



Heterocyclic and Alkyne Terpenoids by Terpene Synthase-Mediated Biotransformation of Non-Natural Prenyl Diphosphates: Access to New Fragrances and Probes

Benjamin Weigel^{+, [b]}, Jeanette Ludwig^{+, [b]}, Roman A. Weber^{+, [b]}, Steve Ludwig,^[b] Claudia Lennicke,^[b] Paul Schrank,^[b] Mehdi D. Davari,^[b] Mohamed Nagia,^[b, c] and Ludger A. Wessjohann^{*[a]}

Two terpene cyclases were used as biocatalytic tool, namely, limonene synthase from *Cannabis sativa* (CLS) and 5-*epi*-aristolochene synthase (TEAS) from *Nicotiana tabacum*. They showed significant substrate flexibility towards non-natural prenyl diphosphates to form novel terpenoids, including core oxa- and thia-heterocycles and alkyne-modified terpenoids. We

elucidated the structures of five novel monoterpene-analogues and a known sesquiterpene-analogue. These results reflected the terpene synthases' ability and promiscuity to broaden the pool of terpenoids with structurally complex analogues. Docking studies highlight an on-off conversion of the unnatural substrates.

Introduction

Terpenoids constitute one of the largest classes of secondary metabolites comprising more than 80000 compounds, many of which have high biological and pharmaceutical importance as reported in the Dictionary of Natural Products (<http://dnp.chemnetbase.com>).^[1] They act as attractants, toxins, repellants, or antibiotics among other biological activities.^[2] Volatile mono- and sesquiterpenes are the main constituents of flower scents and natural aromas. Thus, these terpenes have significant commercial value as food additives, and in the cosmetic and cosmeceutical industry for fragrances.^[3] However, subtle variations in the terpene composition of a flower bouquet can cause dramatic changes in its fragrance. Even more intriguing, minor

changes in the chemical structure of terpenes can have a substantial influence on smell perception.^[4] All terpenoids originate from a few very simple, linear, non-chiral prenyl diphosphates.^[5] By action of terpene synthases (terpene cyclases), these substrates are folded and converted into complex, mostly enantioselective and multicyclic structures by complex mechanisms involving cation rearrangements, stabilization cascades, and selective deprotonation, or nucleophile addition.^[6] Thus terpene cyclases catalyze one of the most complex biochemical reactions. On average, one-half of the carbon atoms of the substrate are subject to the new bond formation, and/or change in stereochemistry via a multistep cyclization scheme.^[5b] Due to the type of prenyl modifications, the enzymes are classified as prenyltransferases, which attach a prenyl chain to a second aliphatic or an aromatic substrate,^[2a,7] or terpene synthases (terpene cyclases), which catalyze an intramolecular cyclization of geranyl- (GPP), farnesyl- (FPP), or geranylgeranyl diphosphate (GGPP), etc. to mono-, sesqui- and diterpenes, respectively.^[6b,8]

In our ongoing research to utilize non-natural substrates with prenyl-diphosphate converting enzymes,^[7b,9] we will concentrate, in this paper, on highly unusual substrates with hetero atoms or triple bonds in the prenyl chain which can serve as indicators for reaction pathways, are suitable for labeling to follow them in biological systems, and have the power to generate totally novel, complex heterocycles (and eventually fragrances). Here, this is exemplified by the action of two well-characterized terpene cyclases: (–)-limonene synthase from *Cannabis sativa* (CLS), which catalyzes the conversion of GPP to limonene,^[10] and 5-*epi*-aristolochene synthase from *Nicotiana tabacum* (TEAS), a sesquiterpene synthase converting FPP to 5-*epi*-aristolochene.^[11] As typical for terpene cyclases, both enzymes generate a variety of other cyclic and acyclic products.^[10,12] In addition to the main product of TEAS, several low abundant carbocycles are formed via the tightly bound intermediate germacrene A.^[12–13] O'Maille and co-workers

[a] Prof. Dr. L. A. Wessjohann
Natur- und Wirkstoffchemie
Leibniz-Institut für Pflanzenbiochemie
Weinberg 3, 06120 Halle/Saale (Germany)
E-mail: wessjohann@ipb-halle.de

[b] Dr. B. Weigel,⁺ Dr. J. Ludwig,⁺ Dr. R. A. Weber,⁺ Dr. S. Ludwig, Dr. C. Lennicke,
P. Schrank, Dr. M. D. Davari, Dr. M. Nagia
Natur- und Wirkstoffchemie
Leibniz-Institut für Pflanzenbiochemie
Weinberg 3, 06120 Halle/Saale (Germany)

[c] Dr. M. Nagia
Additional address:
Department of Chemistry of Natural Compounds
Pharmaceutical and Drug Industries Research Institute
National Research Center
El Buhouth St. 33, 12622 Cairo (Egypt)

[⁺] These authors contributed equally to this work.

 Supporting information for this article is available on the WWW under <https://doi.org/10.1002/cbic.202200211>

 This article is part of a Special Collection dedicated to the Biotrans 2021 conference. Please see our homepage for more articles in the collection.

 © 2022 The Authors. ChemBioChem published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

reported the formation of additional minor products from (*Z,E*)-FPP via the cisoid pathway of TEAS beside the sesquiterpenes originating from its original transoid mechanism of the (*E,E*)-FPP.^[13]

Furthermore, some similarities are apparent in the reaction mechanism of both terpene synthases.^[5a] Cleavage of the diphosphate moiety of the substrate is facilitated by complexation to divalent metal ions such as Mg²⁺.^[6b,14] Computational studies suggested a protonation-dependent diphosphate cleavage in terpene synthases, highlighting the critical effects of Mg²⁺, the diphosphate protonation states, and the coordination interactions on promoting the diphosphate activation as leaving group.^[15] Subsequently, the resulting (transient) carbocation is converted to terpenoid hydrocarbons by the intramolecular attack of double bonds or rearrangements, eventually terminated by deprotonation (\rightarrow terpene) or nucleophile attack, mostly by the addition of water to the terpenoid alcohol.^[16] The carbocation formation and stabilization thus is the key step in the catalysis, which opens the door for (excessive) rearrangements, bond formations, methyl and proton shifts, etc. and serves as the basis for the diversity of terpenes.^[16a] Indeed, rearrangement is crucial for product formation in monoterpene synthases, but also other terpene synthases.^[5a] For a direct product formation, the intermediate cation has to be stabilized, sometimes with increasing intensity along the reaction pathway in different positions, and eventually a directed deprotonation/nucleophile addition has to render the process irreversible.

Understanding detailed catalytic mechanisms of TEAS and CLS is crucial for understanding their product specificity. Computational studies proposed three critical chemical control factors to be responsible for the catalytic promiscuity and fidelity in most sesquiterpene cyclases, including substrate folding mode, intermediate flexibility, and key residues.^[17] QM/MM calculations revealed that the catalytic promiscuity of TEAS is mostly due to the significant conformational dynamics of the branching intermediate cation and intrinsic intermediate flexibility is highly correlated to the plasticity of the TEAS active site pocket contour.^[17–18] It is also proposed that the Asp444-Tyr520 dyad might act as an additional general acid/base residue pair to increase TEAS promiscuity.^[17–18] Molecular dynamics simulations indicated the role of substrate binding kinetics and protein conformational dynamics responsible for the cyclization reaction in TEAS.^[19] The catalytic mechanism of CLS is not yet fully studied by computational methods, probably because of the lack of a crystal structure. However, QM/MM simulations for a homolog of CLS, i.e., (4*S*)-limonene synthase, showed a concerted-asynchronous (three-steps) reaction pathway, consisting of the isomerization, cyclization process, and the proton-transfer process. In this proposed mechanism, a conserved His residue can act as a general base to deprotonate (4*S*)- α -terpinyl carbocation and generate the limonene product.^[20]

Conversion of substrate analogues by terpene synthases is far from utilizing its full potential.^[21] Enzymatic transformation of artificial substrates is a promising opportunity to enhance or extend the abilities for biological synthesis which, in turn, boosts the natural terpenome with non-natural terpenoids with altered properties. The latter coincidentally gives new insights

into the reaction mechanism of both, mono- and sesquiterpene synthases. While several substrates were tested for other categories of prenyl-converting enzymes, namely prenyltransferases, to enhance their detection limits and examine their substrate specificities,^[22] for TEAS only the conversion of 6-fluorofarnesyl diphosphate has been reported, which proceeded through the formation of the neutral, tightly bound intermediate 1-fluorogermacrene A during the catalysis.^[21a]

Here we report the action of 5-*epi*-aristolochene synthase (TEAS) and limonene synthase (CLS) on eleven artificial organic diphosphates containing potential donor and cation stabilizing atoms and moieties such as oxygen, sulfur, alkyne, or aromatic groups. Theoretically, some of these groups can also serve as nucleophilic terminators (Figure 1). They formed a series of novel cyclic and acyclic, aliphatic, and non-aliphatic products. Strikingly, activity towards two of the substances, namely substrates 5 and 6 in the case of CLS or TEAS, respectively, is even higher than for the natural substrates GPP and FPP. The structures of six reaction products were elucidated, and a mechanism for the enzyme-mediated cyclization is proposed. Computational modeling was performed to gain a molecular understanding of the structural determinants for substrate binding and conversion.

