

further cyclization to cation C (Table 1 and *SI Appendix*, Fig. S8*A*). Consequently, we could assign the monocyclic W753H product to (−)-(*R*)-cembrene A (CM) (*SI Appendix*). Alternatively, H753 may act in deprotonating proximal C830, which leads to a reactive anionic sulfur acting as deprotonating base comparable to the catalytic diad mechanism in cysteine proteases (Fig. 3*C*) (42). The double-mutant W753H/C830A yields the same product distribution as W753H. Thus, it seems likely that the histidine residue acts as a deprotonating base. Additional substitutions of W753L/V/E/C/A resulted in inactive TXS variants, demonstrating the necessity for an aromatic residue at this position in the wild-type enzyme (see below) (*SI Appendix*, Table S4).

Bicyclic and Tricyclic Cations C–E: Steric and Electrostatic Control of the Carbocation Cascade. The next step in the catalytic cascade involves formation of a bicyclic verticillen-12-yl cation (cation C) (Fig. 1) (10, 12, 15). Analogous to the process applied to cation B above, the TXS•cation C complex was modeled using the previous TXS•cation B complex as a template. Interestingly, a verticillen-12-yl cation formed spontaneously in TXS with an equivalent structure to the QM-derived cation C (Fig. 4*A*) (15). To our knowledge, this is the first experimental confirmation of this intermediate using the native substrate.

Furthermore, docking studies confirmed the localization of cation C as distinctly derived from the catalytically relevant TS B-C structure (*SI Appendix*, Figs. S3*C* and S6, and Table S3). Subsequent analysis of this complex indicates that residues V584 and V610 are close to the positively charged C12 atom of cation C. Although V610H/S/F/A mutations resulted in inactive TXS variants, V584M/L yielded a product mixture in in vitro experiments including T (13.8%) and three bicyclic verticillene-type structures (84%, 1.4%, and 0.8%) (Table 1 and *SI Appendix*, Fig. S8*A*). Notably, we detected the same product profile in in vitro experiments using native TXS, albeit with different product variant percentages (Table 1 and *SI Appendix*, Fig. S8*A*). In vivo experiments followed by detailed NMR analysis allowed assignment of the major verticillene-type structure (84%) to (+)-(*R*,*R*)-verticilla- 3,7,12(13)-triene (V) (*SI Appendix*). The absolute configuration of V was derived by comparison of its CD spectrum with the spectra of the 7-fluoro-verticillene analogs (*SI Appendix*) (10).

To unravel the deprotonation events leading to V, we conducted in silico comparisons of TXS-V584M/L•cation C complexes with its corresponding wild-type equivalent. Although cations A and B exhibit identical shapes in the TXS–V584M/L complex, cation C is localized closer to the R580-PP_i moiety due to steric restraints imposed by M and L. As a consequence, C13 hydrogens are shifted 0.8 Å toward the R580-PP_i motif in comparison with native TXS, which likely mediates premature cation C deprotonation (Fig. 4*B* and *SI Appendix*, Fig. S3*D*). In addition to V, we examined the formation of the minor bicyclic products V1 and V2 using MS-based analyses. We also observed these products in spectra affiliated with TXS [V1 (0.2%), V2 (1.1%)] (Table 1 and *SI Appendix*, Fig. S8*A*). Interestingly, the V1 MS spectrum is identical to the spectrum reported for verticillia-4(20),7,11-triene extracted from *Boswellia carterii* (*SI Appendix*) (43). This analogy not only confirms the presence of cation F but also designates V1 as its intermediate quenching product. Our confirmation of cation F linked to the observation of V and V1 now allows a tentative assignment of the V2 mass spectrum to the formation of verticillia-3,7,11(12)-triene (*SI Appendix*), as in silico studies indicated that V1 and V2 formation could be induced through variable, time-resolved positioning of cation F (Fig. 1), accompanied by deprotonation of C3 or C20, respectively. Due to the positioning of cation F in TXS, deprotonation by the bifunctional R580-PP_i motif seems likely (*SI Appendix*, Fig. S3*E*).

Single amino acid switches, able to "short-circuit" the cyclization cascade by electrostatic stabilization, have previously been reported for other terpene synthases (44–46). Analysis of the closed complex indicates that the TXS dipole network is shielded or spatially too far removed from cationic intermediates, thereby preventing a direct electrostatic interaction with the cyclization cascade. The TXS model further suggests that S587 and Y841, which are spatially connected via an OH-mediated H bond, are in proximity to the cationic positive charge (Fig. 3*C*). Interestingly, in silico-guided substitution of S587A and Y841F resulted in deprotonated intermediates (see below). Subsequent in vitro

Table 1. Product distributions of wild-type TXS and important variants

Target	Mutation	Activity*	T†	T1	CM	V	V1	V2
TXS		100.0	93.2	4.7	N.D.	0.8	0.2	1.1
V584	V584M	92.3	13.8	0.6	N.D.	83.4	0.8	1.4
V584	V584L	92.1	13.8	0.6	N.D.	83.4	0.8	1.4
S587	S587A	21.8	8.9	N.D.	58.9	32.2	N.D.	N.D.
W753	W753H	51.3	N.D.	N.D.	100.0	N.D.	N.D.	N.D.
Y841	Y841F	41.3	N.D.	N.D.	56.4	43.6	N.D.	N.D.

*Enzyme activities of mutants (in percentage) in comparison with TXS (=100%). The activities were determined by peak integration of the enzyme products in relation to the peak area of the internal standard α-humulene (*Materials and Methods*).

†Compound numbers CM–T1 refer to compound numbers in Fig. 1 and *SI Appendix*, Fig. S8*A*. The tabulated values represent the product distributions in percentage for TXS and each mutant (sum = 100%). N.D., not detectable.

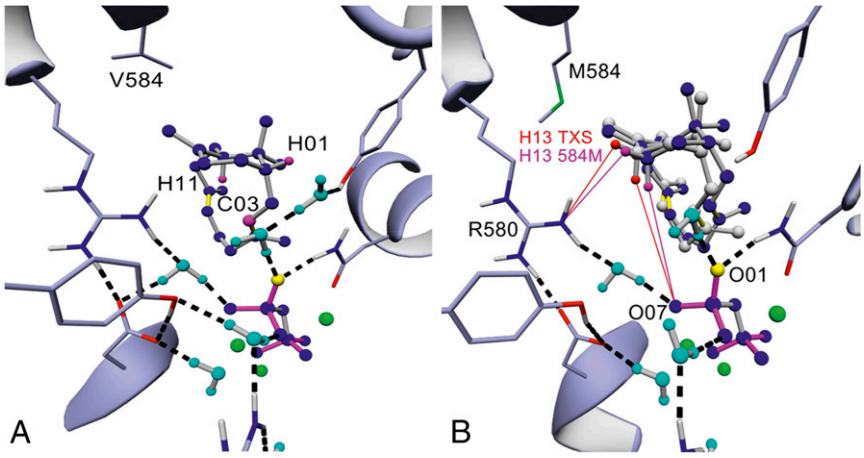

Fig. 4. Wild-type and mutant cation C complexes. (*A*) TXS harboring manually built cation C (H01equatorial, H11axial, and C03: magenta). (*B*) Superposition of TXS-V584M harboring manually built cation C (blue) with manually built cation C of TXS (gray). C13 hydrogens abstracted during R580-PP_i–assisted deprotonation resulting in formation of V are shown (C13 hydrogens of TXS in red; C13 hydrogens of TXS-V584M in magenta). Distance between H13equatorial (H13e) of TXS-V584M and NH2 of R580 is 3.05 Å (magenta line); distance between H13e of TXS and NH2 of R580 is 3.87 Å (red line). Distance between H13axial (H13a) of TXS-V584M and O07 of PP_i is 4.25 Å; distance between H13a of TXS and O07 of PP_i is 4.83 Å.

