

# **New Insights into Terpene Synthase Catalysis for Rational Enzyme Engineering**

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

**Houchao Xu**

aus

Ganzhou

Bonn 2023



Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät  
der Rheinischen Friedrich-Wilhelms-Universität Bonn

Gutachter\*in/Betreuer\*in: Prof. Dr. Jeroen S. Dickschat

Gutachter\*in: Prof. Dr. Dirk Menche

Gutachter\*in: Prof. Dr. Daniel B. Werz

Tag der Promotion: 20.10.2023

Erscheinungsjahr: 2023



## Acknowledgements

First of all, I am very grateful to Prof. Dr. Dickschat for his great help during my PhD study. He is a very self-disciplined and hard-working scientist from whom I learnt how to think professionally and how to perform experiments efficiently. I also appreciate the chance to work in AK Dickschat where I harvested fruitful research publications. Jeroen, thank you very much for being my supervisor, my first German teacher and my good friend. Days will never be forgotten when we had our discussions about my running projects and you can always provide helpful suggestions. Days will never be forgotten when we talked with each other in simple German on the way to Mensa. Days will never be forgotten when we cooked together for our group and you told me a lot about wine, food, cycling...

I also want to thank Prof. Dr. Dirk Menche for spending time reading and refereeing my thesis. Thanks also go to Prof. Dr. Robert Glaum and Prof. Dr. Matthias Wüst for joining the thesis examination committee.

Many thanks to my colleagues with whom I stay for most of the time during my PhD study. Especially I would like to say thanks to Lukas Lauterbach for offering me a lot of help when I started my work in 2019. In addition, I also want to thank Dr. Jan Rinkel, Dr. Anwei Hou, Dr. Zhiyang Quan, Anuj Chhalodia, Dr. Geng Li, Zhiyong Yin, Dr. Lin-Fu Liang, Heng Li, Dr. Binbin Gu, Kizerbo Taizoumbe, Alain Tumma, Dr. Georges Bellier Tabekoueng, Dr. Fulong Lin, Neran Reuber and Alina Oseguera Jaramillo for the time we spent together in the lab.

I also express my gratitude to our collaboration partners with whom we published several great papers: Prof. Dr. Bernd Goldfuss, PD. Dr. Tobias G. Köllner, Prof. Dr. Heike Brötz-Oesterhelt, PD Dr. Bertolt Gust, Prof. Dr. Feng Chen, Prof. Dr. Ruiibo Wu, Prof. Dr. Ping-Hua Sun, Prof. Dr. Yong-Heng Wang, Prof. Dr. Russell J. Cox, Prof. Dr. Ming Ma and Prof. Dr. Donghui Yang.

In addition, I need to thank all the analytical facilities especially the NMR department for measuring numerous samples efficiently.

Finally, I want to thank my parents, my wife, my son, my sister and my brother for their big support as always.



## **Preamble**

This cumulative doctoral thesis “New Insights into Terpene Synthase Catalysis for Rational Enzyme Engineering” includes 21 chapters. First of all, a general introduction of natural products, terpenoids and terpene biosynthesis is presented in Chapter 1. Chapters 2–20 show brief summaries of all the publications for this thesis. Specifically, Review articles are introduced in Chapters 2–4 and 20. Original research studies on the biosynthesis of sesquiterpenes, diterpenes and sesterterpenes are presented in Chapters 5–19. Corresponding publications are all attached in Appendices A–S. Chapter 21 provides a summary and outlook based on the projects I have worked on.





## Table of contents

1	The history of research on natural products	1
2	Terpene biosynthesis	4
2.1	Sesquiterpene biosynthesis	9
2.1.1	Sesquiterpenes formed by 1,10-cyclisation	13
2.1.2	Sesquiterpenes formed by 1,11-cyclisation	15
2.1.3	Sesquiterpenes formed by 1,6- and 1,7-cyclisations	16
2.1.4	Non-canonical terpene cyclisations	16
2.2	Diterpene biosynthesis	18
2.3	Sesterterpene biosynthesis	19
2.4	Structural biology of terpene synthases	21
3	Germacrene A – A Central Intermediate in Sesquiterpene Biosynthesis	25
4	Hedycaryol – Central Intermediates in Sesquiterpene Biosynthesis, Part II	31
5	Germacrene B – A Central Intermediate in Sesquiterpene Biosynthesis	37
6	Isotopic Labeling Experiments Solve the Hedycaryol Problem	43
7	Mechanistic Divergence between (4 <i>S</i> ,7 <i>R</i> )-Germacrene-(1(10) <i>E</i> ,5 <i>E</i> )-dien-11-ol Synthases from <i>Dictyostelium purpureum</i> and <i>Streptomyces coelicolor</i>	49
8	Catalytic Role of Carbonyl Oxygens and Water in Selinadiene Synthase	53
9	Isoishwarane synthase from <i>Streptomyces lincolnensis</i>	59
10	The Sesquiterpene Synthase PtTPS5 Produces (1 <i>S</i> ,5 <i>S</i> ,7 <i>R</i> ,10 <i>R</i> )-Guaia-4(15)-en-11-ol and (1 <i>S</i> ,7 <i>R</i> ,10 <i>R</i> )-Guaia-4-en-11-ol in Oomycete-Infected Poplar Roots	65

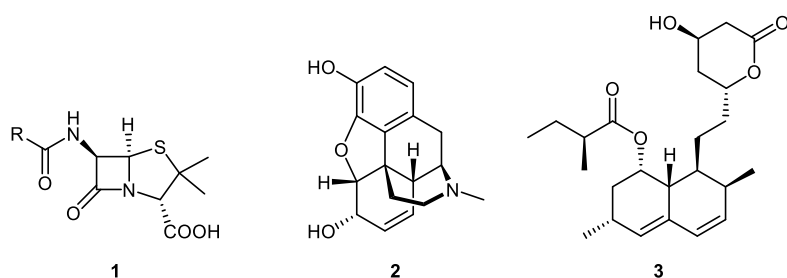
11	1,2- or 1,3-Hydride Shifts: What Controls Guaiane Biosynthesis?	71
12	The enzyme mechanism of patchoulol synthase	77
13	Stereochemical characterisation of the non-canonical $\alpha$ -humulene synthase from <i>Acremonium strictum</i>	83
14	Biosynthesis of the Sesquiterpene Kitaviridene through Skeletal Rearrangement with Formation of a Methyl Group Equivalent	89
15	Fragmentation and [4 + 3] cycloaddition in sodorifen biosynthesis	95
16	A Detailed View on Geosmin Biosynthesis	101
17	Revision of the Cyclisation Mechanism for the Diterpene Spiroviolene and Investigations on its Mass Spectrometric Fragmentation	107
18	Mechanistic Characterisation of the Bacterial Sesterviridene Synthase from <i>Kitasatospora viridis</i>	113
19	Structural Insights into Three Sesquiterpene Synthases for the Biosynthesis of Tricyclic Sesquiterpenes and Chemical Space Expansion by Structure-Based Mutagenesis	119
20	Crystal Structure Based Mutagenesis of Cattleylene Synthase Leads to the Generation of Rearranged Polycyclic Diterpenes	125
21	Mechanistic Investigations on Microbial Type I Terpene Synthases through Site-Directed Mutagenesis	131
22	Summary and outlook	137
23	References	143
24	Appendices A–S	153

# Chapter 1

## Introduction

### 1. The history of research on natural products

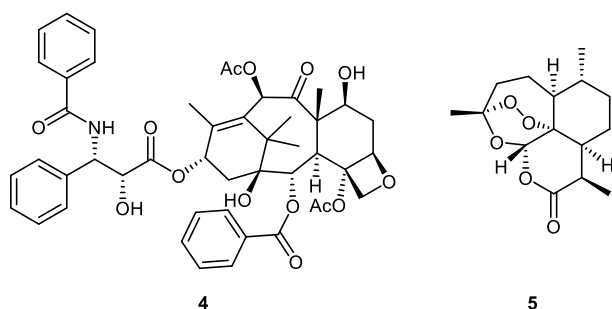
Natural products are compounds, including primary and secondary metabolites, produced in living organisms. Primary metabolites, e.g. nucleotides, carbohydrates, amino acids and vitamins, play essential roles in maintaining the normal growth, development and reproduction of organisms, while secondary metabolites are usually generated by certain organisms to function in ecological interactions, or to protect themselves from the attack of their predators and pathogens. The first natural product studies were performed on plants. In the middle of the 19th century, Justus von Liebig extended the range to animals. Later on, micro-organisms were also regarded as precious sources to produce natural products due to the discovery of penicillin (**1**, Figure 1) from *Penicillium* moulds by the Scottish scientist Alexander Fleming in 1928.<sup>[1][2]</sup>



**Figure 1.** Structures of bioactive natural products representing different classes. Penicillin (**1**) is a non-ribosomal peptide, morphine (**2**) is an alkaloid, and lovastatin (**3**) is a representative of polyketides.

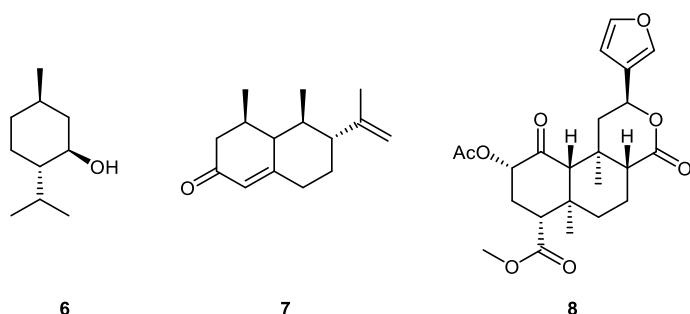
Before the isolation technology was developed, natural products appeared as a mixture used as traditional medicines and essential oils. After the first pure natural product, morphine (**2**, Figure 1), was obtained by the German pharmacist Friedrich Sertürner in 1805,<sup>[1]</sup> numerous compounds have been isolated, purified and structurally elucidated. According to the early concept of “vitalism”, natural products were made by a mysterious life force from nature, meaning that these kinds of compounds could not be produced artificially. However, in 1828, Friedrich Wöhler succeeded in synthesizing urea,<sup>[3]</sup> a natural

product discovered in urine, which was the first scientific experiment to falsify this vitalism idea.



**Figure 2.** Structures of bioactive terpenoids: Paclitaxel (**4**) and artemisinin (**5**).

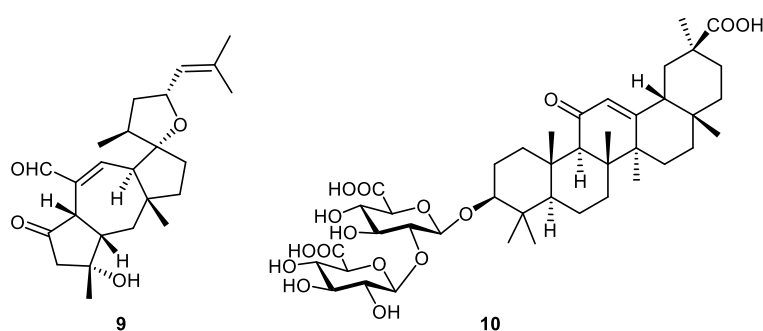
Many secondary metabolites show therapeutic effects and are used as drugs or sources of drug discovery. Examples include, but are not limited to penicillin (**1**, Figure 1), one of the earliest antibiotics discovered from the fungus *Penicillium notatum*,<sup>[4]</sup> morphine (**2**, Figure 1), a potent analgesic obtained from the opium poppy *Papaver somniferum*,<sup>[5]</sup> and lovastatin (**3**, Figure 1), a cholesterol-lowering medication derived from the fungus *Aspergillus terreus*.<sup>[6]</sup>



**Figure 3.** Structures of representative monoterpene (**6**), sesquiterpene (**7**) and diterpene (**8**).

Classes of secondary metabolites include non-ribosomal peptides such as **1**, alkaloids as exemplified by **2** and polyketides as represented by **3**. Terpenoids are the largest class of natural products found abundantly in plants, as well as in fungi, bacteria, and animals. Up to present, approximately 100,000 terpenoids with diverse skeletons have been discovered. Terpenoids contribute to the distinct scents and tastes of fruits, spices, and herbs. These compounds enhance the sensory experience and enjoyment of foods. Terpenoids also

exhibit a wide range of pharmaceutical and industrial functions. Apart from the potential to be developed into drugs such as paclitaxel (**4**, Figure 2) used in cancer treatment,<sup>[7][8]</sup> and artemisinin (**5**, Figure 2) applied to cure malaria,<sup>[9]</sup> terpenoids are also major components of essential oils used in aromatherapy, perfumes, personal care products, and as natural flavourings. Besides those benefits to humans, terpenoids also serve various functions in plants such as natural defences against herbivores, insects, and pathogens, and attracting pollinators such as bees and butterflies to help in the process of pollination and reproduction.<sup>[10]</sup>