Results and Discussion

CLS and TEAS catalyze the cyclization of GPP and FPP

CLS and TEAS were produced by heterologous expression in *E. coli*, with typical yields of 10 and 40 mg of protein per liter of culture, respectively. The enzymes were assayed using the naturally occurring prenyl diphosphates dimethylallyl diphosphate (DMAPP), GPP, FPP, and GGPP. The lipophilic products were separated from substrates by liquid-liquid extraction with *n*-heptane/*n*-hexane, and the organic phase was analyzed by GC/MS. CLS was most active towards GPP, which was cyclized into (*S*)-limonene and a variety of other aliphatic and hydroxylated monoterpenes, as reported (see Supporting Table S7).^[10] Hardly any activity was observed with FPP as substrate (Figure 2).

On the other hand, TEAS converted GPP in addition to its physiological substrate FPP. GPP was utilized to a considerable extent (20% of relative conversion vs. FPP), whereby, a broad spectrum of products is formed. The main reaction product was linalool, but also linear, mono- and bicyclic monoterpenes were produced (see Supporting Table S8). Such conversion was not reported before. Cyclization of the transoid geranyl cation does not allow formation of cyclic monoterpenes. Thus in accordance with common knowledge in the field, we propose that these products arise from the cisoid neryl cation, which would be formed analogously to the cisoid (*Z,E*)-farnesyl cation during FPP conversion (Supporting Figure S2). This second way of cyclization was first described for TEAS by O'Maille in 2006.^[13]

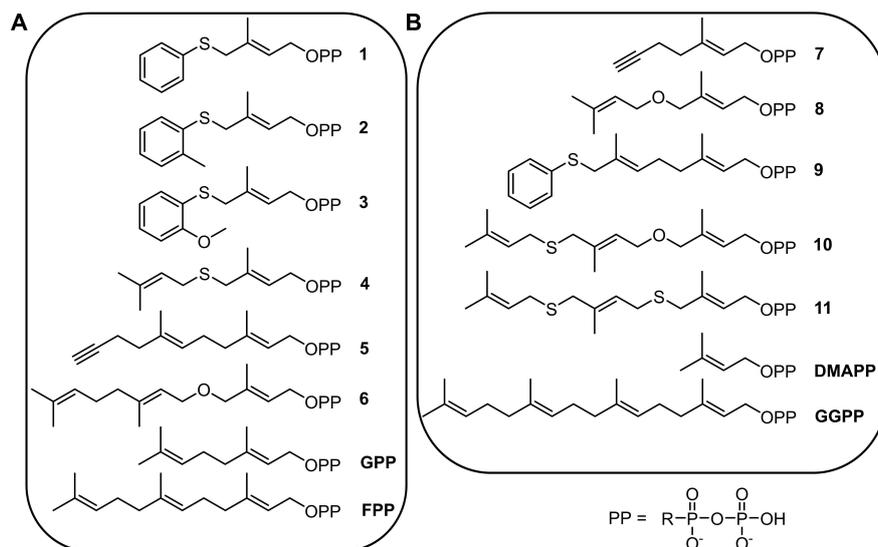


Figure 1. Prenyl diphosphates applied for transformations by CLS and TEAS. A) Prenyl diphosphates accepted by CLS or TEAS. B) prenyl diphosphates which were not transformed by CLS and/or TEAS (B).

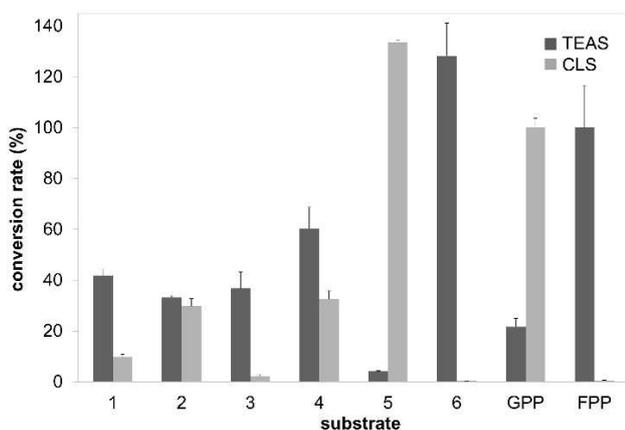


Figure 2. Overall substrate conversion rates of six artificial and two natural isoprenoid substrates. Shown are relative conversion rates to the natural substrate GPP for reactions of CLS, or FPP for reactions of TEAS. Enzymatic assays were performed in triplicates.

Both CLS and TEAS show enzymatic activity towards novel prenyl diphosphates

Eleven synthetic prenyl diphosphates (compounds 1–11; Figure 1) were applied for transformations by CLS and TEAS. The conversion of the artificial substrates was analyzed by GC/MS (for chromatograms and GC/MS spectra, see Supporting Information), followed by the identification of the dominant products by mass spectrometry and NMR where required.

No transformation of substrates 7–11 was observed by any of the enzymes. However, substrates 1–6 were converted by both enzymes (Figure 2). Abiotic conversion, i.e. reaction without or with denatured enzyme was absent in all cases, while products of the enzymatic reaction turned out to be non-racemic (see Supporting Information for chiral chromatography

data). Differences in the total conversion of substrates were observed between either of the enzymes where CLS accepted substrate 5 even better than the natural substrate GPP, while substrate 6 turned out to be the least acceptable one. For TEAS the order was inverse, as would be expected from the natural chain length preferences. (CLS: 5 > GPP ≫ 6; TEAS: 6 > FPP ≫ 5; Figure 2).

In the transformation of the artificial substrates, both enzymes produced multiple reaction products. This catalytic promiscuity was to be expected since the enzymatic conversion of the natural substrates GPP and FPP already produces a spectrum of terpenes. Fortunately, biotransformation of the artificial substrates yielded one main product (> 50%), except for substrate 4 which produced several minor compounds (Table 1).

Differences in the substrate specificity were also reflected by the conversion of the aromatic prenyl diphosphates 1–3, which showed higher conversion with TEAS than with CLS (Figure 2). Notably, TEAS almost exclusively transformed these substrates into the hydrolysis products shown in Scheme 2A, whereas conversion with CLS also yielded a substantial amount of other, non-cyclized products. By comparison of data obtained from GC/MS experiments with those from literature,^[23] we surmise one of the products from CLS conversion of substrate 1, i.e. (1a) to be 2-(phenylthiomethyl)-1,3-butadiene.

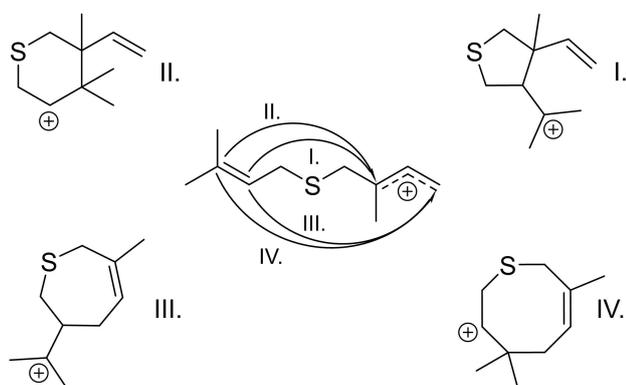
Conversion of artificial diphosphates affords structures analogous to natural monoterpenes and sesquiterpenes

For structure elucidation, enzymatic product 4b was isolated by preparative GC to be analyzed by ¹H NMR spectroscopy. Based on hypothetically possible cyclization patterns, four intermediates for the conversion of 4 could be theorized (Scheme 1). The ¹H NMR spectrum of 4b shows 16 protons, five of which were

Table 1. Product distribution for different organic diphosphate substrates for CLS and TEAS.