and in vivo experiments allowed structural assignment of Y841F reaction products as CM and V, indicating complete catalysis termination at cation C (Table 1 and *SI Appendix*, Fig. S8*A*). Complementary, the S587A mutant resulted in formation of CM and V as well as minor amounts of T (Table 1 and *SI Appendix*, Fig. S8*A*).

Carbocation Cascade Mechanism: Steering Effect of R580-PP_i Determines Promiscuity. Formation of tricyclic **T** proceeds from cation C via F, D, and E in the gas phase (Fig. 1), although the direct route of cation C→D→E cannot be excluded (15). Transition to cation D, which adopts two different conformations D1 (cation D1) and D2 (cation D2) in the gas phase represents the fifth step in the carbocation cascade (15). MM-based modeling and docking of cations F, D1, D2, and E into TXS demonstrate that each adopts a conformation that is distinctly derived from the respective former catalytically relevant cascade cation structure (Fig. 5*A* and *SI Appendix*, Figs. S3 *E–G* and S6, and Table S3). Complex models of TXS•cations D1 and D2 indicate that the transition of cation D1→D2, i.e., the movement of the D1-C8 methyl group toward an aligned position with the C4 methyl group in –D2, is induced by an electrostatic effect of R580-PP_i. Direct spatial proximity of negatively charged O atoms of R580-PP_i and Δ3,4 of the substrate suggest electrostatic repulsion forces resulting in D1→D2 transition (*SI Appendix*, Fig. S3*F*). Cumulative analyses of the TXS•cation complexes and observed product distributions indicate a pronounced electrostatic steering effect of R580-PP_i as the cascade enters the bicyclic carbocation stage.

In contrast to TXS•cations C–E complexes, TXS•cation A and B complexes reveal a high degree of amino acid-assisted stabilization. Specifically, π–π interactions between the aromatic residues W753, F834, Y835, and Y841 and the substrate as well as cation–π interactions between the positive charge and these residues stabilize early- but not late-stage cations in TXS, which is in line with previous reports of other terpene synthases (*SI Appendix*, Fig. S4) (25, 47, 48) and further corroborated by docking cluster analyses (*SI Appendix*, Table S3). Furthermore, product distributions of V584M/L and S587A mutants show ongoing cyclization to T compared with W753H and Y841F (Table 1). This, in turn, suggests that the reduced π-interaction in TXS•cations C–E complexes are not able to counteract a continuous R580-PP_i–induced electrostatic attraction that promotes the cascade to cation E. This steering effect further corroborates the experimentally observed native existence of cation F, as cation C→F transition would bring the positive charge into proximity to R580-PP_i (*SI Appendix*, Fig. S3*E*) (see above). This implicates that this transition is thermodynamically favored due to ion-pairing effects in contrast to the gas phase (15).

However, as the cascade does not stop at cation F, other effects must counteract those exerted by R580-PP_i and provide the driving force for reaction to T. The TXS•cation F complex demonstrates that R580-PP_i and C4 hydrogens of cation F are still too far removed for a complete terminal deprotonation at this stage in contrast to the TXS•cation E complex (Fig. 5*B*), which allows the native hydride shift resulting in cation F→D1 transition (*SI Appendix*, Fig. S3*E*). This indicates that kinetic control, i.e., spatial positioning and movement of cations restricted by the active site, leads to the thermodynamically not favored cation F→D1 transition, as the positive charge is located distal to the R580-PP_i in cations D1 and D2 compared with cation F (*SI Appendix*, Fig. S3 *E–G*). In line with these data, it seems feasible that R580-PP_i–steered positive charge migration toward itself induces the conformational change of D1→D2 and the final transition of cation D2→E due to the observed lack of π-stabilization. Concordantly, the TXS•cation E complex shows the global distance minimum of the positive charge as well as H5 and C20 hydrogens with respect to R580-PP_i and exhibits no substrate-induced electrostatic forces that counteract proton abstraction (Fig. 5*B*). This, in turn, indicates that cation E is the global energy minimum of the reaction in relation to cation A, contrary to QM gas phase calculations (15). Docking cluster analyses further corroborate this indication (*SI Appendix*, Table S3).

The model further indicates that R580-PP_i, in addition to its oversized active-site cavity (8), may also be responsible for the enzyme's product promiscuity. Although initial excess of spatial freedom allows GGPP misfolding, spatial freedom is shown to lead to predominant carbocation tumbling and imprecise barrier crossings at later stages of catalysis resulting in alternative product formation (21, 22). Therefore, tumbling or imprecise barrier crossing in an oversized active site cavity may also be the reason for alternative products observed during the catalytic cascade, as alternative products in TXS and mutants seem exclusively to be derived from later catalytic stages, i.e., cations C–E (Fig. 1 and Table 1). These effects in combination with reduced π-interactions in cations C–E and R580-PP_i–induced effects, can thus explain premature R580-PP_i–mediated deprotonation in cation F