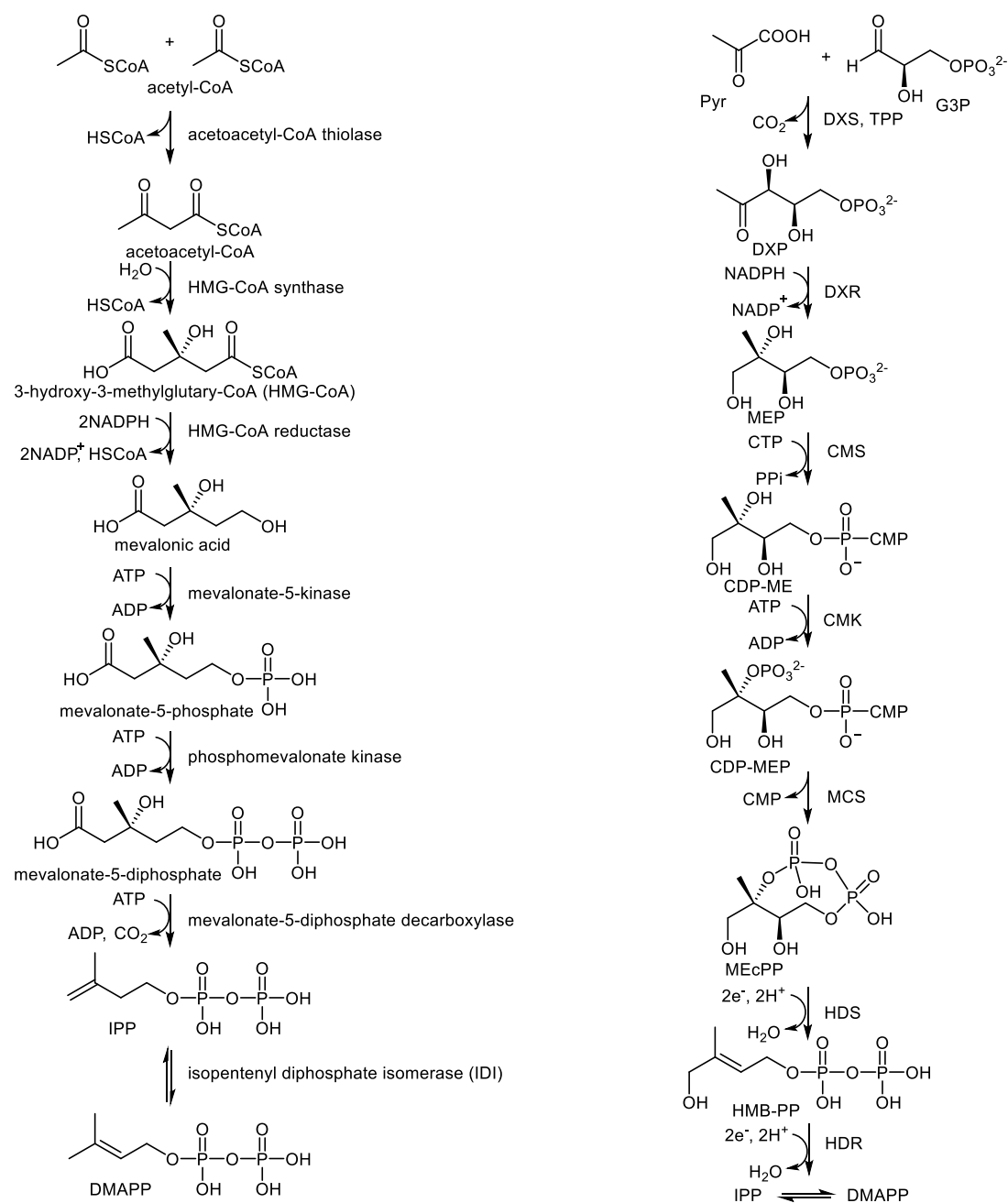


**Figure 4.** Structures representative sesterterpenoid ophiobolin A (**9**) and triterpenoid glycyrrhizic acid (**10**).

Terpenoids are composed of hemiterpenoids (C<sub>5</sub>), monoterpenoids (C<sub>10</sub>), sesquiterpenoids (C<sub>15</sub>), diterpenoids (C<sub>20</sub>), sesterterpenoids (C<sub>25</sub>), triterpenoids (C<sub>30</sub>), etc. For instance, (–)-menthol (**6**, Figure 3) is a monoterpenoid found in peppermint (*Mentha piperita*) or other mint plants.<sup>[11]</sup> This compound is known for its cooling sensation that can be used as a pain relief to throat irritation.<sup>[12]</sup> Nootkatone (**7**, Figure 3) was characterised as a sesquiterpenoid that is a dominant constituent contributing to the smell of grapefruits.<sup>[13][14]</sup> As a representative of diterpenoids, salvinorin A (**8**, Figure 3) isolated from *Salvia divinorum* functions as a potent hallucinogen.<sup>[15][16][17]</sup> In addition, ophiobolin A (**9**, Figure 4), a sesterterpenoid found in *Bipolaris oryzae*,<sup>[18]</sup> was investigated as a potential treatment of rhabdomyosarcoma (RD).<sup>[19]</sup> Last but not least, glycyrrhizic acid (**10**, Figure 4) obtained as a triterpenoid from *Glycyrrhiza glabra*,<sup>[20]</sup> is widely applied in foods as a natural sweetener.<sup>[21]</sup>

## 2. Terpene biosynthesis

Terpenes refer to hydrocarbons or alcohols that are the direct products of terpene synthases, while terpenoids are usually oxidised or otherwise modified terpenes. Therefore, terpenes serve as the precursors of terpenoids biosynthetically. Since the biosynthesis of terpenes will be in the focus of this thesis, the term terpene will be mainly discussed and used in the following chapters.

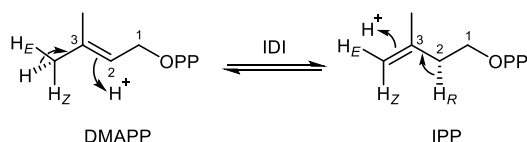


**Scheme 1.** Mevalonate pathway (left) and non-mevalonate pathway (right) towards IPP and DMAPP.

Otto Wallach pointed out that terpenes are constructed from two or more isoprene units,<sup>[22]</sup> which is known as the isoprene rule nowadays. Biosynthetically, terpenes originate from two basic C<sub>5</sub> units, the electrophile dimethylallyl pyrophosphate (DMAPP) and the nucleophile isopentenyl pyrophosphate (IPP) (Scheme 1).

### **Mevalonate pathway and non-mevalonate pathway**

As the prime units for terpenes, DMAPP and IPP can be obtained mainly from two different biosynthetic pathways (Scheme 1). The first one is called mevalonate pathway, also known as the HMG-CoA reductase pathway. This pathway is the predominant pathway for isoprenoid biosynthesis in most organisms, including animals, plants, fungi, and some bacteria. The pathway begins with a Claisen condensation of two molecules of acetyl-CoA to acetoacetyl-CoA by the acetyl-CoA thiolase. Acetoacetyl-CoA then undergoes an aldol addition with a third unit of acetyl-CoA catalysed by HMG-CoA synthase to form 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA). HMG-CoA is then reduced with NADPH to mevalonic acid by the enzyme HMG-CoA reductase. Two sequential adenosine triphosphate (ATP) dependent phosphorylations of mevalonic acid catalysed by the mevalonate-5-kinase and phosphomevalonate kinase give rise to mevalonate-5-pyrophosphate that can finally be converted into IPP by the ATP-dependent mevalonate-5-pyrophosphate decarboxylase. The isopentenyl pyrophosphate isomerase (IDI) is responsible for the interconversion between IPP and DMAPP through a protonation-deprotonation sequence. Stereochemically, the isomerisation of DMAPP to IPP is initialised by the protonation of the C2/C3 double bond from the *Re* side of C2, followed by a deprotonation of the (*E*)-methyl group. In the reverse reaction, the conversion of IPP into DMAPP proceeds with a protonation of the terminal double bond at its *Re* side and then deprotonation at C2 with the *pro-R* proton being lost (Scheme 2).<sup>[23]</sup>



**Scheme 2.** The stereochemical course of the interconversion of DMAPP into IPP.

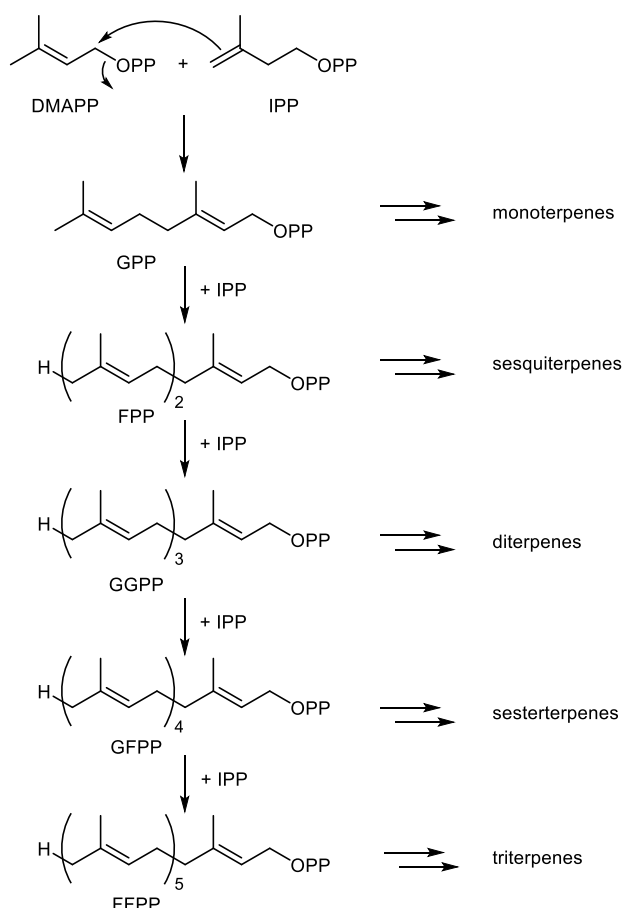
The alternative metabolic pathway towards DMAPP and IPP is called the non-mevalonate pathway, also known as the mevalonate-independent pathway or the 2-C-methyl-D-erythritol 4-phosphate/1-deoxy-D-xylulose 5-phosphate (MEP/DOXP) pathway.<sup>[24]</sup> The pathway starts with the thiamine pyrophosphate (TPP) dependent fusion of pyruvate and glyceraldehyde 3-phosphate (G3P) catalysed by the enzyme 1-deoxy-D-xylulose 5-phosphate synthase (DXS), leading under decarboxylation to the formation of 1-deoxy-D-xylulose 5-phosphate (DXP). DXP then proceeds with a rearrangement reaction and NADPH-dependent reduction, catalysed by the enzyme DXP reductoisomerase (DXR) to form 2-C-methyl-D-erythritol 4-phosphate (MEP). This compound is subsequently converted with cytidine triphosphate (CTP) into 4-diphosphocytidyl-2-C-methyl-D-erythritol (CDP-ME) by 2-C-methyl-D-erythritol 4-phosphate cytidyltransferase (CMS). CDP-ME is then phosphorylated using ATP by the enzyme MEP kinase (CMK), followed by ring closure with extrusion of cytidine monophosphate (CMP) to 2-C-methyl-D-erythritol 2,4-cyclopyrophosphate (MEcPP) in a reaction catalysed by the 2-C-methyl-D-erythritol 2,4-cyclopyrophosphate synthase (MCS). Then a reduction happens to MEcPP by the enzyme HMB-PP synthase (HDS), leading to the formation of (*E*)-4-hydroxy-3-methylbut-2-en-1-yl pyrophosphate (HMB-PP). HMB-PP can be eventually converted into IPP and DMAPP by HMB-PP reductase (HDR).

### Prenyltransferases

After the formation of DMAPP and IPP, DMAPP can be coupled with one IPP molecule to give geranyl pyrophosphate (GPP), the precursor of monoterpenes. GPP can then be elongated to farnesyl pyrophosphate (FPP) that serves as the precursor to sesquiterpenes. Further elongation can happen to FPP to form geranylgeranyl pyrophosphate (GGPP), the precursor of diterpenes, and to GGPP to produce geranylgeranyl pyrophosphate (GFPP) for the formation of sesterterpenes (Scheme 3). Triterpenes are well known to be derived from



squalene or oxidosqualene, products generated via a head-to-head condensation of two FPP molecules.<sup>[25]</sup> However, it has been found recently that farnesylfarnesyl pyrophosphate (FFPP) can also be formed and converted into triterpenes catalysed by the bifunctional fungal enzymes talaropentaene synthase (TvTS) from *Talaromyces verruculosus* and macrophomene synthase (MpMS) from *Macrophomina phaseolina*.<sup>[26]</sup>

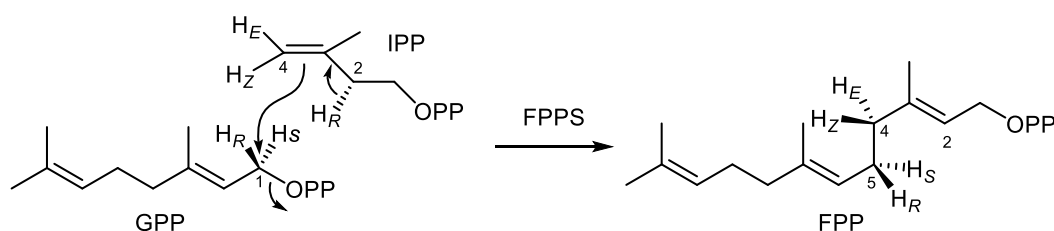


**Scheme 3.** Biosynthesis of terpene precursors by oligoprenyl pyrophosphate synthases.

The couplings of IPP with DMAPP and other building blocks are catalysed by different prenyltransferases (PTs), namely isoprenyl pyrophosphate synthases (IPPSs). According to the products of PTs in the aspect of stereochemistry, PTs can be classified as *cis*- and *trans*-PTs.<sup>[27]</sup> *Trans*-PTs usually catalyse the formation of linear *trans*-polyprenyl pyrophosphates which are the precursors of terpenes (C<sub>10</sub>–C<sub>30</sub>). In addition, *trans*-PTs such as hexaprenyl pyrophosphate synthase (HexPPS),<sup>[28][29]</sup> octaprenyl pyrophosphate synthase (OPPS),<sup>[30][31][32]</sup>

solanesyl pyrophosphate synthase (SPPS),<sup>[33][34][35]</sup> and decaprenyl pyrophosphate synthase (DPPS)<sup>[36]</sup> can produce compounds with 35, 40, 45 and 50 carbons, respectively. On the other hand, *cis*-PTs can biosynthesise products with 50 carbons or even longer chains. For instance, C<sub>55</sub>, C<sub>55</sub>–C<sub>100</sub> and C<sub>120</sub> products are made by undecaprenyl pyrophosphate synthase (UPPS),<sup>[37][38]</sup> dehydrodolichyl pyrophosphate synthase (DDPPS)<sup>[39][40]</sup> and C<sub>120</sub> polymerprenyl pyrophosphate synthase (PPPS).<sup>[41]</sup> These products with different and specific numbers of carbons play varied roles in living organisms. For example, in the biosynthesis of cell wall peptidoglycan, the product of UPPS functions as a lipid carrier.<sup>[37][38]</sup> Moreover, the natural polymer rubber biosynthetically arises through the action of a *cis*-PT. Despite the fact that the *cis*-PTs are mainly responsible for synthesising long-chain products, interestingly some exceptions do also exist. For example, a farnesyl pyrophosphate synthase (FPPS) from *Mycobacterium tuberculosis* can produce both (2*Z*)- and (2*E*)-FPP using GPP and IPP.<sup>[42]</sup>