Substrate	Enzyme	Product distribution [%] ^[a]																Yield [times] ^[b]
		a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	
1	CLS	48.1	1.1	1.0	27.6	1.2	19.5											0.1
	TEAS	7.1	ND	ND	89.9	ND	1.4											0.4
2	CLS	66.2	28.8	ND	4.3													0.3
	TEAS	4.4	94.6	1.0	ND													0.3
3	CLS	51.0	49.0															0.02
	TEAS	ND	97.3															0.4
4	CLS	1.6	31.5	23.5	ND	8.9	6.9	15.0	1.9	5.7	ND	1.2	3.0	ND				0.3
	TEAS	ND	29.6	31.1	1.6	ND	2.4	3.3	8.4	8.3	2.4	4.7	1.3	2.5				0.6
5	CLS	ND	7.8	1.1	4.6	ND	57.6	ND	ND	1.9	ND	2.4	ND	ND	2.0	14.1	ND	1.3
	TEAS	31.5	ND	ND	ND	2.3	ND	4.9	4.7	ND	2.4	ND	39.8	6.9	ND	ND	7.5	0.04
6	CLS	ND																0.004
	TEAS	99.3																1.3
GPP	CLS	5.7	3.8	76.9	ND	ND	ND	ND	1.7	3.8	2.1	ND	4.2	ND				1
	TEAS	ND	23.1	7.0	3.5	2.7	2.3	8.5	46.4	ND	ND	1.0	ND	4.5				0.2
FPP	CLS	ND	41.2	ND	ND	ND	ND	58.8										0.01
	TEAS	9.9	ND	1.1	78.0	6.5	1.3	ND										1

ND = product not detected or relative abundance below 1%. Enzymatic products are numbered by small letters according to their retention times in GC-MS chromatograms. A product is given the same letter when it appears in the GC chromatograms of both enzymes with the same substrate (see Supporting Information). [a] Only products with an abundance of more than 1% are considered. All product distributions were determined from triplicates and a hydrolysis control was performed to determine non-enzymatic reactions. [b] Conversion of the named substrate relative to that of the natural substrate (GPP for CLS, or FPP for TEAS). Marked in bold are new, artificial terpene synthase products (structure elucidated in this study).



Scheme 1. Possible cyclization patterns of substrate 4. Based on hypothetically possible cyclization mechanisms, four different prenyl cations could be assumed as initial intermediates for the cyclization of this substrate (center).

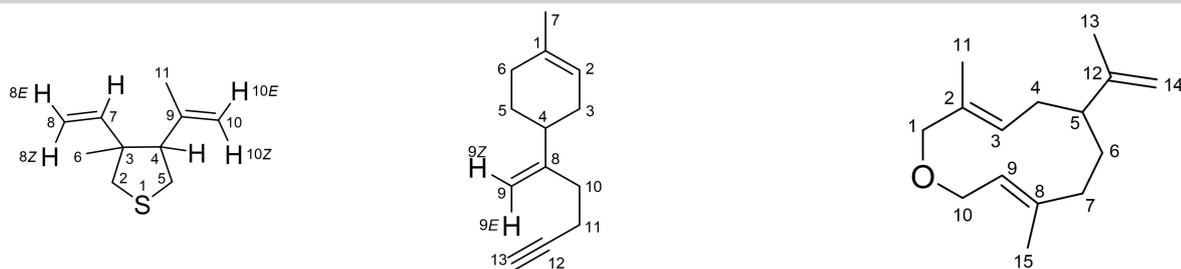
olefinic protons. Based on this information, cyclic scaffolds III and IV (Scheme 1) having a maximum of three olefinic protons, could be discarded as core structures. On the other hand, cations I and II would both retain five olefinic protons after deprotonation. However, the coupling pattern strongly suggests two exocyclic terminal double bonds. In contrast to scaffold II, these data strongly support cyclization of 4 via scaffold I, leading to the bis-vinyl compound 4b after deprotonation (Table 2).

The downfield signal at 5.96 ppm (H-7) shows vicinal couplings ($^3J_{8E,7} = 17.5$ Hz, $^3J_{8Z,7} = 10.9$ Hz) with protons H-8E ($\delta = 5.09$ ppm) and H-8Z ($\delta = 5.16$ ppm). Protons H-8E and H-8Z in turn also display a small geminal coupling ($^2J_{8E,8Z} = 1.3$ Hz). Another geminal coupling ($^2J_{10E,10Z} = 2.0$ Hz) is found for protons H-10E ($\delta = 4.83$ ppm) and H-10Z ($\delta = 4.91$ ppm). Also protons H-10E/Z display 4J -couplings ($^4J_{10E,11} = 0.8$ Hz, $^4J_{10Z,11} = 1.5$ Hz) with the methyl group at position 11 ($\delta = 1.74$ ppm) consistent with

our proposed structure. The methylene protons H-5a and H-5b solely present a geminal coupling ($^2J_{2a,2b} = 10.6$ Hz) with each other. The second pair of methylene protons adjacent to the sulfur (H-5a/b) each show a geminal ($^2J_{5a,5b} = 10.5$ Hz), as well as a 3J coupling to proton H-4 ($\delta = 2.51$ ppm). Using these NMR data together with the EI-fragmentation, we could confirm our proposed structure (Table 2).

Based on (chiral) GC/MS measurements, we could confirm the formation of four possible stereoisomers of 4b with CLS distributed with the following ratios (48%, 18%, 20%, 14%) and three thereof with TEAS (60%, 22%, 18%). GC/MS measurements revealed the formation of two diastereomers but based on the TIC of chiral GC/MS two additional stereoisomers were separated. Stereoisomers were formed in disparate ratios which confirms an enzymatic formation of 4b (see Supporting Information).

The constitution of 5f was validated by a combination of different NMR spectroscopy methods, including correlation spectroscopy (COSY), heteronuclear multiple-bond correlation spectroscopy (HMBC), and heteronuclear single-quantum correlation spectroscopy (HSQC; Table 2). The key resonances in the corresponding 1H NMR and HSQC spectra are the two methylene protons H-9Z ($\delta = 4.83$ ppm) and H-9E ($\delta = 4.78$ ppm), as well as the methine proton H-2 ($\delta = 5.4$ ppm) that exhibit a distinct downfield shift. Furthermore, the methyl group at position 7 ($\delta = 1.65$ ppm) is crucial. Standard measurement conditions did not show a correlation signal H-13/C-13. However, an optimized HSQC experiment without ^{13}C decoupling revealed said correlation with a $^1J_{H,C}$ coupling constant of 247 Hz, which is characteristic for methine groups. ^{13}C signals at 84.3, 133.8, and 152.4 ppm do not show HSQC correlations and were thus assigned as quaternary carbon atoms C-12, C-1, and C-8, respectively. The protons of a methylene group (H-11a/b) with an unusual ^{13}C upfield shift of 17.7 ppm show HMBC

Table 2. ¹H and ¹³C-NMR data of enzyme products.


4b			5f				6a				
Pos.	δ_{H} , mult.	(J in Hz)	Pos.	δ_{H} , mult.	δ_{C}	HMBC corr.	Pos.	δ_{H} , mult.	(J in Hz)	δ_{C}	HMBC corr.
2 a	2.77, d	(10.6)	1		133.8	3, 5, 6, 7	1 a	4.04, d	(12.5)	79.9	2, 3, 10, 11
2 b	2.86, d	(10.6)	2	5.40, m	120.6	3, 4, 6, 7	1 b	3.86, d	(12.5)		
4	2.51, dd	(9.1, 7.5)	3 a	2.12, m	31.3	1, 2, 4, 5, 8	2			135.6	1, 4, 11
5 a	2.95, dd	(10.5, 9.1)	3 b	1.89, m			3	5.38, m		126.6	1, 4, 5, 11
5 b	2.97, dd	(10.5, 7.5)	4	2.10, m	39.6	2, 3, 5, 6, 8, 9, 10	4 a	2.05, m		33.4	2, 3, 5, 6, 12
6	1.22, s		5 a	1.82, m	28.2	1, 3, 4, 6, 8	4 b	1.90, m			
7	5.96, dd	(17.5, 10.9)	5 b	1.48, m			5	1.88, m		48.6	3, 4, 6, 7, 12, 13, 14
8 E	5.09, ddq	(10.9, 1.3, 0.4)	6 a	2.04, m	30.7	1, 2, 4, 5, 7	6 a	1.65, m		32.8	4, 5, 7, 8, 12
8 Z	5.16, dd	(17.5, 1.3)	6 b	1.96, m			6 b	1.39, m			
10 E	4.83, ddq	(2.0, 0.8, 0.8)	7	1.65, s	23.4	1, 2, 6	7 a	2.12, dd	(13.1, 6.1)	39.8	5, 6, 8, 9, 15
10 Z	4.91, dq	(2.0, 1.5)	8		152.4	3, 4, 5, 9, 10, 11	7 b	1.73, td	(13.1, 2.1)		
11	1.74, dd	(1.5, 0.8)	9 E	4.78, m	108.1	4, 8, 10	8				
			9 Z	4.83, s			9	5.28, m		124.7	7, 10, 15
			10 a	2.30, m	33.6	4, 8, 9, 11, 12	10 a	4.22, dd	(13.2, 8.5)	69.2	1, 8, 9
			10 b	2.30, m			10 b	4.17, dd	(13.2, 6.1)		
			11 a	2.34, m	17.7	8, 10, 12, 13	11	1.67, s		14.5	1, 2, 3
			11 b	2.28, m			12			151.8	4, 5, 6, 13, 14
			12		84.3	10, 11, 13	13	1.73, dd	(1.4, 0.8)	19.6	5, 12, 14
			13	1.96, m	68.3	11, 12	14 a	4.69, dq	(2.2, 0.8)	107.7	5, 12, 13
							14 b	4.60, dq	(2.2, 1.4)		
							15	1.69, s		17.7	7, 8, 9