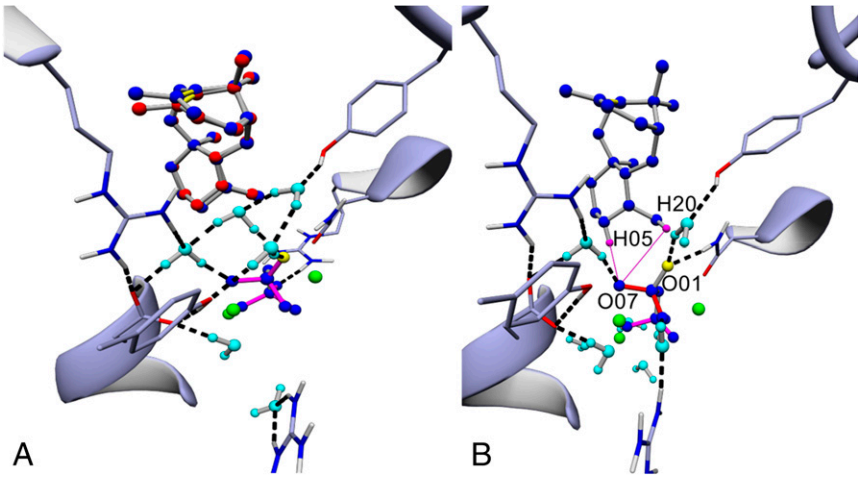

Fig. 5. Molecular docking of QM-derived cation E, product promiscuity in cation E. (*A*) Superposition of TXS harboring manually built cation E (blue) with the docked QM gas phase-derived cation E (red). (*B*) TXS harboring cation E. Differing distances of H05 equatorial (deprotonation results in formation of T) and closest H20 (deprotonation results in formation of T1) with respect to O07 of R580-PP_i are shown as magenta lines. Distance of equatorial H5 of the newly formed cyclohexane ring (twist-boat conformation) to O07 of R580-PP_i is 2.75 Å; distance of closest H20 to O07 is 3.35 Å. The distance between the positively charged C atom with respect to R580-PP_i reaches its global minimum of the cascade at formation of cation E (3.71 Å).

resulting in formation of V1 and V2 in TXS and V584M/L mutants (Table 1 and *SI Appendix*, Fig. S8*A*). This view is supported by the adjacent positioning of R580-PP_i with respect to the C3 and C20 hydrogens (*SI Appendix*, Fig. S3 *E* and *H*). Experimentally, this is further corroborated by the active site enlarging F834G substitution, which results in various verticillene and taxadiene isomers derived from cations C, F, and E, but no monocyclic structure derived from cation B (*SI Appendix*, Fig. S8*A* and Table S4). By contrast, multiple product formation derived from cation E in TXS [T (93.2%), T1 (4.7%)] cannot be explained by tumbling or imprecise barrier crossing alone but may also involve R580-PP_i–induced electrostatic effects. Therefore, we suggest a simple explanation for that observed product distribution. In fact, differing relative distances of R580-PP_i with respect to cation E H5β or C20 methyl hydrogens, seem to lead to selective deprotonation events that govern the formation and distribution of T or T1, respectively (Fig. 5*B*).

Biotechnological Implications: Production of Cembranoid- and Verticillene-Type Synthons. We successfully used our in vivo terpene production system in 30-L–scale fermentations, yielding CM (W753H) and V (V584M) at concentrations of 8 and 11 $mg{\cdot}L^{-1}$, respectively, with glycerol as carbon source (*SI Appendix*, Fig. S10 *A* and *B*). To our knowledge, this is the first time that these compounds were produced in recombinant microbial production systems. Although cembranoid synthases have been reported (49, 50), to our knowledge, this is the first report of a terpene synthase capable of producing verticillene-type skeletons. The sustainable fermentative microbial production of the macrocycle can be combined with further chemical or enzymatic derivations. It therefore represents a decisive step on the way to sustainable synthesis of bioactive cembranoid- and verticillene-type diterpenoids (16–20).

Discussion

Studying the dynamic carbocation cascades and protein structure–function correlations involved in terpene synthase catalysis is challenging. In this study, we introduce an MM-based computational methodology that allows the elucidation of important structure–function relationships in terpene synthases and the identification of essential amino acids, whose substitution lead to alternate terpene macrocycles. Moreover, a catalytically relevant, closed complex model of a class I terpene synthase containing a productive GGPP is introduced that can serve as a model for any other eukaryotic class I terpene synthase comprising a functional class I associated α-domain (1). GGPP can easily be substituted by geranyl (GPP) or farnesyl (FPP) pyrophosphate (Fig. 6).

We successfully constructed MM-based models of the catalytically relevant, closed TXS complex containing each reaction intermediate. The predictive power of this in silico methodology was corroborated by experimental site-directed mutagenesis data and verified by docking of QM gas phase intermediates. The methodologies described herein can be applied to identify catalytically important residues and localization of carbocationic intermediates for crystallized open-complex models, closed complexes without productive substrate, and high-class homology models of prokaryotic and eukaryotic class I terpene synthases. We have verified our in silico tool kit using the prokaryotic diterpene synthase CotB2, as its proposed carbocationic intermediates and the open-form crystal structure have recently been reported (*SI Appendix*, Fig. S7*A*) (36, 51, 52). Using MM-based modeling, we were able to elucidate the localizations of productive GGPP and the cationic intermediates in the active site as well as get insights into the structural basis for the two 1,5-hydride shifts and the carbon rearrangement proposed by NMR-based labeling experiments and QM gas phase calculations (Fig. 7 *A* and *B*, and *SI Appendix*, Fig. S7 *B*–*F*) (51, 52). Analysis of the cyclization trajectories of closed CotB2 also enabled the elucidation of the molecular basis for the generated alternate macrocycles yielded upon mutagenesis that have recently been reported (*SI Appendix*, Fig. S7 *G* and *H*) (36, 53). The closed, catalytically active CotB2•GGPP complex could therefore serve as a productive, catalytically relevant template model for any other microbial class I terpene synthase comprising the class I-associated α-domain (Fig. 6).

The identification of a universal, H-bond donor amino acid network in class I terpene synthases independent of their taxonomic origin and function strongly argues for common mechanistic features particularly in the initial stages of catalysis. PP_i charge stabilization in conjunction with carbocation shielding from remaining solvent water through an H-bond donor amino acid network constitutes a common catalytic trigger upon active-site closure. Requirements for charge stabilization and separation argue that ionized PP_i is retained in the active site during the entire catalytic cascade suggested by previous reports (21, 26–32). The lack of essential Y residues results in hydroxylated products as in the case of CBTS, which suggests an amino

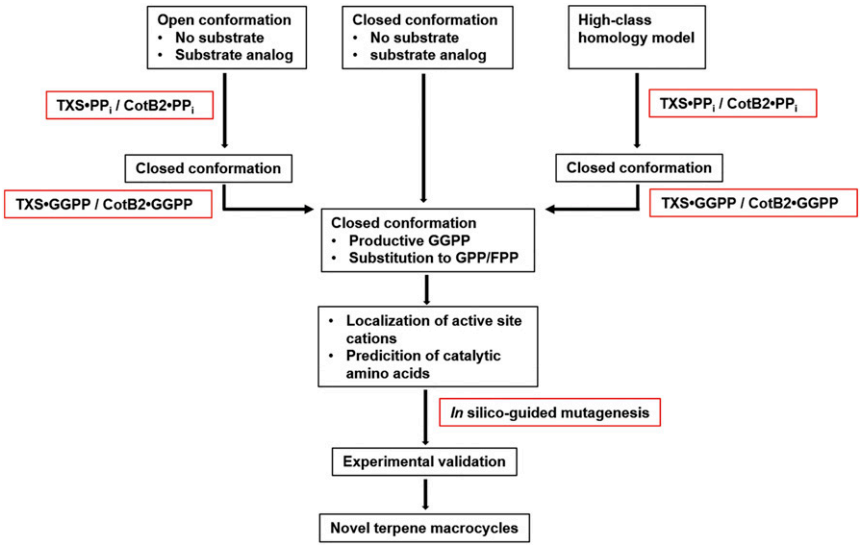

Fig. 6. Workflow for the generation of catalytically relevant class I terpene synthase conformations. The workflow that can be used to generate catalytically relevant, closed conformations of class I terpene synthase comprising a class I-associated α-domain is shown. (The detailed modeling procedure is depicted in *Materials and Methods*.) Subsequent manual formation of respective catalytic cascade steps enables the prediction of important catalytic residues in the active site.