Stereochemically, the elongation to the linear (*E*)-prenyl pyrophosphates also follows a specific rule. Taken the elongation of GPP with IPP catalysed by FPPS as an example (Scheme 4), the olefinic double bond of IPP attacks C1 in GPP from its *Si* face after the pyrophosphate is abstracted from GPP. Meanwhile a deprotonation happens to C2 in IPP with the *pro-R* proton being lost to form an (*E*)-double bond in FPP. Notably, the configuration at C1 in GPP undergoes an inversion during the elongation step to FPP.<sup>[23][23a]</sup>



**Scheme 4.** The stereochemical course of the elongation of GPP with IPP catalysed by FPPS.

### **Class I and class II terpene synthases**

It is known that two major classes of enzymes, namely class I and class II terpene synthases (TSs), are involved in the biosynthesis of terpenes.<sup>[43][44]</sup>

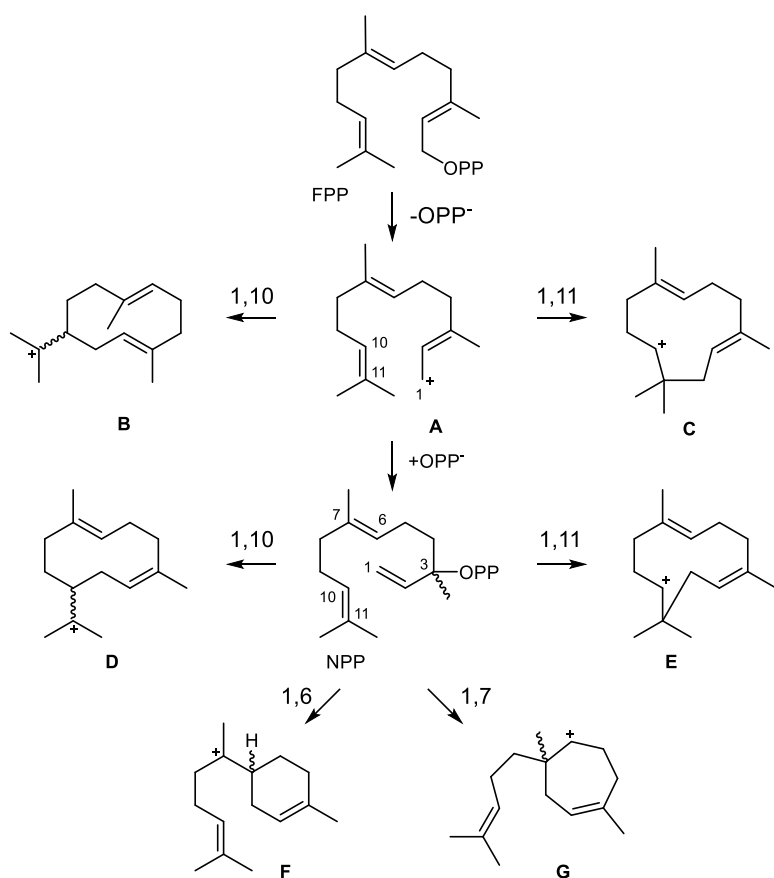
These enzymes differ substantially in their structures. Class I TSs usually exhibit an  $\alpha$ ,  $\alpha\beta$ , or  $\alpha\beta\gamma$  fold, while the class II TSs possess a  $\beta\gamma$  or  $\alpha\beta\gamma$  fold. Class I TSs are responsible for biosynthesising monoterpenes, sesquiterpenes, diterpenes, sesterterpenes and triterpenes, while class II TSs have been reported to mainly catalyse the biosynthesis of diterpenes and triterpenes. Terpene cyclisations proceed through cationic cascade reactions for both classes of terpene synthases, but the mechanism of substrate ionisation is fundamentally different. Class II terpene synthases catalyse the protonation of a terminal alkene or epoxide of the substrate, usually GGPP or squalene (epoxide), to trigger the cascaded reactions. For class I terpene synthases, the cyclisation cascade starts with the abstraction of diphosphate from an oligoprenyl pyrophosphate to form a reactive intermediate with an allyl cation. Although these two types of terpene synthases are mechanically distinct, similar terpene carbon skeletons can be formed by both classes.<sup>[45]</sup>

After the cascaded reactions are triggered by type I or type II synthases, direct sequential ring closures, hydride or proton shifts and Wagner-Meerwein rearrangements can happen. The cascades are terminated by a final deprotonation or nucleophilic attack of water to end up with a terpene hydrocarbon or alcohol. Notably, in some cases these neutral products can undergo a reprotonation to proceed with a second cascaded reaction to reach more complex polycyclic compounds. As for the type II terpene synthase catalysing the conversion of GGPP, the cyclised products still contain a pyrophosphate moiety. On the one hand, they can be hydrolysed by certain hydrolases to generate terpene alcohols. On the other hand, the type I synthases can take them over from type II terpene synthase and catalyse another reaction cascade starting with the abstraction of diphosphate to eventually form an alcohol or a hydrocarbon compound.

### 2.1. Sesquiterpene biosynthesis

As the direct precursor of sesquiterpenes, FPP can undergo the departure of diphosphate followed by a 1,10- or 1,11-cyclisation to give the (*E,E*)-germacradienyl cation (**B**) or the (*E,E*)-humulyl cation (**C**). Besides that, FPP can also be isomerised to nerolidyl pyrophosphate (NPP) through abstraction

of diphosphate and reattack at C3 (allylic transposition of diphosphate). Starting with NPP in an *anti*-S<sub>N</sub>2' reaction, the possible 1,10-, 1,11-, 1,6- and 1,7-cyclisations can happen to give the (*Z,E*)-germacradienyl cation (**D**), the (*Z,E*)-humulyl cation (**E**), the bisabolyl cation (**F**) or the cycloheptenyl cation (**G**) (Scheme 5). The thus formed reactive cationic intermediates, i.e. **B–G**, play a crucial role in biosynthetic pathways towards diverse sesquiterpenes.<sup>[43]</sup> How these downstream reactions can further proceed will be exemplified here for the sesquiterpene hydrocarbon pentalenene (**12**).



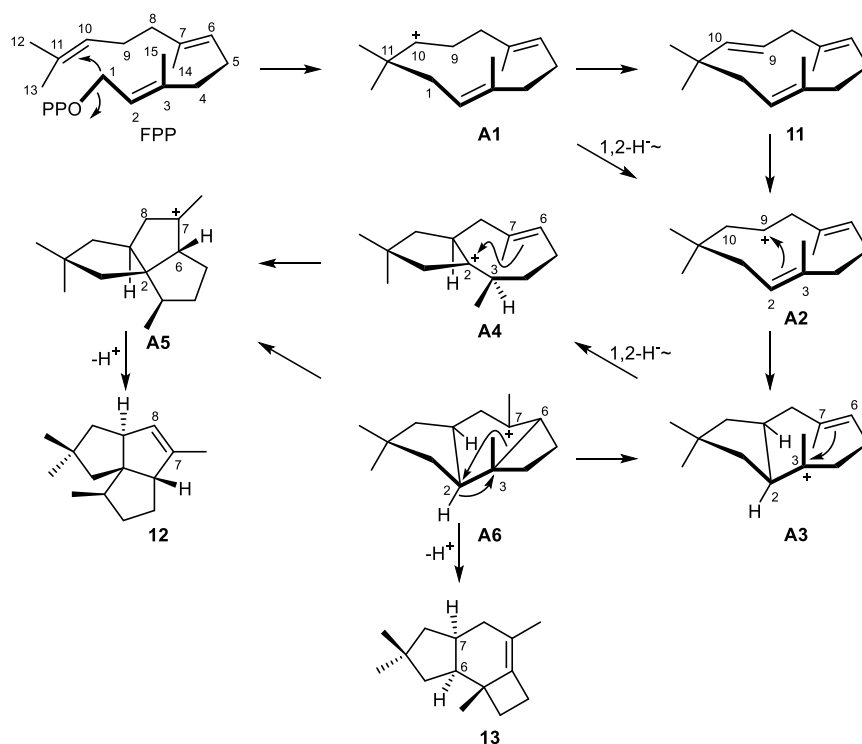
**Scheme 5.** Terpene cyclisation modes for FPP.

Pentalenene (**12**, Scheme 6) is a triquinane sesquiterpene that features a unique 5/5/5-membered ring system. This compound was first isolated from *Streptomyces griseochromogenes* in 1980.<sup>[46]</sup> Due to the structural novelty of this hydrocarbon, pentalenene has attracted a lot of chemists to work on its total synthesis.<sup>[47][48][49][50]</sup> Apart from that, its biosynthesis has also been uncovered based on a series of continuous studies. In 1983, Cane et al. worked with a cell-free extract of *Streptomyces* UC5319 that could convert FPP into

pentalenene.<sup>[51]</sup> Later on Cane and coworkers found out that the conversion was catalysed by a single enzyme and then they proposed a biosynthetic pathway towards pentalenene based on enzymatic conversions with isotopically labelled substrates (Scheme 6).<sup>[52]</sup> The cyclisation mechanism begins with an ionisation of FPP. Then a 1,11-cyclisation by electrophilic attack on the *Si* side of C11 and deprotonation at C9 generate humulene (**11**). Reprotonation of humulene at C10 followed by a 3,9-cyclisation was suggested to give the cationic intermediate **A2**. This intermediate then undergoes a 2,9-cyclisation to **A3**, a 1,2-hydride shift to **A4**, a 2,6-cyclisation to **A5** and a deprotonation at C7 to yield the final product pentalenene. The usage of (9*R*)- and (9*S*)-(9-<sup>3</sup>H,4,8-<sup>14</sup>C)FPP together with (1*S*)- and (1*R*)-(1-<sup>2</sup>H)FPP for the enzymatic reactions with pentalenene synthase further confirmed this established mechanism and indicated a formal S<sub>E</sub>' reaction taking place with net *anti* stereochemistry and a net inversion occurring at C1 during cyclisation.<sup>[53][54]</sup> However, the deprotonation and reprotonation via humulene was modified to a 1,2 hydride shift due to the experimental result that no free **11** was released during the cyclisation cascade.<sup>[52]</sup> Another reason is that the residue H309 was conjectured to serve as a base to abstract the proton at C9 to complete the deprotonation from **A1** to **11** and then to function as an acid to reprotonate C10 (Scheme 6), but the exchange of H309 against alanine, cysteine, serine or phenylalanine through site-directed mutagenesis merely influenced the catalytic activity.<sup>[55]</sup>

After the pentalenene synthase gene had been cloned giving access to the pure enzyme by heterologous expression, the crystallisation and X-ray diffraction analysis were performed successfully in 1997,<sup>[56]</sup> together with *epi*-aristolochene synthase from *Nicotiana tabacum*<sup>[56a]</sup> representing the first example showing the structure of a terpene synthase. The pentalenene synthase structure provided deep mechanistic insights into the cyclisation reactions to pentalenene and allowed for site-directed mutagenesis studies on terpene synthases targeting active site residues. It was experimentally proved that the alteration of D80, D81 and D84 in the aspartate-rich motif (DDXXD) resulted in a sharply decreased efficiency of FPP conversion. The Asp-rich motif, together with the NSE triad (NDLXSXXE) binds the trinuclear (Mg<sup>2+</sup>) cluster

that in turn binds to the substrate's pyrophosphate. Upon active site closure this leads to substrate ionisation, explaining these results of the mutational studies. The aromatic residues F76 and F77 were believed to be responsible for stabilising carbocations in the active cavity. Their exchange through site-directed mutagenesis also caused a reduced catalytic activity.<sup>[57]</sup>



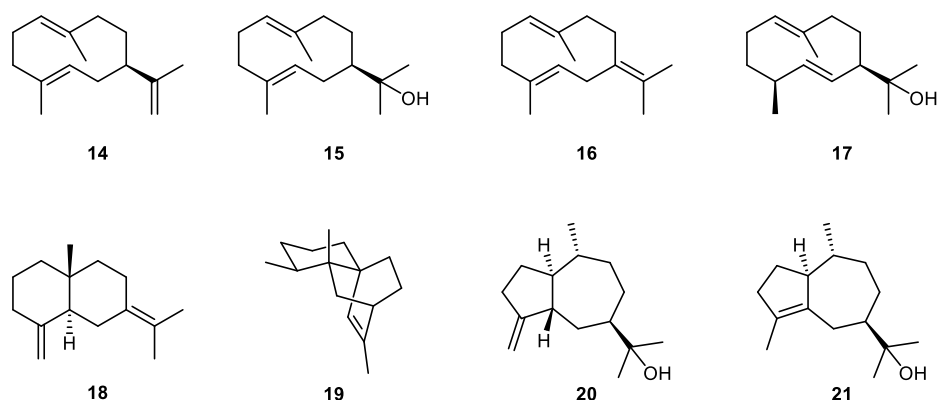
**Scheme 6.** Proposed biosynthetic pathways towards **12** and **13**.