correlations with quaternary carbons C-12 and C-8, which, owing to its downfield shift, is apparently sp^2 hybridized. Protons H-11a/b also display correlations with C-10 ($\delta = 33.6$ ppm) and C-13 ($\delta = 68.3$ ppm). This demonstrates the connection of C-11 to the terminal triple bond (C-12/C-13). The methylene protons H-10 exhibit HMBC correlations with C-8, C-9, C-11, and C-12. In turn, H-9a and H-9b correlate with C-8 and C-10. Methine signal H-4 ($\delta = 2.10$ ppm) also shows an HMBC correlation with C-8. C-4 correlates with protons H-9a/b. This provides the reason for the proposed hex-5-en-1-ynyl moiety, which is connected to C-4. The protons of the methyl group at position 7 display correlations with C-6, the quaternary carbon C-1 as well as with the methine carbon C-2. H-2, in turn, correlates with C-7, C-4, and C-3. Both protons of the methylene group C-5 show HMBC correlations with C-4 and C-8. Furthermore, the vicinal coupling of H-2 with H-3a and H-3b, observed by COSY, and the NOEs between H-9a and H-5a/b (H-3a/b) can only be explained by a connection of C-2 to C-4 via C-3. The entirety of 1D- and 2D-NMR data led to the structure elucidation of propargyl-limonene (**5f**). The chiral GC/MS chromatogram of compound **5f** hints to the formation of one enantiomer only as it showed only one sharp peak. This data should, however, be considered with caution due to the unavailability of a racemic standard to prove the separability of the two enantiomers (see Supporting Information).

The structure determination of **6a** was accomplished similarly, using HSQC and HMBC experiments in addition to the corresponding proton and carbon spectra. The latter revealed the existence of six sp^2 hybridized carbons based on six signals between 126.0 (C-3) and 151.8 (C-12) ppm. Three of them did not show any HSQC correlations and thus, seem to be quaternary carbon atoms ($\delta = 135.6$ (C-2), 137.4 (C-8), 151.8 ppm (C-12)). The remaining sp^2 hybridized carbons show correlations with either one proton ($\delta = 126.6$ (C-3), 124.7 ppm (C-9)) or two protons ($\delta = 151.8$ ppm (C-14)) which evidences a terminal methylene group at position 14. The HMBC correlations of H-3 and H-9 with the corresponding quaternary carbons at positions 2 and 8 are not visible due to the selected parameters for this experiment. However, it can be proved indirectly that C-2 and C-3, as well as C-8 and C-9 are connected to each other, taking a look at correlations of atoms in close vicinity. In the first case, H-3 shows correlations with C-1, C-4, C-5, and C-11, whereas C-2 shows correlations with H-1, H-4, and H-11. Moreover, H-11 shows correlations with C-1, C-2, and C-3 while H-1 shows correlations to C-2, C-3, C-10, and C-11. Combined with the corresponding chemical shifts, these correlations verify the connection of C-2 and C-3 via a double bond. The interconnection of C-8 and C-9 can be investigated likewise, yielding the same result. The connection of positions 1 and 10, via an oxygen atom, can be estimated by the chemical shifts of the involved atoms ($\delta = 4.04$ (H-1a), 3.86 (H-1b), 4.22

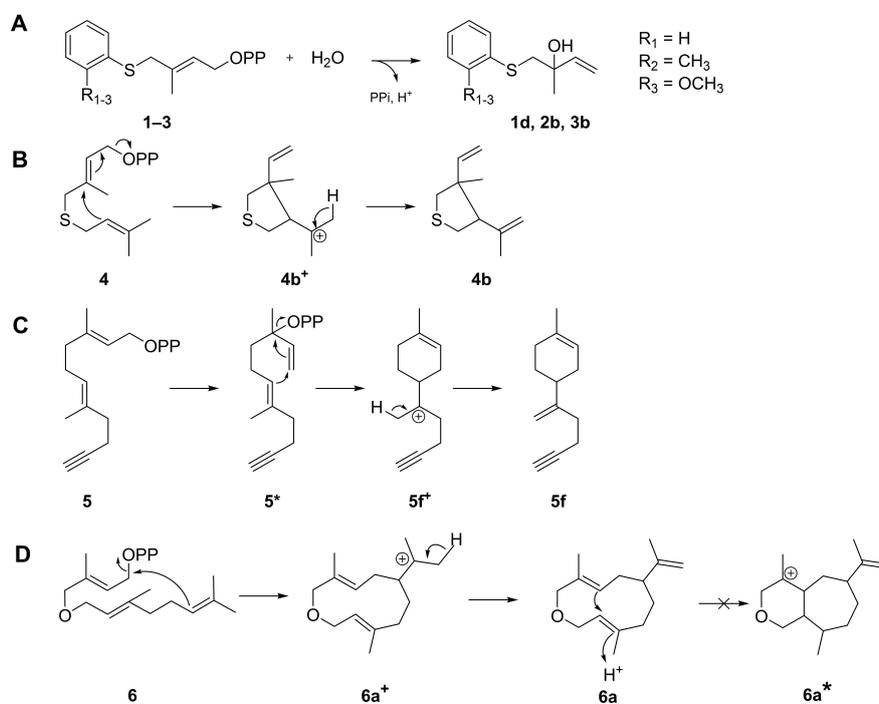
(H-10a), 4.17 (H-10b), 79.9 (C-1), 69.2 ppm (C-10) and is proved by the presence of an HMBC correlation for these positions and the concurrent absence of correlations for position 1 and 9 as well as position 2 and 10, respectively. Finally, the size of the ring, which is formed during the enzymatic reaction, is determined by the HMBC correlation of H-5 with C-3, C-4, C-6, C-7, C-12, C-13, and C-14. This attests to a ring closure at position 5, leading to the formation of an eleven-membered ring and, therefore, to the formation of the homo-oxa-germacrene (**6a**). Since the double bonds of substrate **6** are in *E*-configuration it can be assumed that they are in the same configuration in the product. Additionally, ^1H NMR data show a fixed conformation of the 11-membered ring of **6a** as protons of the CH_2 -groups of the ring can be distinguished into axial and equatorial protons since they generate separate signals. Due to the fixed ring conformation, terpenoid **6a** can only occur in the form of two diastereomers, albeit not as enantiomers. As only one set of ^1H - and ^{13}C -NMR signals is obtained from compound **6a**, as well as (chiral) GC/MS measurements show only one sharp signal of the terpenoid and a specific rotation of $[\alpha]_{589}^{24,6} = -37.5^\circ$ was determined, the isolated (-)-homo-oxa-germacrene (**6a**) is proven to be one single diastereomer. The macrocyclic ether **6a** also was discovered in a parallel study by Kirschning et al. as an enzymatic product from the ether derivative of FPP (**6**) with several sesquiterpene cyclases from plants and fungi.^[24] Further heterocyclic germacrene derivatives were reported from FPP-analogues with a 10,11-epoxide or allylic alcohols by germacrene synthases, but in the latter cases, the oxygen is not

introduced in the prenyl chain backbone, and does not show an interruption of the native cyclization as in our or Kirschning's case.^[21d]

Interestingly, the enzymatic conversion of the substrates **1–3** generated products with a characteristic prominent m/z of 71 in their EI-MS spectra. Thus, the formation of the products **1d**, **2b**, and **3b** (Scheme 2A) was proposed, which yield a fragment with the corresponding mass in MS experiments due to α cleavage of the 1-hydroxy-1-methylprop-2-enyl moieties. The formation of the suggested products was confirmed by comparison with synthetic standards in GC/MS experiments. The enantiomeric distribution was investigated by chiral GC/MS in comparison to the racemic authentic standards (see Supporting Information). Both CLS and TEAS showed enantiomeric excess of one stereoisomer, although both possible enantiomers of **1d**, **2b**, and **3b** are formed.

Putative mechanism of the formation of novel terpenoids

The terpenoid alcohols **1d**, **2b**, and **3b** likely originate from the aromatic diphosphates **1–3** (Figure 1) by an $\text{S}_{\text{N}}2'$ -like (or $\text{S}_{\text{N}}1'$ -like) mechanism. After activating the diphosphate group by the Lewis-acidic Mg^{2+} ions in the active sites of TEAS or CLS, water (or hydroxide) can attack the allylic position of the substrate, expelling inorganic diphosphate and leaving the corresponding alcohol as the product. The mechanism proceeds via the formation of a linalyl cation analogue with synchronized ($\text{S}_{\text{N}}2'$) or later ($\text{S}_{\text{N}}1'$) attack of a water molecule to form the



Scheme 2. Proposed mechanisms for the formation of reaction products from substrates **1–6**. A) Hydrolytic conversion of aromatic diphosphates **1–3** into linalool derivatives. B) Cyclization of thia-homo-GPP **4** into **4b** likely via a transoid 5-exo-trig reaction. C) Formation of **5f** by isomerization of the transoid intermediate into the cisoid neryl-like diphosphate (**5***) by the standard 6-exo-trig-cyclization. D) Cyclization of oxa-homo-FPP substrate **6** into germacrene-like cation **6a⁺**, inhibiting transannular follow-up reactions.

corresponding alcohol from each substrate as described by Degenhardt and co-workers.^[5a]

Due to the high structural similarity with GPP, conversion of **5** with CLS affords propargyl limonene (**5f**) as the main and hoped-for product, as such terminal alkynes are valuable products for biological assays. Alkynes and their tissue distribution can be directly followed without disturbing dyes by spatially resolving CARS (coherent anti-Stokes Raman spectroscopy) microscopy experiments. They can also be dye-labeled by bio-orthogonal reaction with fluorescent azides in a click reaction.^[25] Most likely, cyclization of this substrate proceeds analogously to the mechanism proposed for the natural substrate, which involves a linalyl diphosphate-like intermediate (**5***) and a terpinyl-like cation (**5f⁺**, Scheme 2C).^[26] The same mechanism is described for a similar (-)-limonene synthase from *Mentha* sp. whose crystal structure is already known.^[5b,27] This hypothesis is supported by the high conversion observed for this substrate (Figure 2). However, the triple bond is not involved in the reaction, possibly because of the instability of the vinyl cation that would be formed.