acid-assisted mechanism that provide control for water-assisted quenching products. We could further demonstrate by molecular-docking calculations that, in the initial reaction phase, the acyclic cation A can only assume a single conformation, which strongly argues for the template model (22). The data confirm the existence of a monocyclic cation B, specifically as in TXS-W753H deprotonation of C17 results in formation of CM. This indicates a concerted, asynchronous cyclization reaction progressing via GGPP→cation A→B→C, involving cations A and B as distinct short-lived intermediates. Our trajectories suggest the route cation C→F→D as energetically more favorable compared with the direct route cation C→D. This is demonstrated by proximity between R580-PP_i and the positive charge in cation F in the closed-complex model. Together with the finding of V1 and presumably V2 in chromatograms affiliated with TXS and V584M/L, we present direct experimental evidence for the native existence of cation F.

In silico-guided mutagenesis enabled identification of steric effectors that allow formation of bicyclic verticillenes. These deprotonation events seem to be assisted by the bifunctional R580-PP_i motif. Complementary, in silico studies with 7-fluoro-GGPP as substrate corroborated this hypothesis, which is in line with previous detection of an *exo*-7-fluoroverticillene and other 7-fluoroverticillene–type structures (*SI Appendix*, Fig. S5*B*) (10). Electrostatic stabilization of the positive charge in cations C and F induced by the adjacent 7-fluoro ligand permits perturbation of the fine-tuned kinetics in transition to cations F and D, respectively, whereas R580-PP_i directs the carbocation toward itself, resulting in subsequent deprotonation (*SI Appendix*, Fig. S5*A*). Therefore, an R580-PP_i–mediated deprotonation event in conjunction with carbocation tumbling due to reduced π-interactions is able to explain formation of V observed in the product profile of TXS (Table 1). In addition to steric effectors modulating TXS product formation, electrostatic dipole-mediated effects lead to alternative macrocycles. Based on in silico studies, we generated S587A and Y841F mutants, thereby perturbing the shielded active-site dipole network. These mutant complexes indicated again that disruption of the fine-tuned, R580-PP_i–coordinated cyclization timeline is able to result in premature deprotonation of cations C and B by R580-PP_i (Table 1 and *SI Appendix*, Fig. S3 *I* and *J*).

Interestingly, incomplete deprotonation and observed ongoing cyclization to cation E in V584M/L and S587A indicate that cations C–E exist in a well-defined equilibrium, with cation C comprising the global energy minimum and formation of T being promoted by site-specific deprotonation as previously suggested (15). An alternative hypothesis would suggest cation E being the global minimum as well as the cyclization of cations A–E as electrostatically governed by R580-PP_i and assisted by steric control imposed by the active site.

Our cumulative data support the alternative hypothesis that is generally applicable upon elucidation of class I terpene synthase catalysis. The results suggest that (*i*) formation of olefinic macrocycles in terpene synthases is a sequential event, which in TXS comprises formation of monocyclic, bicyclic, and tricyclic intermediates. (*ii*) After GGPP hydrolysis, the initial intermediate (cation A) and PP_i are spatially separated and stabilized in the hydrophobic active-site cavity (25, 47). (*iii*) Acyclic cation A and the monocyclic secondary intermediate are stabilized by π–π and cation–π interactions imposed by the active site. (*iv*) Amino

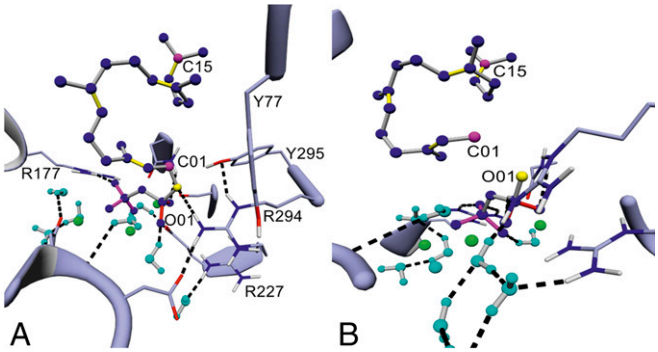

Fig. 7. Closed complex of CotB2 harboring productive GGPP and cation A. (*A*) Closed-complex conformation of CotB2 harboring productive GGPP and the identified H-bond donor network consisting of Y77, R177, R227, R297, and Y295 (C01 and C15 in magenta, O01 in yellow, Mg^{2+} ions in green, and coordinated water molecules in cyan). (*B*) Closed-complex conformation of CotB2 harboring cation A (*SI Appendix*, Fig. S7*C*).

acid-assisted stabilization in early stages of catalysis prevents carbocation tumbling and imprecise barrier crossings, thereby prohibiting premature deprotonation. (*v*) Upon progression of catalysis to the bicyclic cation C stage, π-interaction is reduced, which results in increased cation tumbling and imprecise barrier crossing while the R580-PP_i mediated steering effect becomes significant. (*vi*) This steering and cation-attracting role of R580-PP_i leads to direct premature deprotonation in cations C–F. (*vii*) It further leads to direct or indirect premature deprotonation when the fine-tuned and electrostatically balanced timeline of cyclization and hydride shifts is perturbed. (*viii*) Instead, in the absence of counteracting amino acid-assisted electrostatic effects, steric restriction imposed by the active site and steering of R580-PP_i provide for cyclization from cation A→E, which corroborates the template model and dominating kinetic instead of thermodynamic control (24, 54). (*ix*) Due to proximity of the positively charged carbocation to R580-PP_i and in agreement with docking cluster analyses, cations F and E comprise lower energy levels in TXS compared with cation C in relation to cation A. (*x*) The R580-PP_i bifunctional motif, only present after active-site closure and arguably characterized by an inverse protonation equilibrium, appears to be the prime candidate for the active-site base in TXS and other terpene synthases examined in this study.

Materials and Methods

Quantum Chemical Calculations (QM). All calculations were performed with Gaussian 03 (55). All cationic intermediates A–E and GGPP in productive conformation were optimized using the B3LYP/6-31+G(d,p) method as described previously (15).