Although the proposed cyclisation mechanism for **12** was continuously reinforced by labelling experiments, based on quantum chemical studies, Gutta et al. raised that the structures and inherent reactivities of the cationic intermediates involved in this pathway may be different from what was initially proposed. Accordingly, a revised pathway was suggested.<sup>[58]</sup> The major difference of the revised route is the step from **A3** to **A5** via **A6**. Namely, the intermediate **A3** proceeds with a ring closure instead of a 1,2-hydride transfer to give **A6** which serves as the direct precursor of protoilludene (**13**, Scheme 6), a byproduct of the enzyme variants H309A, H309S, H309C and H309F.<sup>[55]</sup> Then a dyotropic rearrangement takes place converting **A6** into **A5**. This mechanism is also completely in agreement with all of the reported experimental results on the reaction cascade to pentalenene. Finally in 2012,

the usage of (6-<sup>2</sup>H)FPP provided evidence that the latter mechanism should be favoured.<sup>[59]</sup> This was concluded from the observation that with the deuterated substrate the product distribution with the H309A enzyme variant was shifted to less **13** and more **12**, as expected for a normal kinetic isotope effect, if **A6**, but not **A3**, is the last common intermediate.

This over 30-year biosynthesis study of pentalenene, a structurally charming sesquiterpene, is a hallmark example showing how the biosynthesis of terpenes was investigated step by step by diverse developing approaches and technologies. In the following paragraphs, more biosynthesis studies of various terpenes will be introduced and the majority of these cases was investigated by me during the course of this doctoral study.

### 2.1.1. Sesquiterpenes formed by 1,10-cyclisation



**Figure 5.** Structures of compounds **14–21** that were investigated in this doctoral study.

As discussed above, 1,10-cyclisations can happen to FPP and NPP, leading to the formation of the intermediates **B** and **D** (Scheme 5). Representative compounds derived from these intermediates are germacrene A (**14**, Figure 5), hedycaryol (**15**, Figure 5), germacrene B (**16**, Figure 5) and (4*S*,7*R*)-germacra-(1(10)*E*,5*E*)-dien-11-ol (**17**, Figure 5). It is worth mentioning that germacrene A, hedycaryol and germacrene B are important neutral intermediates in the biosynthetic pathways towards many sesquiterpenes. Three review articles summarising all the sesquiterpenes derived from these compounds will be presented in Chapters 2–4, respectively. Due to the large strained ring system, some products such as germacrene A and hedycaryol exhibit three major

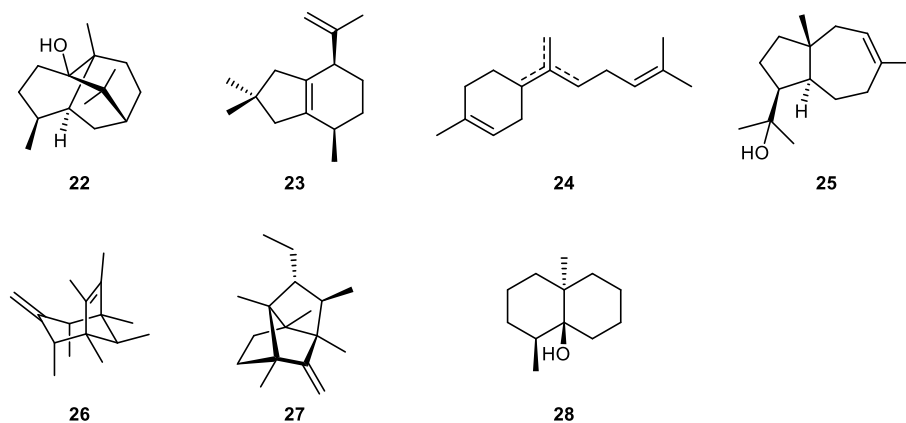
conformers and three sets of broadened peaks can be observed from their  $^{13}\text{C}$  NMR spectra. This phenomenon causes a challenging NMR assignment which was only partially solved.<sup>[60]</sup> This problem was addressed for hedycaryol using a labelling strategy, which will be introduced in Chapter 5. Interestingly, germacrene A and hedycaryol are thermally unstable. Both compounds can react in a Cope rearrangement to give elemene and elemol as rearranged products, respectively. By investigating the biosynthesis of (4*S*,7*R*)-germacra-(1(10)*E*,5*E*)-dien-11-ol by (4*S*,7*R*)-germacra-(1(10)*E*,5*E*)-dien-11-ol synthases from *Dictyostelium purpureum* and *Streptomyces coelicolor* through isotopic labelling experiments (Chapter 6), it was found that mechanistically different terpene cyclisation reactions can lead to the same compound.

Apart from these monocyclic sesquiterpenes, bicyclic or polycyclic compounds can also be generated after 1,10-cyclisation of FPP and NPP, such as selina-4(15),7(11)-diene (**18**, Figure 5) and isoishwarane (**19**, Figure 5). Selina-4(15),7(11)-diene is a eudesmane-type sesquiterpene which features a 6/6-membered ring. It was discovered that germacrene B serves as an intermediate in the biosynthetic pathway towards selina-4(15),7(11)-diene. QM/MM molecular dynamics (MD) simulations and isotopic labelling experiments were hereby performed to study the deprotonation and reprotonation steps in its biosynthesis, which will be discussed in Chapter 7. Compared to the biosynthesis of selina-4(15),7(11)-diene, the biosynthesis of isoishwarane also proceeds with a neutral intermediate, i.e. germacrene A. However, isoishwarane shows a more complex carbon skeleton because one more step of cyclisation happens in its biosynthetic pathway compared to that of selina-4(15),7(11)-diene (Chapter 8).

In addition, the biosynthesis of guaiane-type sesquiterpenes also starts with a 1,10-cyclisation of FPP. The representative compounds are guaia-4(15)-en-11-ol (**20**, Figure 5) and guaia-4-en-11-ol (**21**, Figure 5) which are involved in the defense against *P. cactorum* infection in poplar roots (Chapter 9). Based on the biosynthesis of guaia-4(15)-en-11-ol and guaia-4-en-11-ol, 1,2- or 1,3-hydride shifts in guaiane biosynthesis were investigated systematically using extensive DFT calculations (Chapter 10). Although it is known that the biosynthesis of patchoulol (**22**, Figure 6) also proceeds with guaiane intermediates, several



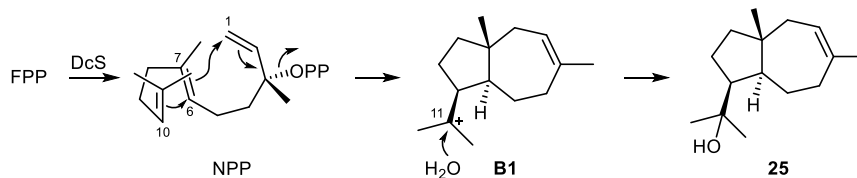
different cyclisation mechanisms were published in the literature<sup>[61][62][63]</sup> causing a confusing situation. By the use of isotopic labelling experiments, two of these mechanisms were excluded and the third one was favoured, leading to a refined mechanistic hypothesis (Chapter 11).



**Figure 6.** Structures of compounds **22–28**. Compounds **22**, **23**, **26** and **28** were investigated in this doctoral study.

### 2.1.2. Sesquiterpenes formed by 1,11-cyclisation

Some sesquiterpene synthases can also produce compounds from FPP initialised by a 1,11-cyclisation.  $\alpha$ -Humulene (**11**) is a typical example that is generated via 1,11-cyclisation. This compound is one of the few achiral terpenes and exhibits  $C_s$  symmetry. These features make it challenging to understand the stereochemical course for the 1,11-cyclisation and the terminal deprotonation in its biosynthesis. However, in conjunction with a desymmetrisation strategy, these problems were resolved based on isotopic labelling experiments (Chapter 12). Kitaviridene (**23**, Figure 6) is another example of which the biosynthesis begins with a 1,11-cyclisation of FPP. Kitaviridene is a rearranged bicyclic sesquiterpene with a 5/6-membered ring system. After the 1,11-cyclisation to form the (*E,E*)-humulyl cation, further cyclisation, hydride shifts, Wagner-Meerwein rearrangements, and a terminal deprotonation were proposed for the biosynthetic pathway towards **23** and confirmed based on labelling experiments and DFT calculations (Chapter 13).



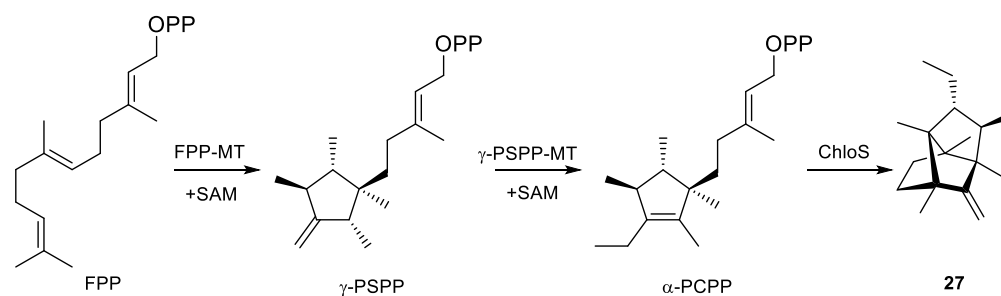
**Scheme 7.** Biosynthesis of dauc-8-en-11-ol (**25**).

### 2.1.3. Sesquiterpenes formed by 1,6- and 1,7-cyclisations

The biosynthesis of terpenes proceeding with a 1,6- or 1,7-cyclisation also contributes to the diversity of sesquiterpene skeletons. Representative compounds arising through a 1,6-cyclisation are the bisabolenes (**24**, Figure 6), for which the terpene synthases from sunflower *Helianthus annuus* have been discovered.<sup>[64]</sup> The intermediate **G** (Scheme 5) generated by a 1,7-cyclisation of NPP is a rare case since the secondary cation in **G** is usually not preferred in terpene biosynthesis. An example for a compound arising via **G** is dauc-8-en-11-ol (**25**, Figure 6).<sup>[65][66]</sup> However, **G** may be a transient species and can be avoided in a concerted 1,7- and 6,10-cyclisation to **25** (Scheme 7).

### 2.1.4. Non-canonical terpene cyclisations

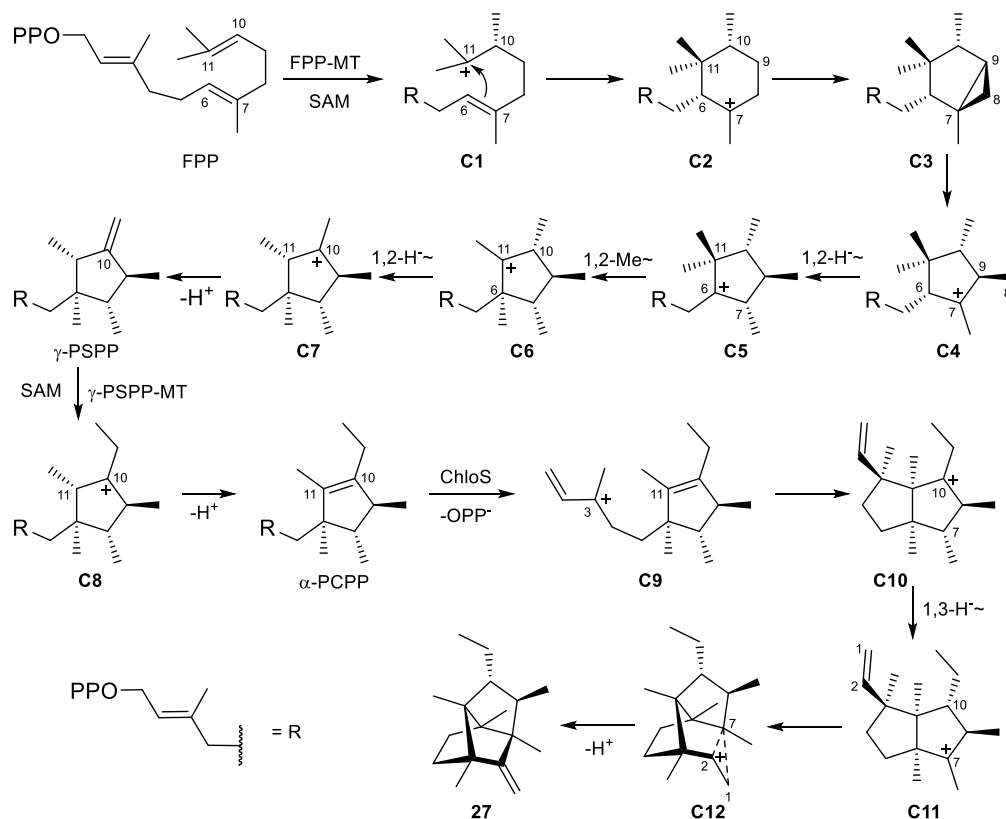
Derived from FPP, sesquiterpenes are supposed to have 15 carbons. Interestingly, there are some exceptional compounds that originate from FPP, but exhibit more or less than 15 carbons. Sodorifen (**26**, Figure 6)<sup>[67]</sup> is an outstanding natural product that is formed through a peculiar biosynthetic pathway. This compound is a  $C_s$  symmetrical methylated sesquiterpene whose biosynthesis requires the involvement of two enzymes. The first one acts as a C-methyl transferase (MT) that is responsible for the formation of presodorifen pyrophosphate. The second one is a non-canonical terpene synthase that catalyses the conversion of presodorifen pyrophosphate into sodorifen.<sup>[68]</sup> As one of the main achievements of the study on **26** for this thesis, the biosynthetic pathway from presodorifen pyrophosphate to sodorifen was deeply investigated. The terpene cyclisation step was proposed to undergo hydrogen migrations, a highly unusual fragmentation, [4+3] cycloaddition and deprotonation. In addition, the biosynthesis of presodorifen pyrophosphate was also studied and the absolute configuration of this compound was revised (Chapter 14).