While substrate **8**, the oxygen analogue of **4**, shows no conversion with either enzyme. The higher electronegativity and lower polarizability of oxygen compared to sulfur could hinder the ionization of the organo-diphosphates and destabilize the prenyl cations. TEAS can convert substrate **6**, which is the structurally highly similar oxygen-inserted homolog to its natural substrate FPP, into the germacrene-like product **6a** (Scheme 2D). Interestingly, additional products are almost nonexistent. During the initial reaction step, the germacrene-like cation (**6a⁺**) is produced.^[14] Here the initial positive charge is located at least six bonds away from the oxygen atom. The reaction is completed by deprotonation of **6a⁺**, yielding homo-oxa-germacrene (**6a**). Thus, a premature termination of the natural reaction path occurs, and no further activation of **6a** and further cyclization of the eudesmane-like cation (**6a***) is observed. In contrast to the conformation of the ten-membered ring forming as an intermediate in the cyclization of the natural substrate FPP, the conformation of the corresponding eleven-membered ring from substrate **6** could be inappropriate for further cyclization. Thus **6** can be seen as a probe to verify that the path to eudesmane/aristolochene passes through a medium-sized ring, a ring size that usually is disfavored both kinetically and thermodynamically in chemical cyclizations.^[28] Furthermore, the oxygen atom might have inhibitory effects when in close vicinity (neighboring or transannular) to the positive charge.

The unreactivity of oxa-homo-GPP **8** prompted us to substitute oxygen by sulfur. The electronegativity of sulfur is almost equal to that of a methylene (CH₂) group. Should electronegativity differences be responsible for the non-acceptance of the oxa-analogue **8**, the thia analogue **4** should react, and it did. However, in contrast to the enzyme reactions described above, conversion of **4** does not follow the "natural" path and yields an unexpected cyclization product (**4b**, Scheme 2B), which constitutes 30% of the total. This product is a tetrahydrothiophene derivative that contains two exocyclic double bonds. Possibly, cyclization is initiated by an S_E-like

attack of the initial/developing allylic cation to the proximal carbon of the second double bond in a 5-exo-trig manner that is favored over the six-endo-trig alternative, yielding cation **4b⁺** (Scheme 2B). Finally, enzymatic deprotonation of **4b⁺** affords 3-methyl-3-vinyl-4-isopropenyl tetrahydrothiophene (**4b**) as the product. An isomerization into the cisoid conformation is not required for this path of product generation. An analogous reaction path is not possible for GPP, since it would result in 4-exo-trig or 5-endo-trig cyclizations, both are less favorable to the six-endo-trig path followed in terpinyl cation formation towards limonene.^[6b] However, we reckon that some of the minor products produced may also stem from a cisoid cyclization pathway, where the "natural path" analogous structure III (i.e. 7-exo-trig-cyclization) is the first cyclic cationic intermediate (Scheme 1). Overall, this shows that the enzyme does activate substrate **4**, but following this initiation does not guide product formation very well, as evidenced by formation of many products including stereoisomers. Thus we did not bother to dig deeper into the minor components.

Structure-activity relationship of CLS and TEAS for prenyl diphosphate substrates

TEAS adopts a monomeric structure while CLS acts as a homodimer (Figure 3). As can be seen in Figure 3, both enzymes coordinate three Mg²⁺ ions in the active site for binding, coordination, and activation of substrates. For the binding and coordination of Mg²⁺ ions, in turn, they contain motifs of highly conserved aspartate, glutamate and arginine residues.^[5a] The calculated volume of the active site pocket in CLS (1076 Å³) is much smaller than in TEAS (1940 Å³) (Figure S5). Comparison of the substrate access tunnels shows that CLS has a longer but narrower substrate-binding pocket (bottleneck radius = 1.3 Å, length = 33.5 Å) compared to TEAS (bottleneck radius = 3.2 Å, length = 18.8 Å; Figure S6). Flexibility analysis of CLS and TEAS protein structures (Figure S7) indicated that loops 1–5 forming the entrance tunnels to the substrate-binding pocket are highly flexible which in turn allow the entrance of large(r) substrates but might also lead to the diffusion of water molecules. Short-chain substrates such as DMAPP and **7** are too small to serve as substrates for either enzyme since their cyclization in a relatively large active site pocket would be highly disfavored vs. water attack. On the other hand, more bulky substrates (**9–11** and **GGPP**) even though they are fitting into the catalytic pockets of the enzymes, cannot form a catalytically-competent binding pose or cyclization-competent substrate conformation and thus no conversion was observed. A hydrophobicity analysis of the binding pockets shows that CLS and TEAS have mainly hydrophobic pockets that include a key tyrosine residue for guidance and stabilization of the carbocation intermediate(s) in the active site,^[6b,14] in addition to side chains of aromatic residues like phenylalanine and tryptophan which also were reported to stabilize the charged intermediate(s).^[5a]

To gain molecular insight into the substrate specificity of CLS and TEAS, a mechanism-based substrate docking was

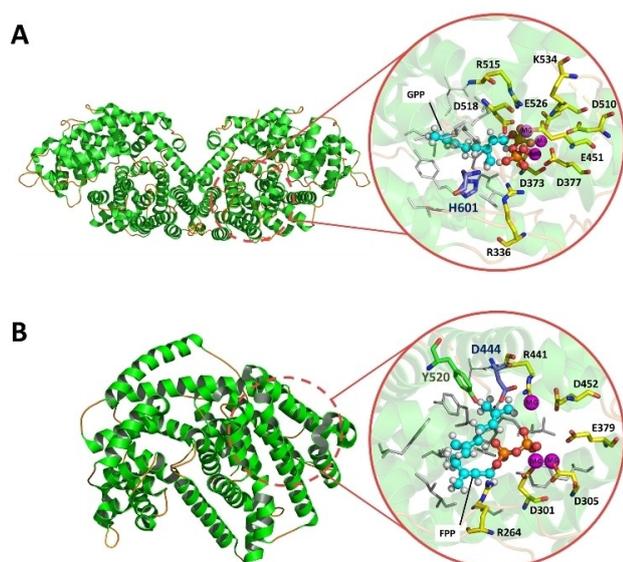


Figure 3. Cartoon model of limonene synthase from *Cannabis sativa* (CLS) (A) and 5-*epi*-aristolochene synthase (TEAS) from *Nicotiana tabacum* (B). Alpha helices in the cartoon representations of the respected models are shown in green. Loops are shown in orange. Zoom-in view of the active site and the amino acid residues that are involved in the substrate binding. GPP and FPP are shown as ball and stick. Residues involved in the GPP and FPP binding sites are highlighted in yellow. Proposed residues acting as general bases for deprotonation of intermediates during the catalysis are colored in blue. Residues acting as proton shuttles for protonation of intermediates are shown in green. Residues involved in binding the carbon moiety of GPP and FPP through hydrophobic interactions are shown in grey. Mg^{2+} ions are colored in magenta. Presentation of the model and the active site was designed using PyMOL Molecular Graphics System, Version 2.3.3 Schrödinger, LLC.

carried out, using all prenyl diphosphate substrates (natural and artificial) and the corresponding carbocations or uncharged intermediates (Figure 3). In our mechanism-based docking, the binding and stabilization of intermediates were analyzed according to the proposed mechanism for CLS and TEAS.^[5b,15,17–19,29] The criteria for choosing the productive docking poses (catalytically competent docking poses) were based on the suitable coordination of substrate or intermediate in the active site, presence of key interactions with amino acid residues expected to be involved in the reaction mechanism, and having the highest docking score.