Model of TXS in Closed Conformation Harboring Productive GGPP. Molecular-dynamics (MD) simulation and energy minimization steps were performed in a periodic neutralized TIP3P water box at pH 6.8 of YASARA Structure (version 14.12.2; YASARA Biosciences) using the second-generation self-parametrizing YASARA2 force field with default settings (56). Assigned hydrogens, bond orders, and charges for GGPP and cations A–E were obtained using the AutoSmiles force field parameter assignment that is part of YASARA Structure (57, 58). Analysis of simulation and energy-minimized trajectories were performed using YASARA Structure.

Preparation of Closed Conformation of TXS. Modeling of the N-terminal residues (81–110) and the J–K loop (residues 834–849) in closed conformation was performed with BPPS in complex with three Mg^{2+} ions and 3-azageranyl diphosphate (PDB ID code 1N20) as template (26) and modeled as described previously (8), followed by 10,000 steps of energy minimization and a 10-ns MD simulation in water using YASARA Structure. During energy minimization, all residues, heavy atoms, and 2-FGGPP were unconstrained. Nonbonded cutoff were set to 7.5 Å. Particle mesh Ewald method was used for long-range electrostatic forces (59).

Integration of GGPP in Productive Conformation. The geometry-optimized GGPP in productive conformation obtained by B3LYP/6-31+G(d,p), as described previously (15), was superposed onto the 2-FGGPP in the modeled closed conformation (see 2.1), the 2-FGGPP deleted, and the bond between O1 and C1 linked manually. The obtained complex was subsequently subjected to 10,000 steps of energy minimization and a 10-ns MD simulation in water using YASARA Structure.

Molecular Docking of QM-Derived Cations A–E. Molecular docking of QM-derived cations A–E was performed using the AutoDockVina program environment of YASARA Structure (60). Cations were docked into a simulation cell (size: *X* size = 14 Å, *Y* size = 14 Å, *Z* size = 14 Å; angles: $\alpha = 90°$, $\beta = 90°$, $\gamma = 90°$) around the following eight residues: Y89, E583, F602, F612, S752, L827, A764, and Y836. For each cation, 999 docking runs were performed while all atoms of the corresponding cations were set as rigid. Cluster analysis was performed in the AutoDockVina program environment, and clusters were characterized by binding energy (in kilocalories per mole), dissociation constant (in picomolar concentration), and binding energy spread (in kilocalories per mole) (*SI Appendix*, Table S3). Two enzyme–ligand complexes belonged to different clusters if the ligand root-mean-square deviation (RMSD) was larger than 2.0 Å.

Molecular Modeling of Cations in the Active Site of TXS. An approach of force field-based MM modeling combined with molecular docking to derive localization of cations A–E was applied. First, using the AutoDockVina program environment of YASARA Structure, the QM cation A was docked into the active site of the closed conformation bearing the retained diphosphate group bound to the trinuclear magnesium cluster and held in place by the established H-bond donor amino acid network. From this initial structure, the bonds that reflect cyclization and hydride shifts en route to cation, following the path suggested by in vacuo geometry optimization (15), were built manually. After construction of each closed conformation harboring its corresponding cation ("transient enzyme complex"), the whole system was subjected to several rounds of energy minimization and a 1-ns MD simulation using the self-parametrizing YASARA2 force field of YASARA Structure. Subsequently, the corresponding cations derived by quantum chemical calculations were docked into the active sites of these transient enzyme complexes to elucidate the existence of additional conformations. Consecutively, localization of docked cations were compared with the manually constructed corresponding cations by structural superposition of Cα atoms using YASARA Structure (5). Conformational clusters that are not derived from the conformation of productive GGPP (Fig. 3*B*) or the respective former cascade carbocation according to Fig. 1 were not considered in cluster analysis. Docking and cluster calculations of QM-derived cations B–E were used to validate the MM-based manual method. Note that simple docking of cationic intermediates into the closed conformation technically harboring unionized GGPP would not reflect the dynamic events during the carbocation cascade.

Homology Modeling of Terpene Synthases. Homology modeling was performed by the homology-modeling module of YASARA Structure, according to the manufacturer's protocol (www.yasara.org/homologymodeling.htm). For homology modeling of casbene synthase from *Jatropha curcas*, the GenBank accession number AB687998 was used as template. For corresponding homology modeling of isopimaradiene synthase from *Picea abies*, the GenBank accession number AY473620.2 was used.

Model of Closed Conformation of CBTS from *N. tabacum*. A genomic truncation of CBTS (GenBank accession number AAS46038.1) lacking the predicted N-terminal transit sequence (amino acids 1–50) was subjected to the homology-modeling module of YASARA Structure. The refined model was subsequently subjected to a structural alignment with the closed conformation of TXS as template. The structural alignment had a RMSD of 0.31 Å over 498 residues with 26.78% primary sequence identity. The unfolded A–C and J–K loops as well as the N terminus of CBTS were then replaced by its corresponding secondary structure elements of TXS followed by "backmutating" to the corresponding CBTS residues and subsequent 10,000 steps of energy minimization and a 10-ns MD simulation in water using YASARA Structure. A second structural alignment between closed conformation of TXS and the "backmutated" and minimized CBTS were generated, and the productive GGPP of TXS were integrated into this CBTS complex. Thereafter, water molecules were transferred from TXS to CBTS and dative bonds between PP_i, water, and the corresponding amino acids were established. A second energy minimization and MD simulation procedure yielded the productive closed conformation of CBTS.

Model of Closed Conformation of CotB2 from *Streptomyces melanosporofaciens*. The modeling approach was according to that of CBTS. The structural alignment of chain D of selinadiene synthase in complex with dihydrofarnesyl pyrophosphate (PDB ID code 4OKZ_D) (37) and the open complex of CotB2 (PDB ID code 4OMG) (36) had a RMSD of 0.18 Å over 287 residues with 13.47% primary sequence identity. After replacement and backmutating of the corresponding loops followed by energy minimization and MD simulation, the productive GGPP from TXS was integrated into this CotB2 complex. Establishment of dative bonds between PP_i, water, and the corresponding amino acids followed by energy minimization and MD simulation yielded the productive closed conformation of CotB2. Localization of cations A–F in the active site of CotB2 were established by manually breaking and linking the corresponding bonds according to the recently published carbocation cascade mechanism (52) followed by energy minimization using the YASARA2 force field of YASARA Structure.

General Experimental Procedures, Bacterial Strains, Genes, Vectors, and Mutagenesis Procedures. See *SI Appendix*.