**Scheme 8.** Biosynthesis of chlororaphen (**27**).

Chlororaphen (**27**, Figure 6) is a brexane-type bishomosesquiterpene with 17 carbons. The biosynthesis of this compound needs the involvement of three enzymes, i.e. FPP methyl transferase (FPP-MT),  $\gamma$ -presodorifen pyrophosphate methyl transferase ( $\gamma$ -PSPP-MT) and chlororaphen synthase (ChloS). Firstly, FPP-MT is responsible for the introduction of one methyl group from S-adenosyl methionine (SAM) to FPP that promotes a cyclisation reaction to form  $\gamma$ -PSPP. An second methyl group is subsequently incorporated into  $\gamma$ -PSPP from SAM by  $\gamma$ -PSPP-MT to give the double methylated  $\alpha$ -prechlororaphen pyrophosphate ( $\alpha$ -PCPP) that finally undergoes the cascaded reactions catalysed by ChloS to reach chlororaphen.<sup>[69]</sup>

The cyclisation mechanism of chlororaphen starts with the introduction of a methyl group at C10 to give **C1**, followed by a 7,11-cyclisation to obtain cation **C2**. A ring contraction via the formation of a three-membered ring (**C3**) happens to **C2** to result in the intermediate **C4**. Then a 1,2-hydride shift followed by a methyl group migration and another 1,2-hydride shift of **C4** reaches **C7** that undergoes deprotonation to form the neutral intermediate  $\gamma$ -PSPP. All these reaction steps above are catalysed by FPP-MT. Then  $\gamma$ -PSPP-MT introduces another methyl group and catalyses the deprotonation to generate  $\alpha$ -PCPP. This ethylated product is finally taken over by ChloS to complete the biosynthesis of chlororaphen. The cascaded reactions catalysed by ChloS include the pyrophosphate abstraction to form **C9**, a 3,11-cyclisation to **C10**, a 1,3-hydride transfer to **C11**, a 2,7-cyclisation to **C12** and a terminal deprotonation (Scheme 9).<sup>[69]</sup>

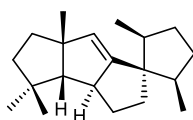


**Scheme 9.** Cyclisation mechanism from FPP to **27**.

Geosmin (**28**, Figure 6) is another unusual secondary metabolite originating from FPP. This terpenoid compound features a 6,6-membered bicyclic skeleton with only 12 carbons. It is known that a terpene synthase with two functional domains is responsible for the biosynthesis of geosmin, and three neutral intermediates are generated in the enzymatic reaction cascade to geosmin.<sup>[70]</sup> Experiments with the terpene scavenger  $\beta$ -cyclodextrin and multiple isotopically  $^{13}\text{C}$  labelled and/or deuterated substrates deciphered how these domains interact with each other and refined the cyclisation mechanism of geosmin in all stereochemical details (Chapter 15).

## 2.2. Diterpene biosynthesis

Diterpenes are terpenes in principle with 20 carbons due to the fact that GGPP serves as their common precursor. As discussed above, both class I and class II synthases are known to produce diterpenes.

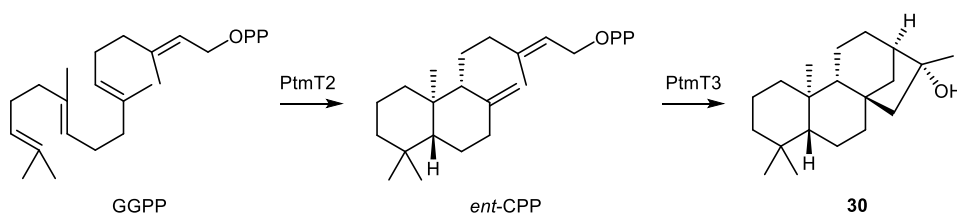


29

**Figure 7.** Structure of compound **29** that was investigated in this doctoral study.

The terpenetriene synthase (Cyc2) from *Streptomyces griseolosporeus* MF730-N reported in 2001, represents the first bacterial type I diterpene synthase.<sup>[71][72]</sup> From then on, more and more bacterial type I diterpene synthases have been revealed.<sup>[73]</sup> Spiroviolene synthase (SvS) from *Streptomyces violens* is one of the noticeable examples.<sup>[74]</sup> Spiroviolene (**29**, Figure 7) is a 5/5/5/5-membered spirotetracyclic diterpene of which the structure was revised based on a total synthesis by Snyder and coworkers.<sup>[75]</sup> Its biosynthesis was thus reinvestigated and modified based on the experimental results with isotopically labelled substrates (Chapter 16).

In addition, some diterpenes biosynthetically require the collaboration of both classes of terpene synthases. For example, the biosynthesis of (16*R*)-*ent*-kauran-16-ol (**30**, Scheme 10) starts with the enzyme-catalysed reaction from GGPP by PtmT2 to afford *ent*-copalyl pyrophosphate (*ent*-CPP). Then PtmT3 acts on *ent*-CPP to fulfil the further cyclisation to afford **30**. Notably, the enzyme PtmT2 is known as a type II terpene synthase while PtmT3 works as a type I terpene synthase.<sup>[76]</sup>



**Scheme 10.** Biosynthesis of compound **30**.

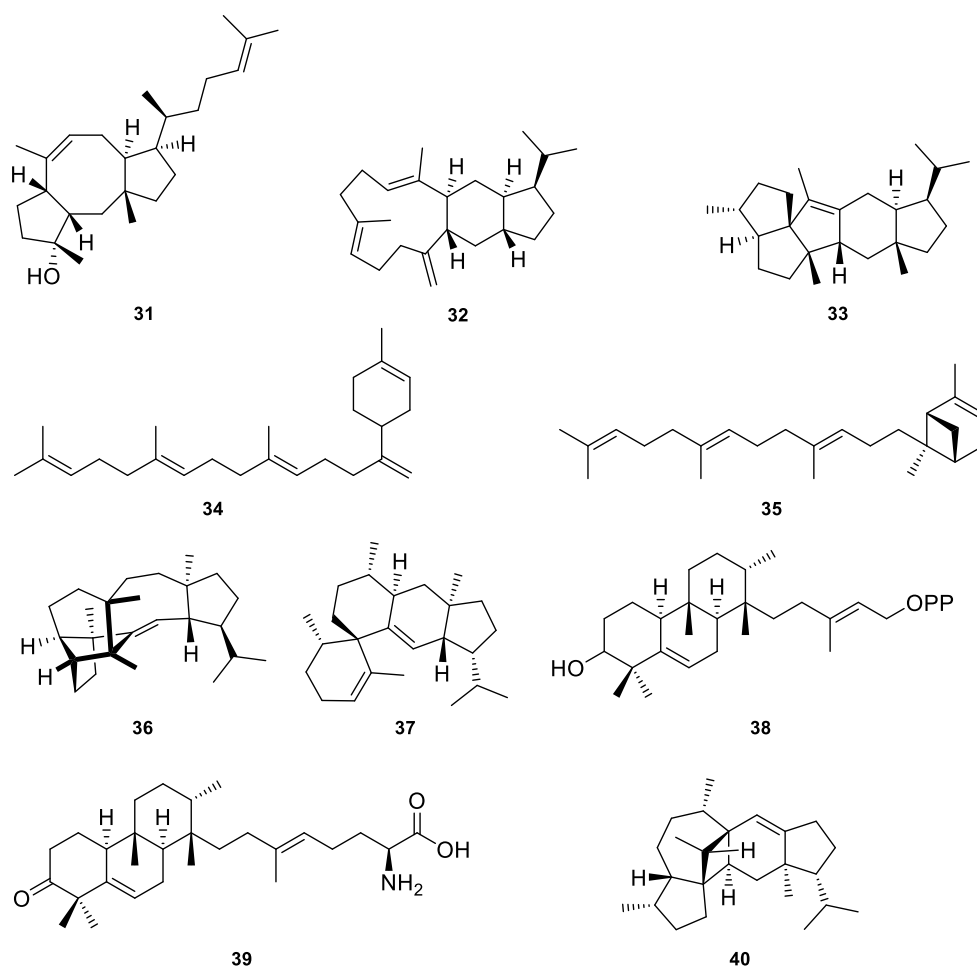
### 2.3. Sesterterpene biosynthesis

Compared to sesquiterpenes and diterpenes, sesterterpenes usually exhibit much more sophisticated structures since its precursor, GFPP has more reactive sites than GGPP and FPP. Due to the complexity of elucidating the structures and the biosynthetic pathways of sesterterpenes, only a limited

number of sesterterpenes together with their synthases has been fully investigated from plants, fungi and bacteria.<sup>[77-82]</sup>

It has been found that usually sesterterpene synthases from fungi are bifunctional enzymes with two domains that serve as a prenyltransferase and a terpene synthase, respectively. Therefore, these enzymes cannot only convert GFPP, but also shorter precursors such as DAMPP, GPP or FPP in conjunction with IPP into sesterterpenes. The biocatalyst for ophiobolin F (**31**, Figure 8), the *A. clavatus* ophiobolin F synthase (AcOS), represents the first sesterterpene synthase that functions as a prenyltransferase and a terpene synthase at the same time.<sup>[77]</sup>

Different from fungi, plants and bacteria often are the sources of monofunctional enzymes that only accept GFPP as a substrate. *AtTPS18* and *AtTPS19* from *Arabidopsis thaliana* were reported as the first plant sesterterpene synthases in 2017. They are capable to catalyse the conversion of GFPP into thalianatriene (**32**, Figure 8) and retigeranin B (**33**, Figure 8), respectively.<sup>[78]</sup> As for the sesterterpene synthases from the bacterial kingdom, the first discovered enzyme was a UbiA related terpene synthase and was designated as StsC. This enzyme was discovered in 2018 from *Streptomyces somaliensis* and produces somaliensenes A (**34**, Figure 8) and B (**35**, Figure 8).<sup>[79]</sup> Later on, two more classical type I sesterterpene synthases, namely SmTS1 generating sestermobaraene A (**36**, Figure 8) as a major product from *Streptomyces mobaraensis* (6),<sup>[80]</sup> and SvSS for sesterviolene (**37**, Figure 8) from *Streptomyces violens*<sup>[81]</sup> have been reported. Interestingly, compound **38** (Figure 8), produced by a type II sesterterpene synthase from *Streptomyces albus* was proposed as one of intermediates for the biosynthesis of atolypene A (**39**, Figure 8).<sup>[82]</sup>



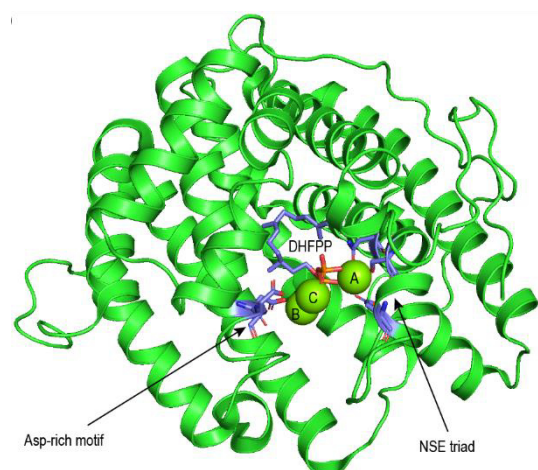
**Figure 8.** Structures of compounds **31–40**.

As part of this doctoral thesis, a new bacterial sesterterpene synthase from *Kitatospora viridis* was characterised. This enzyme catalyses the conversion of GFPP into sesterviridene A (**40**, Figure 8), a pentacyclic sesterterpene, as the major product. The cyclisation mechanism of the sesterterpene synthase for this compound was uncovered using isotopic labelling experiments in conjunction with DFT calculations (Chapter 17).

#### 2.4. Structural biology of terpene synthases

Nowadays, more and more crystal structures of terpene synthases have been uncovered.<sup>[83]</sup> The first crystal structure of the prenyl transferase was reported in 1994 for the farnesyl pyrophosphate synthase (FPPS) from avian liver.<sup>[83a]</sup> The successful elucidation of the three-dimensional structure revealed a novel enzyme fold composed only of  $\alpha$ -helices, with ten core helices around the central cavity containing two highly conserved aspartate-rich sequences on

opposite sides of the cavity. The structure of selina-4(15),7(11)-diene synthase from *Streptomyces pristinaespiralis* represents one of the most insightful structures of sesquiterpene synthases (Figure 9).<sup>[84]</sup> By its open and fully closed structures, an induced-fit mechanism was identified. Strikingly, a novel motif that contains the pyrophosphate sensor Arg178, the linker Asp181, and the effector Gly182-O involved in substrate ionisation was disclosed.



**Figure 9.** Structure of selina-4(15),7(11)-diene synthase (SdS, PDB: 4OKZ) with highlighted Asp-rich motif and NSE triad. Green spheres represent Mg<sup>2+</sup>. DHFPP = 2,3-dihydro-FPP.

Generally, the structures of terpene synthases provide good opportunities to gain a deeper understanding how the complex reactions catalysed by them can proceed in the active cavity of a single enzyme. Terpene synthases share several highly conserved motifs and residues for different functions in their active sites. As mentioned above, the aspartate-rich motif DDX(X)D is regarded as an Mg<sup>2+</sup> cofactor-binding site. The NSE triad ND(L,I,V)XSXX(R,K)E is also responsible for binding of the Mg<sup>2+</sup> cofactor. In addition, the structure of the selina-4(15),7(11)-diene synthase reveals a single Arg that is located upstream of the NSE triad and serves as a pyrophosphate sensor. This sensor is involved in the substrate recognition by forming hydrogen bridges with the substrate, and its importance for enzyme activity was confirmed by site-directed mutagenesis.<sup>[84]</sup> Moreover, the RY pair can also recognise the substrate through hydrogen bridges. There are also other important residues such as the Arg-Glu salt bridge between helices F and G,<sup>[85]</sup>



and the conserved Trp upstream of the RY pair, which exhibit a structural function.