As can be seen in Figure 3A, the docking pose of GPP in CLS reveals that the diphosphate moiety of the natural substrate is anchored in the Mg^{2+} coordination shell, through an extensive network of hydrogen bonds. This interaction is suggested to be important for the conversion of the substrate, as it plays a crucial role in the activation and ionization of the substrate.^[5] In addition to hydrophobic interactions with amino acids deeper in the active pocket, also an adequate distance (4.4 Å) to His601, with its imidazole-group orientated towards the substrate was observed. This histidine is proposed to act as a general base in a CLS homolog, i.e., limonene synthases, as it is involved in the deprotonation of the terpinyl cation during the last step of the reaction forming the final product.^[20]

A closer look at the docking pose of FPP in TEAS (Figure 3B) shows the network of interaction between Mg^{2+} and the coordination shell through the diphosphate group of substrates, and other additional hydrophobic interactions with amino acids in the deeper part of the active site. Adequate distances to the residues Asp444 (3.9 Å) and Tyr520 (4.4 Å) were identified. These residues are proposed to act as a reaction dyad in the catalytic mechanism.^[15] Asp444 most likely is used for deprotonation of the germacrene cation, whereas Tyr520 acts as a proton shuttle for protonation of germacrene A in the following reaction step.^[17]

Additionally, we checked the carbocation conformations to see whether it adopts a helical conformation bound to a key threonine residue side chain, which is involved in guiding and coordinating the charged intermediate core as has been suggested by Starks et al.^[14]

Our substrate docking simulation (Figure S7) shows that all substrates, except diphosphates **8** and **9** for CLS, can be, in principle, accommodated by the two active site pockets. Visual inspection of substrate docking poses shows that all substrates bind diphosphate through Mg^{2+} ions which in turn bind to highly conserved DDXXD – and DTE- motifs, responsible for Mg -coordination, substrate binding, and activation.^[5a]

To discriminate between convertible and non-convertible substrate binding, we analyzed the docking orientations and interactions contributing to the stabilization of the potential allylic carbocation intermediates, which is also an important consideration for the conversion of substrates according to the usual mechanisms expected for CLS and TEAS.^[5b,17–18,20,29] In general, results for carbocation intermediate docking reveal that, in both enzymes, intermediates in their positively charged form dock deeper into the binding pocket than the corresponding starting substrate, except for the carbocation of substrate **6** in CLS (Figure S8). This resembles prior proposals of the dipole-mediated migration of allylic carbocations deeper into the active site through a group of threonine residues.^[14] Further, we identified a general trend that carbocations of natural and artificial substrates form a helical turn-like conformation when bound deeper within the active site (Figure S11). This geometry could be required, possibly, to orient it better towards the threonine residues which in turn improve dipolar stabilization of the cationic intermediate in a conformation suitable for cyclization. In the same fashion, the migration of the carbocation deeper into the active site could be required for a productive catalytic pose, again improving dipolar stabilization and orientation of the intermediates towards possible catalytically active residues. Thus, when docking of the carbocation is unable to generate such a helical confirmation, an absence or a significant reduction of conversion is to be expected. This helical conformation of the carbocations was observed for substrates **1–5** or **1–6** for CLS and TEAS, respectively. In contrast, substrates **7–11** failed to settle in such docking poses. Carbocations of DMAPP are rigid and thus always cause an extended conformation, thus the completion of the mechanism is unsuccessful.

As an example, docking of the allyl carbocation intermediate derived from substrate **6** in CLS (Figure S8) compared to

docking the natural substrate shows an improper positioning in the substrate-binding pocket of CLS, with the geranyl cation binding significantly deeper in it vs. the oxa-homolog. As a result, the interactions with aromatic residues (W345, H601, Y366, Y595) and threonine T370 are missing, which are required to stabilize the allylic carbocation through dipolar interactions and π stacking.^[5a,14] Instead, the charged intermediate of substrate **6** can form hydrogen bonds with the Mg^{2+} -diphosphate binding moiety through the oxygen of the ether group. The retained interactions with the diphosphate binding complex could explain, why the intermediate cannot migrate deeper into the binding pocket for the following reaction steps, which in turn results in a lacking conversion of substrate **6** by CLS.

In addition to the lack of a double bond at C10, necessary to complete the cyclization mechanism by TEAS, our carbocation docking shows that the geranyl cation is located in the binding pocket in an extended conformation, compared to other substrates. A helical conformation can bring the C1 and C10 carbon in close proximity. These structural and conformational properties account for the low-yielding formation of acyclic monoterpenes, mainly linalool, during the conversion of GPP by TEAS. Carbocation docking of substrate **6** in TEAS revealed a shorter distance between its C1 and C10 (4.2 Å) than that between the corresponding carbons of FPP (5.6 Å; Figure S13), which might explain the higher conversion rate of this substrate by TEAS vs. FPP. In principle, also this intermediate should be able to react further in a transannular cyclization, but likely electronic effects forbid this, which cannot be solved by pure MM calculations.

A higher conversion rate for compound **5** was observed compared to the natural substrate GPP in CLS (Figure 2). Docking poses of α -terpinyl cations of GPP and compound **5** into the active site of CLS with diphosphate present in the Mg^{2+} coordination shell identified interactions of intermediates with the proposed base His601. Closer inspection of the docking pose shows that the terpinyl cation of GPP maintains an adequate distance between C9 and πN of His601 (3.8 Å) for deprotonation. However, for compound **5**, the terpinyl cation showed the same distance between C9 and πN of His601, which in turn directly interacts with the retained diphosphate moiety (5.4 Å) (Figure S14). It is proposed that diphosphate migrates towards His601 after protonation of the imidazole side chain through this interaction, thereby abstracting the proton from histidine and completing the deprotonation to yield the final product.^[20] Therefore, deprotonation through the πN could be more optimal for the deprotonation process, as the added proton can be directly abstracted using the diphosphate. In contrast, deprotonation through πN of His601 would additionally require transfer of the added proton to τN for completion of deprotonation. This might explain the higher conversion rate for compound **5** compared to the natural substrate GPP in CLS.

As identified during the pocket analysis of both enzymes, the active pocket of TEAS (1940 Å³) is far more spacious than the active pocket of CLS (1076 Å³). Therefore, TEAS can accept sterically more demanding substrates, such as compounds **1**, **2**, and **3** with their rigid benzene groups. A larger active pocket

enables TEAS to offer more flexibility for these compounds to adapt to a catalytically-competent binding pose or cyclization-competent substrate conformation, leading to a more sufficient cyclization and higher conversion compared to CLS.

Conclusion

Utilizing nature's toolbox to produce complex chemical structures from simple, achiral linear precursors is an elegant way to novel terpenoids and allows access to unusual heterocyclic or labelled terpenoid analogues. The two terpene cyclases used as first model enzymes were CLS and TEAS as they express high promiscuity towards a range of different allylic diphosphates. Eleven extended and oxygen- or sulfur-inserted substrates were tested, six were converted and consequently resulted in novel terpenoids of which six were identified and three studied in more depth. Some products possess distinctive and peculiar smells, while a terminal alkyne derivative opens new possibilities for tracking and labelling. Depending on the substrate, different modes of cyclization are proposed, always following the Baldwin rules and corroborated by molecular modelling studies. By these, the properties of the active site for both enzymes were explored and an ON-OFF explanation for substrate conversion was suggested. Artificial prenyl diphosphate substrates were shown to fit into the active sites of both CLS and TEAS similar to natural substrates; however, for the cyclization, the conformations and stabilization details of the carbocation intermediates appear to be decisive.

We envision the conversion of such substrates with further terpene synthases, which in combination with professional olfactory analyses by flavorists can reveal the full potential of these new hetero-terpenoids. The proposed substrates and processes thus might be useful in fundamental research as well as for industrial applications.

Experimental Section

General remarks: All commercially available chemicals were used without further purification. Syntheses of products **1 d**, **2 b**, and **3 b** were performed according to a method modified after Martin et al. using vinyloxirane with the corresponding thiophenol in a basic medium.^[30] Substrates were synthesized as published elsewhere.^[31] An alternative route to some substrates can also be found in Kirschning's work.^[24]

Bacterial strains, vectors, and chemicals: The expression vector pET101/D-TOPO (Invitrogen, Germany) containing the gene encoding for CLS (pET101-CLS) truncated after amino acid 60 was a kind gift from Prof. Toni Kutchan (Danforth Center, St. Louis, USA).^[10] The plasmid pET-28b(+) containing TEAS (pET28b-TEAS) was a kind gift from Prof. Joe Chappell (University of Kentucky, USA).^[32] *E. coli* strains BL21 (DE3) and BL21 Star™ (DE3) (Invitrogen, Karlsruhe, Germany) were used for gene expression of CLS and TEAS, respectively.

Protein expression and purification: Overexpression and purification of CLS were carried out according to a protocol published previously.^[10] Overexpression and purification of TEAS were performed according to O'Maille et al. (2004) but lysis buffer (25 mM

Tris/HCl, 150 mM NaCl, pH 7.5), wash buffer (25 mM Tris/HCl, 150 mM NaCl, 30 mM imidazole, pH 7.5) and elution buffer (25 mM Tris/HCl, 150 mM NaCl, 300 mM imidazole, pH 7.5) were changed accordingly.^[33] Fractions containing TEAS protein were checked for homogeneity by SDS-PAGE.^[34] Homogeneous protein was pooled and desalted on a PD-10 desalting column (GE Healthcare, USA). The purified protein, which was flash-frozen in liquid nitrogen, was stored in TEAS assay buffer (50 mM HEPES/NaOH, 100 mM NaCl, 20 mM MgCl₂, 1 mM DTT, 10% (v/v) glycerol, pH 7.4) at -20 °C until further use. Protein concentrations were estimated using Bradford reagent according to the protocol of the manufacturer (Roth, Germany).