Expression of TXS and Mutants for in Vitro Screening. Single transformations of pET28b(+) harboring wild-type *txs* or mutations were cultivated for 18 h at 37 °C in LB medium supplemented with 50 $\mu g \cdot mL^{-1}$ kanamycin. These preinocula were used to seed fresh 500 mL of LB media containing 50 $\mu g \cdot mL^{-1}$

kanamycin at a starting OD_{600} of 0.1. After reaching an OD_{600} of 0.6, the cultures were cooled down to 16 °C, induced by 1 mM isopropyl β-D-1-thio-galactopyranoside (IPTG), and grown at 16 °C for 24 h. Cells were harvested by centrifugation at 6,000 × *g* for 15 min and suspended in buffer A [25 mM 3-(*N*-morpholino)propane sulfonic acid (Mops), pH 6.8, 1 mM $MgCl_2$, and 1 mM DTT]. Cells were disrupted by sonification on ice four times (60 s on and 90 s off) at medium power. Cell debris was cleared by centrifugation at 30,000 × *g* for 1 h. The clear supernatant was applied to a preequilibrated HisTrap HP 5-mL column (GE Healthcare) at a flow rate of 1 $mL{\cdot}min^{-1}$ with an ÄKTA Purifier system (GE Healthcare). The loaded column was washed one time with 5 column volumes of buffer A plus 20 mM imidazole and after that eluted with a linear gradient of 20–500 mM imidazole at a flow rate of 2.5 $mL{\cdot}min^{-1}$. SDS/PAGE [12% (vol/vol)] analysis was used to evaluate expression. Selected fractions were combined and dialyzed overnight at 4 °C against buffer A followed by another 6-h dialysis against fresh buffer A. Subsequently, the solution was concentrated to 2 mL and subjected to isocratic size-exclusion chromatography using a Superdex S75 10/300GL column (GE Healthcare) with buffer A at a flow rate of 0.8 $mL{\cdot}min^{-1}$. Fractions were combined and concentrated to 2.5 mL. Concentration was determined by absorbance at 280 nm, and enzyme was stored in 100-μL aliquots at −20 °C for further use.

In Vitro Assay Using Purified Enzymes. One hundred microliters of enzyme solution (1 $mg{\cdot}mL^{-1}$) were added to 400 μL of buffer A. GGPP (Sigma-Aldrich) (43.53 μM) was added to the solution. After gentle mixing, reaction solution was incubated for 12 h at 28 °C and 500 × *g*. For extraction of products, 750 μL of *n*-hexane (VWR) was added to the reaction solution. After vigorous vortexing (2 × 20 s), phases were separated by centrifugation for 10 min, 4 °C and 16,000 × *g*. *n*-Hexane extraction was repeated twice, and organic phases were combined and evaporated to dryness under continuous N_2 stream. The crude extract was resolved in 1 mL of *n*-hexane. One hundred microliters of that resolved extract were evaporated to dryness under continuous N_2 stream, solved in 90 μL, and 10 μL of the internal standard (α-humulene; Sigma-Aldrich; final concentration, 0.889 $mg{\cdot}mL^{-1}$) was added. The samples were analyzed by GC-MS.

In Vivo Production of Deprotonated Intermediates. The in vivo production system was used to generate sufficient amount for NMR analyses. The in vivo approach was based on a culture volume of 3 × 1-L LB medium in 5-L baffled glass flasks. The vectors pColaDuet-1 (*dxr*, *dxs*), pCDFDuet-1 (*ispD/ispF*, *idi*), and the vector pETDuet-1 (*crte*, *txs*) were introduced into *Escherichia coli* BL21(DE3) by standard transformation procedures. For cultivation in shake flasks, single transformants were grown in 3 × 1-L medium containing LB media (10 $g{\cdot}L^{-1}$ tryptone, 5 $g{\cdot}L^{-1}$ yeast extract, and 10 $g{\cdot}L^{-1}$ NaCl) supplemented with 10 $g{\cdot}L^{-1}$ glycerol, 30 $\mu g{\cdot}mL^{-1}$ kanamycin, 50 $\mu g{\cdot}mL^{-1}$ streptomycin, and 50 $\mu g{\cdot}mL^{-1}$ carbenicillin. The cultures were inoculated at OD_{600} of 0.1 from an overnight culture (8-h cultivation 37 °C supplemented with 30 $\mu g{\cdot}mL^{-1}$ kanamycin, 50 $\mu g{\cdot}mL^{-1}$ streptomycin, and 50 $\mu g{\cdot}mL^{-1}$ carbenicillin), grown at 37 °C and 130 rpm (New Brunswick Innova 44/44R Incubator, Eppendorf AG, Hamburg, Germany) until OD_{600} of 0.8 and then cooled down to 25 °C. At 25 °C, 40 $g{\cdot}L^{-1}$ glycerol was added and the cultures were induced by addition of 1 mM IPTG. Cells were grown on 25 °C for 3 d.

Extraction and Isolation of Diterpenes. For diterpene isolation, cells were sedimented (15 min, 17,500 × *g*, 4 °C). Taxa-4,11-diene and its deprotonated intermediates were extracted from *E. coli* cells and supernatants, separately. The cell pellets were washed with water and then suspended in 5 mL of water. Next, cells were lysed by sonification on a Sonoplus HD2070 device (Bandelin Electronic) performing five repeats on ice (5 min on and 3 min off at 80% power) and subsequently extracted three times with 25 mL of *n*-hexane. Supernatants were extracted three times by 200 mL of *n*-hexane. Organic phases were combined, dried with $MgSO_4$, and evaporated under vacuum to dryness. The crude extract was solved in 1 mL of *n*-hexane and analyzed by GC-MS and -flame ionization detector (FID). Purification of verticillia-3,7,12(13)-triene (absolute yield: 3.3 mg out of 3 × 1 LB cultures using the in vivo system described in *In Vivo Production of Deprotonated Intermediates*) was carried out by flash chromatography using an isocratic 90/10 *n*-hexane/ethyl acetate silica step (Silica gel 40; Sigma-Aldrich). Purification of cembrene A was carried out using the same procedure (absolute yield: 2.4 mg out of 3 × 1 LB cultures using the in vivo system described in *In Vivo Production of Deprotonated Intermediates*). The crude extract was resolved in 1 mL of *n*-hexane; 100 μL of that resolved extract was evaporated to dryness under continuous N_2 stream, solved in 90 μL, and 10 μL of the internal standard (α-humulene; Sigma-Aldrich; final concentration, 0.889 $mg{\cdot}mL^{-1}$) was added. The samples were analyzed by GC-MS.