It is remarkable that starting with an acyclic and achiral precursor, a single enzyme can catalyse multiple cascaded reactions to give polycyclic products that are enantiomerically pure. Nowadays, more and more structures of terpene synthases have been crystallised and elucidated,<sup>[83]</sup> providing increasing possibilities to explore the chemical space that can be reached with terpene synthases. The access of an enzyme structure also contributes to the discovery of the active cavity and the conserved residues within. Structure based site-directed mutagenesis studies on the active sites of terpene synthases allow to distinguish the different functions of those conserved residues. In addition, reshaping of the active site of a terpene synthase by site-directed mutagenesis can lead to a functional alteration, such as increase of the production and expansion of the product spectrum. The associated research will be presented in Chapters 18–19 in which structure-based site-directed mutagenesis experiments of sesquiterpene and diterpene synthases, respectively, are discussed, together with a mechanistic investigation through isotopic labelling experiments. Based on previous related work, a review article summarising the mechanistic investigations on microbial class I terpene synthase through site-directed mutagenesis is provided at the end of this thesis (Chapter 20).



## Chapter 2

### Germacrene A – A Central Intermediate in Sesquiterpene Biosynthesis

Houchao Xu<sup>[a]</sup> and Jeroen S. Dickschat<sup>[a],\*</sup>

[a] Kekulé-Institute for Organic Chemistry and Biochemistry, University of Bonn, Gerhard-Domagk-Straße 1, 53121 Bonn, Germany. Email: [dickschat@uni-bonn.de](mailto:dickschat@uni-bonn.de).

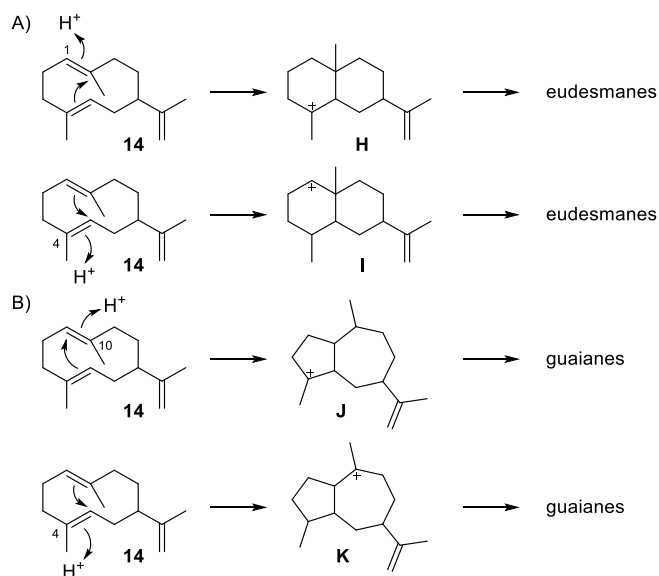
Reprinted from *Chem. Eur. J.* **2020**, 26, 17318 with kind permission from John Wiley & Sons

The publication “Germacrene A – A Central Intermediate in Sesquiterpene Biosynthesis” can be found in Appendix A.



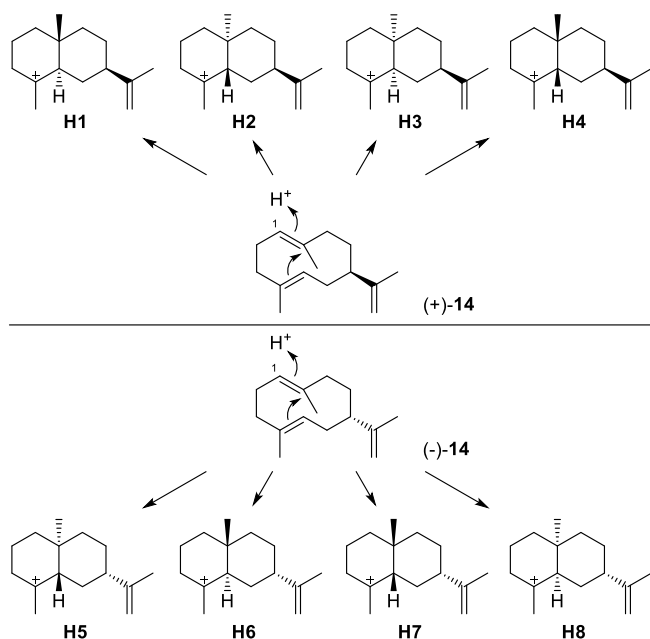
As mentioned already in the introductory Chapter 1, terpenes represent with more than 100,000 known compounds the by far largest class of natural products. Sesquiterpenes are one subclass of the terpenes and are produced by all kingdoms of life. Based on the available genome information, during the past two decades, more and more sesquiterpene synthases have been discovered.<sup>[43]</sup> As explained in detail in Chapter 1, the precursor of sesquiterpenes is farnesyl diphosphate (FPP), which can undergo a direct 1,10- or a 1,11-cyclisation to reach the (*E,E*)-germacradienyl cation (**B**) and the (*E,E*)-humulyl cation (**C**). More cationic intermediates can be obtained when FPP is converted into nerolidyl pyrophosphate (NPP) by isomerisation through *syn*-allylic transposition of diphosphate, allowing 1,10-cyclisations, 1,11-ring closures, 1,6- or 1,7-cyclisations. These reactions can happen to NPP after diphosphate abstraction to give the (*Z,E*)-germacradienyl cation (**D**), the (*Z,E*)-humulyl cation (**E**), the bisabolyl cation (**F**) and the cycloheptenyl cation (**G**) (Scheme 5). These reactive cationic intermediates can proceed with direct sequential ring closures, hydride or proton shifts and Wagner-Meerwein rearrangements, and can finally be quenched by deprotonation or nucleophilic attack of water to give neutral products, i.e. terpene hydrocarbons or alcohols, respectively. Notably, the primarily arising neutral products can be reprotonated to undergo further cascaded reactions to give more complex polycyclic compounds.

Germacrene A (**14**), as was introduced in Figure 5, is a sesquiterpene biosynthetically derived from FPP undergoing a 1,10-cyclisation to give **B** followed by deprotonation at C12 or C13. This natural product was first obtained in 1970 from *Eunicea mammosa*<sup>[86]</sup> and its absolute configuration was determined by chemical correlation with (+)- $\beta$ -elemene through a thermal Cope rearrangement.<sup>[87][87a]</sup> Structurally, germacrene A possesses three major conformers due to its flexible, but highly strained 10-membered ring which resulted in a challenging NMR assignment for this compound<sup>[88]</sup>. Despite of this fact, the NMR data sets (25 °C) for all three conformers were recently fully resolved using a <sup>13</sup>C-labelling strategy.<sup>[89]</sup>

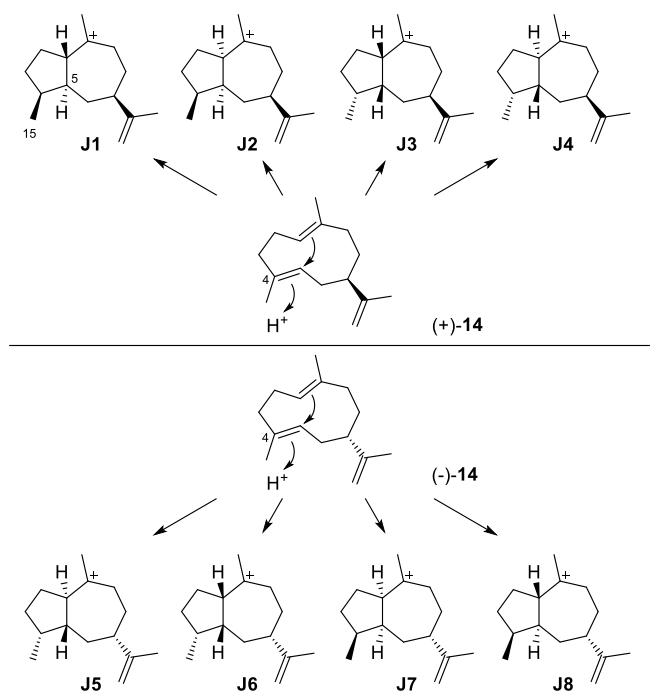


**Scheme 11.** Secondary terpene cyclisations of **14**.

Germacrene A (**14**) is an essential neutral intermediate towards other sesquiterpenes since further cyclisations can happen to this compound initiated by reprotonation. As shown in Scheme 11, reprotonation followed by cyclisation of **1** can generate eudesmanes (**H** and **I**) with a 6-6 bicyclic skeleton and guaianes (**J** and **K**) with a 5-7 bicyclic skeleton. Theoretically, eight stereochemically distinct cationic intermediates are possible for each of these bicyclic intermediates. For example, Scheme 12 shows the eight stereoisomeric intermediates produced by cyclisations induced by reprotonation of **14** at C1, while Scheme 13 indicates the eight stereoisomeric intermediates generated from **14** through cyclisations induced by reprotonation at C4. Then these cationic intermediates can go through cascaded reactions demonstrated above to reach many corresponding compounds.



**Scheme 12.** Cyclisation reactions induced by reprotonation of **14** at C1 leading to intermediates **H1–H8**.



**Scheme 13.** Cyclisation reactions induced by reprotonation of **14** at C4 leading to intermediates **J1–J8**.

In this review, we summarised all the compounds derived from these cationic intermediates together with their structures and known natural sources, explained the way of determination of the absolute configurations, and provided

information about the optical rotations and NMR data, if available. Some related synthetic compounds were also included in this review. In addition, some confusions in the literature regarding structure elucidation and absolute configuration were also described in detail. Totally 369 references were cited in this article.

As for my work on this review, I wrote the sections on the compounds with guaiane skeleton (e.g.  $\delta$ -guaiene, pogostol, etc.) including rearranged guaianes (e. g. patchoulenes, seychellenes, etc.).



## Chapter 3

### Hedycaryol – Central Intermediates in Sesquiterpene Biosynthesis, Part II

Houchao Xu<sup>[a]</sup> and Jeroen S. Dickschat<sup>[a],\*</sup>

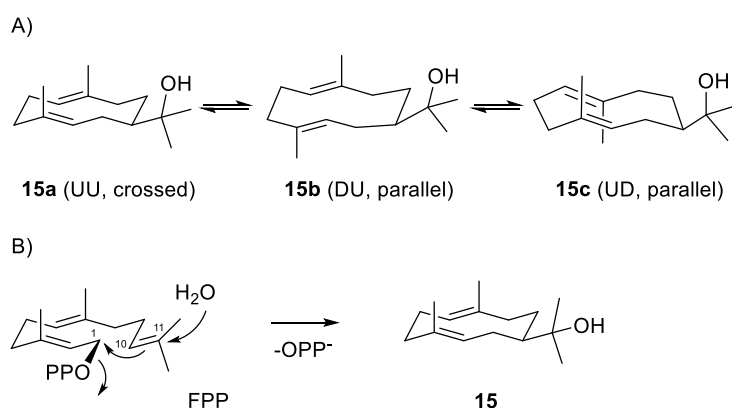
[a] Kekulé-Institute for Organic Chemistry and Biochemistry, University of Bonn, Gerhard-Domagk-Straße 1, 53121 Bonn, Germany. Email: [dickschat@uni-bonn.de](mailto:dickschat@uni-bonn.de).

Reprinted from *Chem. Eur. J.* **2022**, *28*, e202200405 with kind permission from John Wiley & Sons

The publication “Hedycaryol – Central Intermediates in Sesquiterpene Biosynthesis, Part II” can be found in Appendix B.



The first review of this series of these articles summarised all the compounds derived from germacrene A initialised by reprotonation at different positions. Apart from germacrene A, hedycaryol (**15**, Scheme 14), a sesquiterpene alcohol, is also regarded as a vital neutral intermediate towards many sesquiterpenes, as is extensively discussed in this second review article. Both germacrene A and hedycaryol have the same 10-membered ring in their structures which accounts for their similar structural, chemical and biosynthetic properties, and their equally important roles in sesquiterpene biosynthesis.

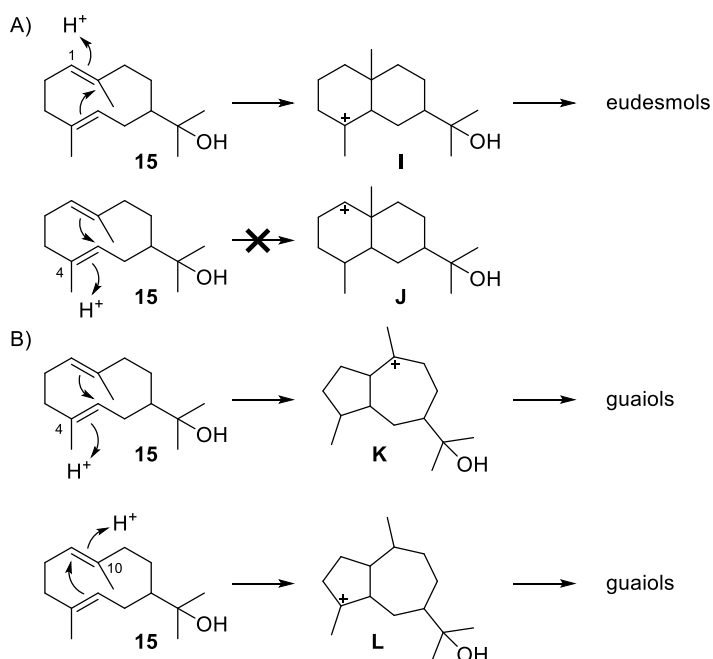


**Scheme 14.** A) Conformers of **15**. U = Me group at 10-membered ring up, D = Me group down. „Crossed“ and „parallel“ refers to relative orientations of double bonds. B) Biosynthetic pathway to **15**.