Determination of CLS and TEAS product spectrum: Reactions (500 µL) containing TEAS assay buffer (50 mM HEPES/NaOH, 100 mM NaCl, 20 mM MgCl₂, 1 mM DTT, pH 7.5) or CLS assay buffer (10 mM MOPSO/NaOH, 20 mM MgCl₂, 1 mM DTT, pH 7.0), 100 µg/mL of enzyme and 0.2 mM of prenyl diphosphate were performed in screw-capped glass vials. The assay mixture was overlaid with 200 µL of organic solvent (*n*-hexane/*n*-heptane (1:1, v/v) containing 25 µM of naphthalene as internal standard). After 3 hours of incubation at 22 °C (TEAS) or 30 °C (CLS), products were extracted by being vigorously vortexed for 30 s. The organic phase was analyzed by coupled gas chromatography/mass spectrometry (GC-MS) as described below.

Preparative scale enzymatic conversion: 2 mg of the substrate was incubated with 1 mg CLS or TEAS in assay buffer (3 mL). The assay mixture was overlaid with 500 µL organic solvent (*n*-hexane/*n*-heptane (1:1, v/v)) and incubated for 15 h at 22 °C (TEAS) or 30 °C (CLS). After the addition of 4 M urea, products were extracted by being vigorously vortexed for 30 s. Reactions were performed in triplicates and the organic phases were pooled. Products in the organic phase of substrate **4** were separated via preparative GC. For preparative reactions of substrates **5** and **6**, the solvent was carefully evaporated with a stream of nitrogen. This afforded 0.8 mg (20%) of biotransformation products of **5f** as a clear oil, as well as 0.5 mg (8.3%) of compound **6a** as an odorous clear oil. The products were solved in CDCl₃ and analyzed by a combination of 1D and 2D-NMR spectroscopy experiments. ¹H, ¹³C NMR spectra were recorded in CDCl₃ at 22 °C at 600 and 150 MHz respectively.

GC/MS analysis of enzymatic products: GC/MS analysis was carried out on a Shimadzu GCMS-QP2010 Ultra (Shimadzu, Japan) with the following settings: 70 eV electron-impact (EI) ionization, source temperature 200 °C, column ZB-5MS (Zebtron, 30 m × 0.25 mm × 0.25 µm), injector temperature 220 °C, interface temperature 300 °C, carrier gas helium, flow rate 1.1 mL min⁻¹, injection volume 1 µL, splitless injection. The temperature program was started at 40 °C and ramped at 10 °C min⁻¹ to a final temperature of 300 °C. The scan rate for mass spectra in the range of 50–300 u was 909 u s⁻¹.

Chiral GC/MS was performed on the above-mentioned GC/EI-MS system with the following settings: 70 eV electron-impact (EI) ionization, source temperature 200 °C, chiral column (Macherey-Nagel, HYDRODEX®-β-6TBDM, 25 m × 0.25 mm), injector temperature 220 °C, interface temperature 250 °C, carrier gas helium, flow rate 1 mL min⁻¹, injection volume 1 µL, splitless injection. The temperature program was started at 40 °C and ramped at 10 K min⁻¹ to 60 °C followed by a temperature increase to 230 °C with 10 K min⁻¹. The scan rate for mass spectra in the range of 50–400 u was 1250 u s⁻¹.

Products were identified by their expected aliphatic molecular masses ($M = M_{\text{Substrate}} - 228.02$, where 228.02 is the molecular mass of the leaving group H₁₂N₃O₂P₂) or the molecular masses of the corresponding alcohols ($M + 16$). Retention indices (RI) of the eluting compounds were calculated after calibration with a C8-C20

alkane standard mix (Sigma-Aldrich, Germany). For compound identification, RI values and data from mass spectra were compared with the NIST 17 EPA/NIH Mass Spectral Database by the program provided by the manufacturer (Shimadzu, Japan) and by AMDIS 2.6 (<http://amdis.software.informer.com/2.6>). Additionally, authentic standards, if available, were compared with the samples to verify compound identity.

Preparative GC separation of compound 4b: For elucidation of products obtained by conversion of the artificial prenyl diphosphates, preparative GC was performed to isolate the corresponding compounds. The separation of the compound mixtures was carried out on an Agilent 6890N GC-system (Agilent, Böblingen, Germany): flame ionization detector (FID), detector temperature 300 °C, column HP-5 (Agilent, 30 m × 0.32 mm × 0.25 µm), splitless injection of 2 µL sample via Gerstel Cooled Injection System, injector temperature 60 °C ramped to 260 °C in 30 seconds, carrier gas helium, flow 2.3 mL min⁻¹. The oven temperature started at 60 °C, was ramped to 180 °C at 9 °C min⁻¹, then to 300 °C at 30 °C min⁻¹ and was kept constant for 5 minutes. The GC-system was coupled to a Gerstel Preparative Fraction Collector (PFC) (Gerstel, Germany) collecting eluting target compound in a cooling trap. For NMR spectroscopy, compounds were dissolved in 700 µL of CDCl₃.

Determination of the optical rotation of compound 6a: The optical rotation was measured in CHCl₃ in a 1 mL quartz glass cuvette (50 mm × 3 mm) on a Jasco P-2000 polarimeter at 598 nm with a concentration of 0.04 w/v % at 25 °C and the specific optical rotation was calculated.

Protein modeling and molecular docking: X-ray crystal structure of 5-epi-aristolochene synthase (TEAS) from *Nicotiana tabacum* (PDB ID: 3M01; 1.85 Å^[35]) was taken from RCSB PDB. TEAS adopts a monomer structure. Crystal structure of CLS (limonene synthase from *Cannabis sativa*) is not available yet, therefore, a homology model was built using the YASARA Structure Version 20.12.24.L.64^[36]. The closest homolog on which the model was mainly built was a (4S)-limonene synthase from *Mentha spicata* (PDB ID: 2ONG; 2.70 Å), with a coverage of 87% and a sequence identity of 43.9%. A hybrid model with an overall quality Z-score of -1.860 was obtained. The modeled CLS structure is a homodimer. We only used chain A for further modeling. Further, models were subjected to energy minimization using AMBER14 force field.^[37] The enzymes' active site and tunnels were computed and localized by HotSpot Wizard 3.0^[38] and MoleOnline^[39] with default settings. Molecular docking of substrates and corresponding intermediates was performed using Molecular Operating Environment (MOE)^[40] with an induced fit method. The docking simulations were further analyzed using PyMOL 2.3.3 (Schrödinger, LLC.)^[41] and LigPlot⁺^[42] to generate 3D and 2D enzyme-substrate interaction diagrams, respectively. For model visualization and analysis, PyMOL 2.3.3 (Schrödinger, LLC.)^[41] was applied.

Acknowledgments

We thank Dr. A. Porzel, G. Hahn (NMR, LC), Dr. J. Schmidt, M. Lerbs and M. Süße (MS, GC) for support, and Dr. Nils Günnewich and Prof. Toni Kutchan for providing GLS plasmid from *Cannabis sativa*, and Prof. Joe Chappell for TEAS plasmid. Part of this work was supported by the EU Project BioNexGen (FP7-KBBE-2010; no. 266025). Open Access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

Keywords: biotransformation · medium-sized rings · molecular docking · terpene synthases · unnatural terpenoids