Batch Bioprocess for the Production of Diterpenes. The 30-L fermentations were performed in a 75-L bioreactor (LP351; Bioengineering AG) using LB media (10 $g{\cdot}L^{-1}$ tryptone, 5 $g{\cdot}L^{-1}$ yeast extract, and 10 $g{\cdot}L^{-1}$ NaCl) supplemented with 10 $g{\cdot}L^{-1}$ glycerol, 30 $\mu g{\cdot}mL^{-1}$ kanamycin, 50 $\mu g{\cdot}mL^{-1}$ streptomycin, and 50 $\mu g{\cdot}mL^{-1}$ carbenicillin. The pH was controlled at 6.8 with 4 M NH_4OH and 5 M H_3PO_4. Oxygen saturation was constantly adjusted to 80%. The 1 $mg{\cdot}mL^{-1}$ Antifoam B (Sigma-Aldrich) was automatically added when necessary. The bioprocess was started by inoculation from 4 × 1-L overnight cultures (8-h cultivation 37 °C supplemented with 50 $\mu g{\cdot}mL^{-1}$ kanamycin, 50 $\mu g{\cdot}mL^{-1}$ streptomycin, and 50 $\mu g{\cdot}mL^{-1}$ carbenicillin) at OD_{600} of 0.25 and run 3.5 h at 37 °C. At OD_{600} of 3.8, 40 $g{\cdot}L^{-1}$ glycerol and 1 mM ITPG were added aseptically, and the culture was cooled down to 25 °C and grown for another 5 d. To determine optical densities and glycerol contents, 10-mL aliquots were taken as triplicates. One milliliter of each triplicate was used for determination of OD_{600}, and 1 mL was used for determination of glycerol content by HPLC analysis. After the bioprocess, 500 mL of culture was used for determination of diterpene contents using *n*-hexane extraction of the supernatant as described in *Extraction and Isolation of Diterpenes*. The crude extract was resolved in 1 mL of *n*-hexane; 100 μL of that resolved extract was evaporated to dryness under continuous N_2 stream, solved in 90 μL, and 10 μL of the internal standard (α-humulene; Sigma-Aldrich; final concentration, 0.889 $mg{\cdot}mL^{-1}$) was added. The absolute yield of cembrene A in the 500-mL bioprocess supernatant was 4 mg; the yield of verticillia-3,7,12 (13)-triene was 5.6 mg. The samples were analyzed by GC-MS and GC-FID.

ACKNOWLEDGMENTS. This work was in part supported by German Federal Ministry of Education and Research (BMBF) Grant SysBioTerp (Support Code 031A305A) (to T.B., W.E., V.S., and R.K.) as well as German BMBF Grant ELP (Support Code 1340/68351/3/11) (to T.B. and P.S.).

1. Gao Y, Honzatko RB, Peters RJ (2012) Terpenoid synthase structures: A so far incomplete view of complex catalysis. *Nat Prod Rep* 29(10):1153–1175.
2. Cao R, et al. (2010) Diterpene cyclases and the nature of the isoprene fold. *Proteins* 78(11):2417–2432.
3. Wani MC, Taylor HL, Wall ME, Coggon P, McPhail AT (1971) Plant antitumor agents. VI. The isolation and structure of taxol, a novel antileukemic and antitumor agent from *Taxus brevifolia*. *J Am Chem Soc* 93(9):2325–2327.
4. Koepp AE, et al. (1995) Cyclization of geranylgeranyl diphosphate to taxa-4(5),11(12)-diene is the committed step of taxol biosynthesis in Pacific yew. *J Biol Chem* 270(15):8686–8690.
5. Lin X, Hezari M, Koepp AE, Floss HG, Croteau R (1996) Mechanism of taxadiene synthase, a diterpene cyclase that catalyzes the first step of taxol biosynthesis in Pacific yew. *Biochemistry* 35(9):2968–2977.
6. Williams DC, et al. (2000) 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. *Arch Biochem Biophys* 379(1):137–146.
7. Williams DC, et al. (2000) Intramolecular proton transfer in the cyclization of geranylgeranyl diphosphate to the taxadiene precursor of Taxol catalyzed by recombinant taxadiene synthase. *Chem Biol* 7(12):969–977.
8. Köksal M, Jin Y, Coates RM, Croteau R, Christianson DW (2011) Taxadiene synthase structure and evolution of modular architecture in terpene biosynthesis. *Nature* 469(7328):116–120.
9. Jin Q, Williams DC, Hezari M, Croteau R, Coates RM (2005) Stereochemistry of the macrocyclization and elimination steps in taxadiene biosynthesis through deuterium labeling. *J Org Chem* 70(12):4667–4675.
10. Jin Y, Williams DC, Croteau R, Coates RM (2005) Taxadiene synthase-catalyzed cyclization of 6-fluorogeranylgeranyl diphosphate to 7-fluoroverticillenes. *J Am Chem Soc* 127(21):7834–7842.
11. Chow SY, Williams HJ, Huang Q, Nanda S, Scott AI (2005) Studies on taxadiene synthase: Interception of the cyclization cascade at the isocembrene stage with GGPP analogues. *J Org Chem* 70(24):9997–10003.
12. Chow SY, et al. (2007) Studies on taxadiene synthase: Interception of the cyclization cascade at the verticillene stage and rearrangement to phomactatriene. *Tetrahedron* 63(27):6204–6209.
13. Tokiwano T, et al. (2005) Proposed mechanism for diterpene synthases in the formation of phomactatriene and taxadiene. *Org Biomol Chem* 3(15):2713–2722.
14. Gutta P, Tantillo DJ (2007) A promiscuous proton in taxadiene biosynthesis? *Org Lett* 9(6):1069–1071.
15. Hong YJ, Tantillo DJ (2011) The taxadiene-forming carbocation cascade. *J Am Chem Soc* 133(45):18249–18256.
16. Kobayashi M, Kobayashi K, Nomura M, Munakata H (1990) Conformational study of the cembranoid sarcophytol A, a potent anti-tumor-promoter. *Chem Pharm Bull (Tokyo)* 38(3):815–817.
17. Kobayashi M, Kondo K, Osabe K, Mitsuhashi H (1988) Marine terpenes and terpenoids. V. Oxidation of sarcophytol A, a potent anti-tumor-promoter from the soft coral *Sarcophyton glaucum*. *Chem Pharm Bull (Tokyo)* 36(7):2331–2341.
18. Sugano M, et al. (1995) Phomactin E, F, and G: New phomactin-group PAF antagonists from a marine fungus *Phoma* sp. *J Antibiot (Tokyo)* 48(10):1188–1190.