Hedycaryol has first been isolated from *Hedycarya angustifolia*<sup>[90]</sup>, followed by isolations from *Phebalium ozothamnoides*,<sup>[91]</sup> *Rubus rosifolius*,<sup>[92]</sup> *Thujopsis dolabrata*,<sup>[93]</sup> *Thymus praecox*,<sup>[94]</sup> *Cryptomeria japonica* and *C. fortunei*,<sup>[95]</sup> and *Chamaecyparis obtusa*.<sup>[96]</sup> Hedycaryol is a thermally unstable compound and it can undergo a Cope rearrangement to form elemol, which can be easily observed in gas chromatography (GC). Due to the highly strained ring system in **15**, structurally this compound possesses three major conformers (Scheme 14A), as germacrene A has, which is the reason for serious obstacles in its NMR assignment. Recently, this problem has been completely solved using an isotopic <sup>13</sup>C labelling strategy.<sup>[97]</sup>

Biosynthetically, after the departure of the diphosphate from farnesyl pyrophosphate (FPP), a 1,10-cyclisation followed by the attack of water affords **15** (Scheme 14B). Up to present, many hedycaryol synthases have been

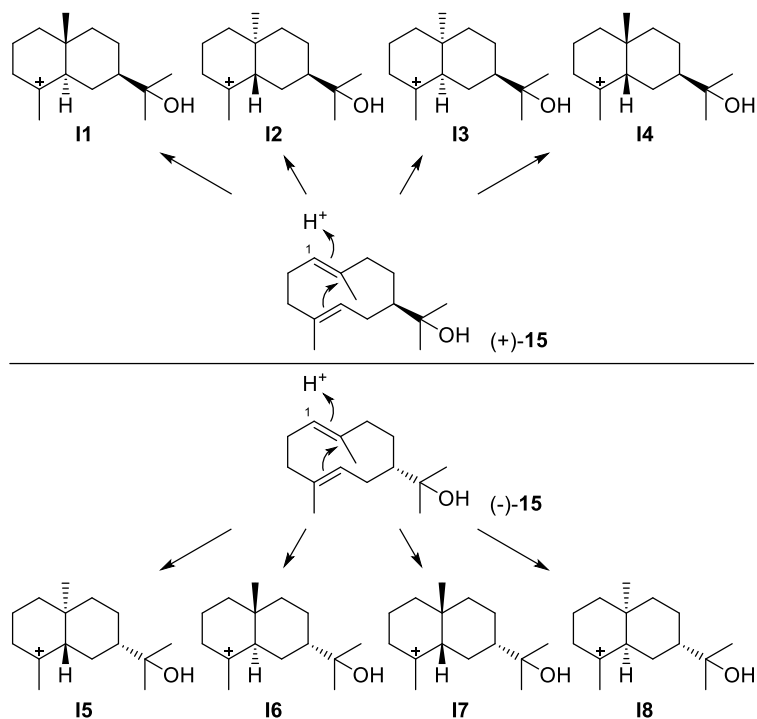
discovered from *Populus trichocarpa* (PtTPS7),<sup>[98]</sup> *Camellia brevistyla* (CbTPS1),<sup>[99]</sup> *Liquidambar formosana* (LfTPS01),<sup>[100]</sup> and *Kitasatospora setae*,<sup>[97]</sup> Interestingly, the diterpene synthase VenA from *Streptomyces venezuelae* <sup>[101]</sup> can also convert FPP into this sesquiterpene alcohol.



**Scheme 15.** Possible terpene cyclisation modes for **15**.

As discussed above, hedycaryol is an essential neutral precursor in sesquiterpene biosynthesis because it can be reprotonated to proceed with downstream enzymatic cascade reactions towards other sesquiterpenes. Since there are two sets of double bonds, i.e. C1=C10 and C4=C5 in this compound, theoretically four positions, C1, C4, C5 and C10, can undergo reprotonation. However, considering the intermediates with secondary cations are not preferred, the only plausible cationic intermediates after cyclisation initialised by reprotonation are eudesmols (**I**) and guaiols (**K** and **L**) (Scheme 15). Intermediate **I** features a 6/6-membered bicyclic ring system with the cation at C4, while **K** and **L** are characterised by a 5/7-membered bicyclic ring system with the cationic charges at C10 and C4, respectively. To be more specific, eight stereoisomers are derived from each intermediate if the stereochemistry is taken into consideration. For instance, the planar structure **I** corresponds to eight possible stereoisomers (**I1** – **I8**) because of three stereogenic centres

in this intermediate (Scheme 16). Guaiols (**K** and **L**) share the situation as in the eudesmols (**I**). Taken together, 24 cationic intermediates from **I**, **K** and **L** are presented in this review.



**Scheme 16.** Cyclisation reactions of **15** induced by reprotonation at C1 towards intermediates **I1 – I8**.

This review summarises all the products derived from these cationic intermediates. Usually these compounds are obtained via hydride shifts, methyl group migrations, attack of water, intramolecular attack of the hydroxyl group in the formation of ethers, Wagner–Meerwein rearrangements (WMR) and deprotonations from these 24 cationic intermediates. In this review, these compounds are introduced by exhibiting their structures, the first isolation, approaches to identify their planar structures and to determine the absolute configurations, their optical rotations, natural sources, essential bioactivities and NMR data, if available. Besides these general points, confusing aspects regarding the structure elucidation and absolute configuration are also elaborated in this review. Some synthetic compounds that can be produced biosynthetically from **15** are also introduced. A total number of 243 references was included in this article. The biosynthetic derivatives from hedycaryol not

only include sesquiterpene alcohols, but also ethers can be formed through intramolecular attack of the alcohol function in hedycaryol to a cationic centre. For the preparation of this review article, I was responsible for summarising the compounds derived from the cationic intermediates **K** and **L** towards guaiols.

## Chapter 4

### Germacrene B – A Central Intermediate in Sesquiterpene Biosynthesis

Houchao Xu<sup>[a]</sup> and Jeroen S. Dickschat<sup>[a],\*</sup>

[a] Kekulé-Institute for Organic Chemistry and Biochemistry, University of Bonn, Gerhard-Domagk-Straße 1, 53121 Bonn, Germany. Email: [dickschat@uni-bonn.de](mailto:dickschat@uni-bonn.de).

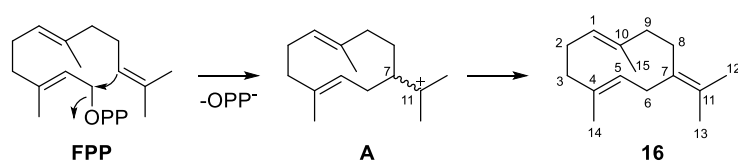
Reprinted from *Beilstein J. Org. Chem.* **2023**, *19*, 186 with kind permission from Beilstein Institute for the Advancement of Chemical Sciences

The publication “Germacrene B – A Central Intermediate in Sesquiterpene Biosynthesis” can be found in Appendix C.





This review focuses on the compounds that biosynthetically originate from germacrene B (**16**, Scheme 17) that is treated as another central intermediate in the biosynthesis of sesquiterpenes, besides the related compounds germacrene A and hedycaryol which were discussed in the previous two chapters. The germacrenes A and B, and hedycaryol are highly related in the aspects of their structures and biosynthetic formations. Structurally, these compounds own the same carbon backbone with a 10-membered ring system. The only difference is that germacrene A and B have a set of double bonds residing at C11/C13 and C7/C11 respectively, while hedycaryol possess a hydroxyl group at C11, which can be explained from the perspective of their biosynthetic pathways. Due to the formation of the double bonds between C7 and C11, the stereogenic centre at C7 which exists in germacrene A and hedycaryol is not present in the case of germacrene B, making **16** an achiral and  $C_s$  symmetric compound. Biosynthetically, these sesquiterpenes are all derived from the same cationic intermediate, namely the (*E,E*)-germacradienyl cation (**A**). Cation **A** is generated by 1,10-cyclisation induced by the departure of diphosphate from farnesyl pyrophosphate (FPP). Germacrenes A and B can be reached, if **A** undergoes alternative deprotonations at C12 or C10, respectively, while hedycaryol can be obtained, if **A** proceeds with the attack of water.

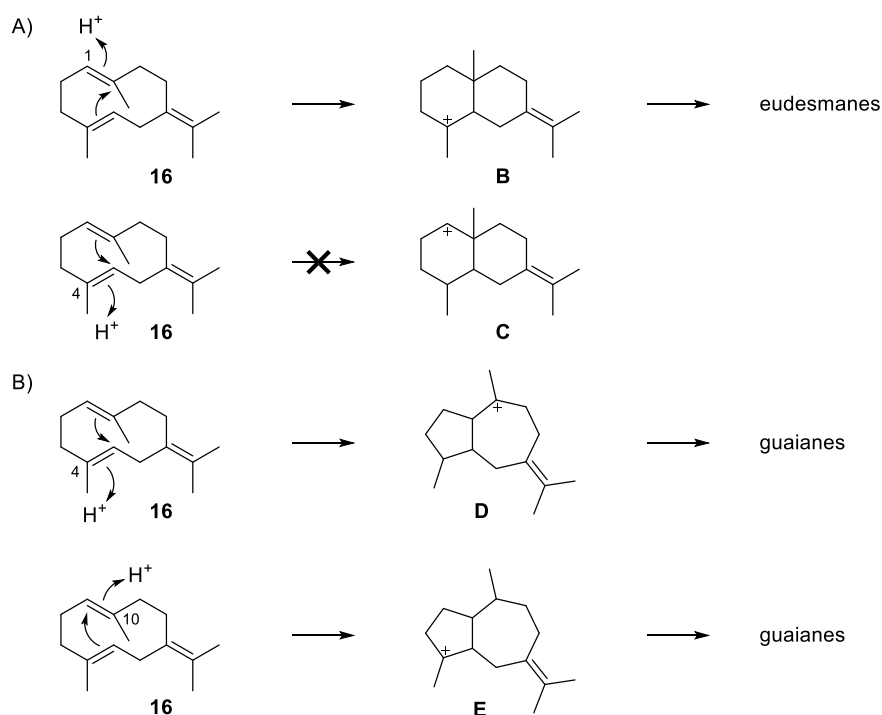


**Scheme 17.** Biosynthesis of germacrene B (**16**).

Germacrene B was first discovered naturally from *Humulus lupulus* <sup>[102]</sup> and *Citrus junos*,<sup>[103]</sup> and before it was already chemically synthesised.<sup>[104]</sup> From then on, the compound has continuously been isolated from *Stenocalyx michelii*,<sup>[105]</sup> *Citrus aurantifolia*, <sup>[106]</sup> and *Solidago canadensis*. <sup>[107]</sup> In addition, germacrene B synthases from *Solanum habrochaites*<sup>[108]</sup> and *Cannabis sativa*<sup>[109]</sup> have also been reported. Interestingly, germacrene B is a side product of many other terpene synthases, e.g. the germacrene C synthase from

*Lycopersicon esculentum*,<sup>[110]</sup> the (+)-germacrene D synthase from *Zingiber officinalis* <sup>[111]</sup>, the avermitilol synthase from *Streptomyces avermitilis* <sup>[112]</sup>, and VoTPS1 from *Valeriana officinalis* <sup>[113]</sup>, which indicates that **16** is involved in many terpene biosynthetic pathways.

Due to the large ring system in this compound, **16** was calculated to have four conformers, suggesting similar structural features as reported for germacrene A and hedycaryol. However, unlike the fact that for germacrene A and hedycaryol three major conformers can be observed, the conformers of germacrene B may undergo a rapid interconversion as manifested by the fact that only one peak for each carbon can be observed in the <sup>13</sup>C-NMR spectrum.<sup>[114]</sup> Although germacrene B seems to be less strained in comparison to germacrene A and hedycaryol, it can be likewise converted into  $\gamma$ -elemene under high temperature conditions through a Cope rearrangement.



**Scheme 18.** Possible cyclisation reactions upon reprotonation of **16**. A) Cyclisations to eudesmane sesquiterpenes, B) cyclisations to guaiane sesquiterpenes.

As one of the important central intermediates in sesquiterpene biosynthesis, germacrene B can undergo reprotonation leading to the formation of other compounds. As Scheme 18 shows, cationic intermediates **B**, **D** and **E** can be

generated upon reprotonation and cyclisation of **16**. Compared to the cationic intermediates produced from germacrene A and hedycaryol, **B**, **D** and **E** only have two stereogenic centres, meaning that each of them corresponds to four instead of eight stereoisomers. Therefore, totally 12 cationic stereoisomers are generated from **B**, **D** and **E**.

This review presents all the compounds arising from these 12 intermediates. Specifically, the elucidation of their structures including relative and absolute configurations, the first isolation together with other natural sources, their optical rotations, biological activities and NMR data, if available, are summarised. Some synthetic compounds are also introduced in this review article if the formations of their structures are biosynthetically logical. Notably, several compounds hypothetically derived from germacrene B can be found in the CAS abstracting system with confusing or obviously erroneous information. For example, more than one compound may be assigned to one CAS number (CAS number 473-04-1 corresponds to (+)-,(-)- and (±)-juniper camphor). In addition, the absolute configurations of some compounds are mentioned, but no relevant publications report these data at all. In this review article, all these kinds of problems are described in detail.