- [1] Z.-Y. Huang, R.-Y. Ye, H.-L. Yu, A.-T. Li, J.-H. Xu, *Bioresour. Bioprocess.* **2021**, *8*, 66.
- [2] a) L. A. Wessjohann, J. Keim, B. Weigel, M. Dippe, *Curr. Opin. Chem. Biol.* **2013**, *17*, 229–235; b) D. Cox-Georgian, N. Ramadoss, C. Dona, C. Basu, in *Medicinal Plants: From Farm to Pharmacy*, Springer **2019**, 333–359.
- [3] a) J. B. Sharmeen, F. M. Mahomoodally, G. Zengin, F. Maggi, *Molecules* **2021**, *26*, 666; b) R. B. Croteau, E. M. Davis, K. L. Ringer, M. R. Wildung, *Naturwissenschaften* **2005**, *92*, 562–577; c) W. Schwab, R. Davidovich-Rikanati, E. Lewinsohn, *Plant J.* **2008**, *54*, 712–732.
- [4] P. Kraft, J. A. Bajgrowicz, C. Denis, G. Fráter, *Angew. Chem. Int. Ed.* **2000**, *39*, 2980–3010; *Angew. Chem.* **2000**, *112*, 3106–3138.
- [5] a) J. Degenhardt, T. G. Köllner, J. Gershenzon, *Phytochemistry* **2009**, *70*, 1621–1637; b) D. W. Christianson, *Chem. Rev.* **2017**, *117*, 11570–11648.
- [6] a) V. Harms, A. Kirschning, J. S. Dickschat, *Nat. Prod. Rep.* **2020**, *37*, 1080–1097; b) W. Brandt, L. Bräuer, N. Günnewich, J. Kufka, F. Rausch, D. Schulze, E. Schulze, R. Weber, S. Zakharova, L. A. Wessjohann, *Phytochemistry* **2009**, *70*, 1758–1775; c) D. Tholl, *Curr. Opin. Plant Biol.* **2006**, *9*, 297–304.
- [7] a) H. F. Schreckenbach, G. N. Kaluderović, L. A. Wessjohann, in *Biocatalysis in Organic Synthesis, Vol. 2* (Ed.: W.-D. Fessner, K. Faber, N. J. Turner), Thieme, Stuttgart, **2015**, pp. 177–211; b) L. Wessjohann, S. Zakharova, D. Schulze, J. Kufka, R. Weber, L. Bräuer, W. Brandt, *Chimia* **2009**, *63*, 340.
- [8] a) L. Heide, *Curr. Opin. Chem. Biol.* **2009**, *13*, 171–179; b) Y. Gao, R. B. Honzatko, R. J. Peters, *Nat. Prod. Rep.* **2012**, *29*, 1153–1175; c) M. Nagia, M. Gaid, E. Biedermann, T. Fiesel, I. El-Awaad, R. Hänsch, U. Wittstock, L. Beerhues, *New Phytol.* **2019**, *222*, 318–334.
- [9] a) C. A. Cotrim, A. Weidner, N. Strehmel, T. B. Bisol, D. Meyer, W. Brandt, L. A. Wessjohann, M. T. Stubbs, *ChemistrySelect* **2017**, *2*, 9319–9325; b) S. Zakharova, M. Fulhorst, L. Łuczak, L. Wessjohann, *Arkivoc* **2004**, *13*, 79–96; c) L. Wessjohann, B. Sontag, *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1697–1699.
- [10] N. Günnewich, J. E. Page, T. G. Köllner, J. Degenhardt, T. M. Kutchan, *Nat. Prod. Commun.* **2007**, *2*, 1934578X0700200301.
- [11] U. Vögeli, J. W. Freeman, J. Chappell, *Plant Physiol.* **1990**, *93*, 182–187.
- [12] J. P. Noel, N. Dellas, J. A. Faraldos, M. Zhao, B. A. Hess, L. Smentek, R. M. Coates, P. E. O'Maille, *ACS Chem. Biol.* **2010**, *5*, 377–392.
- [13] P. E. O'Maille, J. Chappell, J. P. Noel, *Arch. Biochem. Biophys.* **2006**, *448*, 73–82.
- [14] C. M. Starks, K. Back, J. Chappell, J. P. Noel, *Science* **1997**, *277*, 1815–1820.
- [15] F. Zhang, N. Chen, J. Zhou, R. Wu, *ACS Catal.* **2016**, *6*, 6918–6929.
- [16] a) D. J. Tantillo, *Nat. Prod. Rep.* **2011**, *28*, 1035–1053; b) D. J. Miller, R. K. Allemann, *Nat. Prod. Rep.* **2012**, *29*, 60–71.
- [17] F. Zhang, T. An, X. Tang, J. Zi, H.-B. Luo, R. Wu, *ACS Catal.* **2020**, *10*, 1470–1484.
- [18] F. Zhang, Y.-H. Wang, X. Tang, R. Wu, *Phys. Chem. Chem. Phys.* **2018**, *20*, 15061–15073.
- [19] F. Zhang, N. Chen, R. Wu, *J. Chem. Inf. Model.* **2016**, *56*, 877–885.
- [20] J. Yao, F. Chen, H. Guo, *Mol. Simul.* **2018**, *44*, 1158–1167.
- [21] a) J. A. Faraldos, Y. Zhao, P. E. O'Maille, J. P. Noel, R. M. Coates, *ChemBioChem* **2007**, *8*, 1826–1833; b) O. Cascón, S. Touchet, D. J. Miller, V. Gonzalez, J. A. Faraldos, R. K. Allemann, *Chem. Commun.* **2012**, *48*, 9702–9704; c) L. A. Johnson, A. Dunbabin, J. C. R. Benton, R. J. Mart, R. K. Allemann, *Angew. Chem. Int. Ed.* **2020**, *59*, 8486–8490; *Angew. Chem.* **2020**, *132*, 8564–8568; d) F. Huynh, D. J. Grundy, R. L. Jenkins, D. J. Miller, R. K. Allemann, *ChemBioChem* **2018**, *19*, 1834–1838.
- [22] a) M. Nagaki, T. Koyama, T. Nishino, K. Shimizu, Y. Maki, K. Ogura, *Chem. Lett.* **1997**, *26*, 497–498; b) M. Nagaki, S. Sato, Y. Maki, T. Nishino, T. Koyama, *J. Mol. Catal. B* **2000**, *9*, 33–38; c) M. Nagaki, H. Yamamoto, A. Takahashi, Y. Maki, J. Ishibashi, T. Nishino, T. Koyama, *J. Mol. Catal. B* **2002**, *17*, 81–89.
- [23] C. Nájera, J. M. Sansano, *Tetrahedron* **1994**, *50*, 5829–5844.
- [24] C. Oberhauser, V. Harms, K. Seidel, B. Schröder, K. Ekramzadeh, S. Beutel, S. Winkler, L. Lauterbach, J. S. Dickschat, A. Kirschning, *Angew. Chem. Int. Ed.* **2018**, *57*, 11802–11806; *Angew. Chem.* **2018**, *130*, 11976–11980.
- [25] a) Kenry, B. Liu, *Trend. Chem.* **2019**, *1*, 763–778; b) C. J. Pickens, S. N. Johnson, M. M. Pressnall, M. A. Leon, C. J. Berkland, *Bioconjugate Chem.* **2018**, *29*, 686–701.
- [26] D. J. McGarvey, R. Croteau, *Plant Cell* **1995**, *7*, 1015–1026.
- [27] D. C. Hyatt, B. Youn, Y. Zhao, B. Santhamma, R. M. Coates, R. B. Croteau, C. Kang, *PNAS* **2007**, *104*, 5360–5365.
- [28] a) L. A. Wessjohann, E. Ruijter, D. Garcia-Rivera, W. Brandt, *Mol. Diversity* **2005**, *9*, 171–186; b) L. A. Wessjohann, R. Bartelt, W. Brandt, in *Practical Medicinal Chemistry with Macrocycles* (Eds.: E. Marsault, M. L. Peterson), Wiley, **2017**, pp. 77–100.
- [29] J. Degenhardt, T. G. Köllner, J. Gershenzon, *Phytochemistry* **2009**, *70*, 1621–1637.
- [30] G. Martin, J. Sauleau, M. David, A. Sauleau, S. Sinbandhit, *Can. J. Chem.* **1992**, *70*, 2190.
- [31] R. A. Weber, in *Synthese und biokatalytische Umsetzung von Prenyldiphosphaten*. Dissertation Thesis, Martin Luther University Halle-Wittenberg (with Leibniz Institute of Plant Biochemistry), **2011**.
- [32] J. R. Mathis, K. Back, C. Starks, J. Noel, C. D. Poulter, J. Chappell, *Biochemistry* **1997**, *36*, 8340–8348.
- [33] P. E. O'Maille, J. Chappell, J. P. Noel, *Anal. Biochem.* **2004**, *335*, 210–217.
- [34] U. K. Laemmli, *Nature* **1970**, *227*, 680–685.
- [35] J. P. Noel, N. Dellas, J. A. Faraldos, M. Zhao, B. A. Hess Jr, L. Smentek, R. M. Coates, P. E. O'Maille, *ACS Chem. Biol.* **2010**, *5*, 377–392.
- [36] E. Krieger, G. Koraimann, G. Vriend, *Proteins* **2002**, *47*, 393–402.
- [37] D. Case, V. Babin, J. Berryman, R. Betz, Q. Cai, D. Cerutti, T. Cheatham III, T. Darden, R. Duke, H. Gohlke, *Amber* **2014**, *14*, 29–31.
- [38] L. Sumbalova, J. Stourac, T. Martinek, D. Bednar, J. Damborsky, *Nucleic Acids Res.* **2018**, *46*, W356–W362.
- [39] K. Berka, O. Hanák, D. Sehnal, P. Banaš, V. Navratilova, D. Jaiswal, C.-M. Ionescu, R. Svobodová Vařeková, J. Koča, M. Otyepka, *Nucleic Acids Res.* **2012**, *40*, W222–W227.
- [40] Molecular Operating Environment (MOE), 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3 A 2R7, 2022.
- [41] The PyMOL Molecular Graphics System, Schrödinger, LLC.
- [42] R. A. Laskowski, M. B. Swindells, *J. Chem. Inf. Model.* **2011**, *51*, 2778–2786.

Manuscript received: April 14, 2022
Revised manuscript received: August 26, 2022
Version of record online: September 29, 2022