19. Sugano M, et al. (1996) Structure-activity relationships of phomactin derivatives as platelet activating factor antagonists. *J Med Chem* 39(26):5281–5284.
20. Shen YC, et al. (2008) Cespihypotins Q-V, verticillene diterpenoids from *Cespitularia hypotentaculata*. *J Nat Prod* 71(12):1993–1997.
21. Major DT, Weitman M (2012) Electrostatically guided dynamics—the root of fidelity in a promiscuous terpene synthase? *J Am Chem Soc* 134(47):19454–19462.
22. Christianson DW (2008) Unearthing the roots of the terpenome. *Curr Opin Chem Biol* 12(2):141–150.
23. Li R, et al. (2014) Reprogramming the chemodiversity of terpenoid cyclization by remolding the active site contour of epi-isozizaene synthase. *Biochemistry* 53(7): 1155–1168.
24. Zhou K, Peters RJ (2011) Electrostatic effects on (di)terpene synthase product outcome. *Chem Commun (Camb)* 47(14):4074–4080.
25. Srividya N, Davis EM, Croteau RB, Lange BM (2015) Functional analysis of (4*S*)-limonene synthase mutants reveals determinants of catalytic outcome in a model monoterpene synthase. *Proc Natl Acad Sci USA* 112(11):3332–3337.
26. Whittington DA, et al. (2002) Bornyl diphosphate synthase: Structure and strategy for carbocation manipulation by a terpenoid cyclase. *Proc Natl Acad Sci USA* 99(24): 15375–15380.
27. Shishova EY, Di Costanzo L, Cane DE, Christianson DW (2007) X-ray crystal structure of aristolochene synthase from *Aspergillus terreus* and evolution of templates for the cyclization of farnesyl diphosphate. *Biochemistry* 46(7):1941–1951.
28. Shishova EY, et al. (2008) X-ray crystallographic studies of substrate binding to aristolochene synthase suggest a metal ion binding sequence for catalysis. *J Biol Chem* 283(22):15431–15439.
29. Hyatt DC, et al. (2007) Structure of limonene synthase, a simple model for terpenoid cyclase catalysis. *Proc Natl Acad Sci USA* 104(13):5360–5365.
30. Hong YJ, Tantillo DJ (2009) A potential energy surface bifurcation in terpene biosynthesis. *Nat Chem* 1(5):384–389.
31. Hong YJ, Tantillo DJ (2010) Quantum chemical dissection of the classic terpinyl/pinyl/bornyl/camphyl cation conundrum—the role of pyrophosphate in manipulating pathways to monoterpenes. *Org Biomol Chem* 8(20):4589–4600.
32. Tantillo DJ (2011) Biosynthesis via carbocations: Theoretical studies on terpene formation. *Nat Prod Rep* 28(6):1035–1053.
33. Zhou K, et al. (2012) Insights into diterpene cyclization from structure of bifunctional abietadiene synthase from *Abies grandis*. *J Biol Chem* 287(9):6840–6850.
34. Keeling CI, Weisshaar S, Lin RPC, Bohlmann J (2008) Functional plasticity of paralogous diterpene synthases involved in conifer defense. *Proc Natl Acad Sci USA* 105(3): 1085–1090.
35. Nakano Y, et al. (2012) Characterization of the casbene synthase homolog from *Jatropha*. *Plant Biotechnol* 29:185–189.
36. Janke R, Görner C, Hirte M, Brück T, Loll B (2014) The first structure of a bacterial diterpene cyclase: CotB2. *Acta Crystallogr D Biol Crystallogr* 70(Pt 6):1528–1537.
37. Baer P, et al. (2014) Induced-fit mechanism in class I terpene cyclases. *Angew Chem Int Ed Engl* 53(29):7652–7656.
38. Rynkiewicz MJ, Cane DE, Christianson DW (2002) X-ray crystal structures of D100E trichodiene synthase and its pyrophosphate complex reveal the basis for terpene product diversity. *Biochemistry* 41(6):1732–1741.
39. Vedula LS, Cane DE, Christianson DW (2005) Role of arginine-304 in the diphosphate-triggered active site closure mechanism of trichodiene synthase. *Biochemistry* 44(38): 12719–12727.
40. Aaron JA, Lin X, Cane DE, Christianson DW (2010) Structure of epi-isozizaene synthase from *Streptomyces coelicolor* A3(2), a platform for new terpenoid cyclization templates. *Biochemistry* 49(8):1787–1797.
41. Cui H, Zhang ST, Yang HJ, Ji H, Wang XJ (2011) Gene expression profile analysis of tobacco leaf trichomes. *BMC Plant Biol* 11:76.
42. Drenth J, Jansonius JN, Koekoek R, Swen HM, Wolthers BG (1968) Structure of papain. *Nature* 218(5145):929–932.
43. Basar S, Koch A, König WA (2001) A verticillane-type diterpene from *Boswellia carterii* essential oil. *Flavour Fragrance J* 16(5):315–318.
44. Xu M, Wilderman PR, Peters RJ (2007) Following evolution's lead to a single residue switch for diterpene synthase product outcome. *Proc Natl Acad Sci USA* 104(18): 7397–7401.
45. Morrone D, Xu M, Fulton DB, Determan MK, Peters RJ (2008) Increasing complexity of a diterpene synthase reaction with a single residue switch. *J Am Chem Soc* 130(16): 5400–5401.
46. Wilderman PR, Peters RJ (2007) A single residue switch converts abietadiene synthase into a pimaradiene specific cyclase. *J Am Chem Soc* 129(51):15736–15737.
47. Starks CM, Back K, Chappell J, Noel JP (1997) Structural basis for cyclic terpene biosynthesis by tobacco 5-epi-aristolochene synthase. *Science* 277(5333):1815–1820.
48. Lesburg CA, Caruthers JM, Paschall CM, Christianson DW (1998) Managing and manipulating carbocations in biology: Terpenoid cyclase structure and mechanism. *Curr Opin Struct Biol* 8(6):695–703.
49. Kirby J, et al. (2010) Cloning of casbene and neocembrene synthases from Euphorbiaceae plants and expression in *Saccharomyces cerevisiae*. *Phytochemistry* 71(13): 1466–1473.
50. Meguro A, Tomita T, Nishiyama M, Kuzuyama T (2013) Identification and characterization of bacterial diterpene cyclases that synthesize the cembrane skeleton. *ChemBioChem* 14(3):316–321.
51. Hong YJ, Tantillo DJ (2015) The energetic viability of an unexpected skeletal rearrangement in cyclooctatin biosynthesis. *Org Biomol Chem* 13(41):10273–10278.
52. Meguro A, et al. (2015) An unusual terpene cyclization mechanism involving a carbon-carbon bond rearrangement. *Angew Chem Int Ed Engl* 54(14):4353–4356.
53. Görner C, Häuslein I, Schrepfer P, Eisenreich W, Brück T (2013) Targeted engineering of cyclooctat-9-en-7-ol synthase: A stereospecific access to two new non-natural fusicoccane-type diterpenes. *ChemCatChem* 5(11):3289–3298.
54. Vedula LS, et al. (2005) Molecular recognition of the substrate diphosphate group governs product diversity in trichodiene synthase mutants. *Biochemistry* 44(16): 6153–6163.
55. Frisch MJ, et al. (2004) Gaussian 03, Revision (Gaussian, Inc., Wallingford, CT).
56. Krieger E, et al. (2009) Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. *Proteins* 77(Suppl 9):114–122.
57. Jakalian A, Jack DB, Bayly CI (2002) Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J Comput Chem* 23(16): 1623–1641.
58. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general amber force field. *J Comput Chem* 25(9):1157–1174.
59. Krieger E, Nielsen JE, Spronk CA, Vriend G (2006) Fast empirical pKa prediction by Ewald summation. *J Mol Graph Model* 25(4):481–486.
60. Trott O, Olson AJ (2010) AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* 31(2):455–461.
61. Jung J, Lee B (2000) Protein structure alignment using environmental profiles. *Protein Eng* 13(8):535–543.