Taken together, 64 compounds are introduced and 131 references are cited in this review. The compounds derived from germacrene B are much fewer compared to those from germacrene A and hedycaryol. The main reason could be that **B**, **D** and **E** only have two stereogenic centres which results in a limited number of possible stereoisomeric cations.

My contribution to this review was to introduce the guaiane sesquiterpenes obtained from intermediates **D** and **E**.



## Chapter 5

### Isotopic Labeling Experiments Solve the Hedycaryol Problem

Houchao Xu<sup>[a]</sup>, Nathalie D. Lackus<sup>[b]</sup>, Tobias G. Köllner<sup>[b]</sup>, and Jeroen S. Dickschat<sup>[a],\*</sup>

[a] Kekulé-Institute for Organic Chemistry and Biochemistry, University of Bonn, Gerhard-Domagk-Straße 1, 53121 Bonn, Germany. Email: [dickschat@uni-bonn.de](mailto:dickschat@uni-bonn.de).

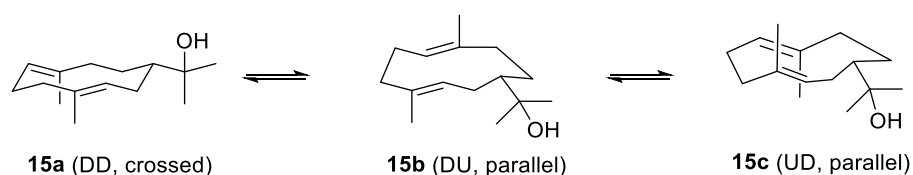
[b] Max Planck Institute for Chemical Ecology, 07745 Jena, Germany.

Reprinted from *Org. Lett.* **2022**, *24*, 587 with kind permission from American Chemical Society

The publication “Isotopic Labeling Experiments Solve the Hedycaryol Problem” can be found in Appendix D.

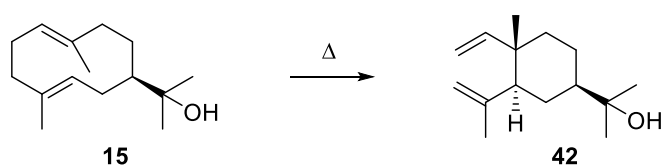


One of our previous review articles presented the accumulated knowledge of sesquiterpenes biosynthetically derived from hedycaryol (**15**). As discussed before, hedycaryol is a monocyclic sesquiterpene alcohol which is characterised by a strained 10-membered ring. This feature allows the existence of three major conformers in this compound (Scheme 19).<sup>[115]</sup> Since these conformers cannot be separated, there are always three peaks, one sharp and two broad ones, for one carbon position observed from the <sup>13</sup>C-NMR, and seriously overlapped and broadened proton signals which are unable to be distinguished from the <sup>1</sup>H-NMR. In addition, **15** is a thermally unstable compound which can undergo a Cope rearrangement to give elemol (**41**) under high temperature (Scheme 20).<sup>[116]</sup> Furthermore, hedycaryol is also sensitive to acidic conditions. Even the mildly acidic silica gel used for column chromatography can induce the chemical conversion of **15** into  $\alpha$ -,  $\beta$ - and  $\gamma$ -eudesmol (**42–44**, Scheme 21) and therefore the purification of this product by column chromatography with silica gel cannot be realised.<sup>[116][117]</sup> Taken together, these properties make it challenging to fully assign the NMR data for **15**, although this compound has been known for a long time.



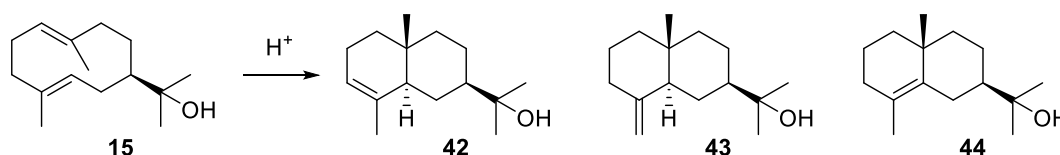
**Scheme 19.** Three major conformers of **15**.

As another key neutral intermediate in the biosynthesis of sesquiterpenes, germacrene A shows the same problems as **15** has. However, the full NMR assignments for its three conformers have been accomplished using an isotopic <sup>13</sup>C and deuterium labelling strategy.<sup>[89]</sup> In this research article, the same approach is applied to solve the problem for hedycaryol.



**Scheme 20.** Cope rearrangement of **15** to form **42**.

In this study, the hedycaryol synthase (HcS) from *Kitasatospora setae* [118] and the enzyme variant PtTPS5\_C403A from *Populus trichocarpa* were investigated. Originally, PtTPS5 can convert farnesyl pyrophosphate (FPP) into (1*S*,5*S*,7*R*,10*R*)-guaia-4(15)-en-11-ol (**20**) and (1*S*,7*R*,10*R*)-guaia-4-en-11-ol (**21**) (Figure 5).<sup>[119]</sup> The site-directed mutagenesis study of PtTPS5 shows that the enzyme variant PtTPS5\_C403A has lost its capability of producing **20** and **21**, but turns out to be an efficient synthase to generate hedycaryol from FPP. It is also known that terpene synthases from plants and bacteria are usually capable to produce opposite enantiomers.<sup>[120][121]</sup> Since HcS and PtTPS5\_C403A originate from a bacterium and a plant, respectively, their enzymatic reaction products were analysed by gas chromatography (GC) using a chiral stationary phase. Hedycaryol from the plant and the bacterial enzyme indeed exhibited different retention times, revealing their enantiomeric nature.



**Scheme 21.** Acid catalyzed reaction of **15** to eudesmols **42–44**.

As discussed above, each carbon of **15** shows one sharp and two broad peaks in the <sup>13</sup>C-NMR spectrum. In order to avoid the signals of all 15 carbons crowded in one spectrum at the same time, a trace of **15** with <sup>13</sup>C labelling for one single carbon was produced by the enzymatic conversion of <sup>13</sup>C labelled substrate with HcS and PtTPS5\_C403A. Due to the <sup>13</sup>C labelling, the signals of the corresponding carbon were enhanced which means the labelled carbon can still be detected by <sup>13</sup>C NMR despite the trace amount of measured sample. The analysis of only trace amounts of labelled material also has one advantage: No signals can be observed for the unlabelled carbons which avoids disturbance from them in the analysis. Eventually, 15 trace samples of **15** with <sup>13</sup>C labelling for 15 different carbons could be prepared and measured by NMR spectroscopy. By using this approach, all these three chemical shifts for each carbon can be assigned. The assignments of all the chemical shifts for the protons can also be addressed through HSQC measurements since the cross-



peaks for the attached hydrogen atoms at the labelled carbon atoms are also significantly enhanced. After all the NMR data were obtained, an enzymatic reaction with ( $^{13}\text{C}_{15}$ )FPP was conducted to solve the problem which 15 signals belong to one conformer by  $^{13}\text{C}$ - $^{13}\text{C}$  COSY measurement.

In order to follow the Cope rearrangement, all the 15 samples of singly  $^{13}\text{C}$  labelled hedycaryol were heated to 130 °C and the corresponding  $^{13}\text{C}$  labelled isotopomers of elemol were obtained. The NMR data of elemol were therefore assigned unambiguously. Notably, the isotopically labelled probes (*E*)- and (*Z*)-( $4\text{-}^2\text{H}, 4\text{-}^{13}\text{C}$ )IPP as well as (*R*)- and (*S*)-( $1\text{-}^2\text{H}, 1\text{-}^{13}\text{C}$ )IPP allowed to elucidate the stereochemical course of the Cope rearrangement from conformer **15a** to elemol.

In this study, I first screened the activities of all the enzyme variants and then performed all the enzymatic reactions with labelled substrates. I also conducted all the chemical reactions to study the Cope rearrangement of hedycaryol.



## Chapter 6

### Mechanistic Divergence between (4*S*,7*R*)-Germacra-(1(10)*E*,5*E*)-dien-11-ol Synthases from *Dictyostelium purpureum* and *Streptomyces coelicolor*

Houchao Xu,<sup>[a]</sup> Jan Rinkel,<sup>[a]</sup> Xinlu Chen,<sup>[b]</sup> Tobias G. Köllner,<sup>[c]</sup> Feng Chen<sup>[b]</sup> and Jeroen S. Dickschat<sup>[a].\*</sup>

[a] Kekulé-Institute of Organic Chemistry and Biochemistry, University of Bonn, Gerhard-Domagk-Strasse 1, 53121 Bonn, Germany. Email: dickschat@uni-bonn.de.

[b] Department of Plant Sciences, University of Tennessee, 2431 Joe Johnson Drive, Knoxville, TN 37996-4561, USA.

[c] Max Planck Institute for Chemical Ecology, Hans-Knöll-Straße 8, 07745 Jena, Germany.

Reprinted from *Org. Biomol. Chem.* **2021**, *19*, 370 with kind permission from Royal Society of Chemistry

The publication “Mechanistic Divergence between (4*S*,7*R*)-Germacra-(1(10)*E*,5*E*)-dien-11-ol Synthases from *Dictyostelium purpureum* and *Streptomyces coelicolor*” can be found in Appendix E.

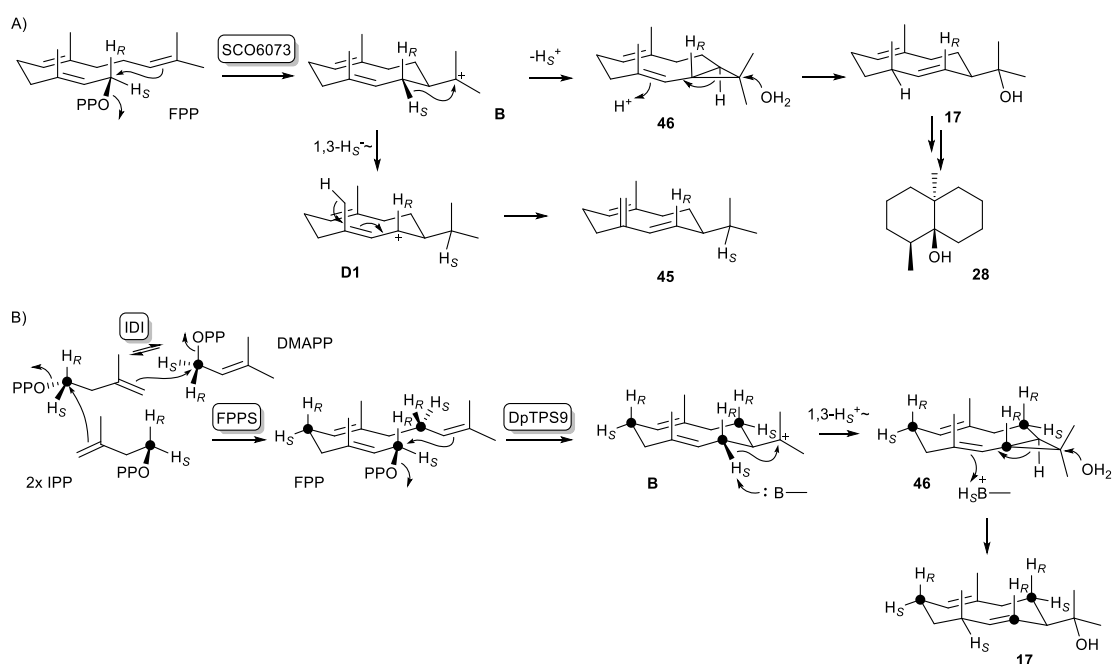


(4*S*,7*R*)-Germacra-(1(10)*E*,5*E*)-dien-11-ol (**17**, Scheme 22) is a sesquiterpene which has been reported from several myxobacteria,<sup>[122][123]</sup> streptomycetes,<sup>[124][125]</sup> cyanobacteria,<sup>[126]</sup> and sponge-associated fungi.<sup>[127]</sup> It was also reported to be an intermediate in the biosynthesis of geosmin (**28**, Scheme 22).<sup>[128]</sup> Recently, it was found that a sesquiterpene synthase, DpTPS9 from the dictyostelid amoeba *Dictyostelium purpureum* also produces this compound together with a minor sesquiterpene hydrocarbon (–)-germacrene (**45**, Scheme 22). The planar structure of **17** was unambiguously identified using extensive one- and two-dimensional NMR spectroscopy. In addition, four stereoselectively deuterated probes (*E*)- or (*Z*)-(4-<sup>13</sup>C,4-<sup>2</sup>H)IPP<sup>[129]</sup> and (*R*)- or (*S*)-(1-<sup>13</sup>C,1-<sup>2</sup>H)IPP<sup>[89]</sup> were used to determine the absolute configuration of **17**. These probes can be further elongated to the corresponding isotopomers of farnesyl pyrophosphate (FPP) with enantioselective deuterations. Then they were converted with DpTPS9 to give stereoselectively labelled products which supported the absolute configuration of (4*S*,7*R*)-**17**. Meanwhile **45** was also identified by comparison to both enantiomers present in the essential oil of *Solidago canadensis* through gas chromatography using a homochiral stationary phase.<sup>[130]</sup>

Compound **17** is known to be produced by the N-terminal domain of the bacterial geosmin synthase SCO6073 from *Streptomyces coelicolor*.<sup>[70]</sup> However, it was found that the amino acid sequences of the amoebal and the bacterial enzyme are different, which indicates that the same enzyme functions have evolved independently. Interestingly, after deeper investigation by isotopic labelling experiments it was discovered that the biosynthetic pathway to **17** by the geosmin synthase SCO6073 from *S. coelicolor* also differs to that by DpTPS9 from *D. purpureum*. As Scheme 22 shows, although both pathways undergo the reprotonation at C4 of the neutral intermediate **46**, geosmin synthase SCO6073 utilizes a proton from the buffer to complete this step while DpTPS9 catches the proton which has been released during the deprotonation step to form **46**.

In addition, the site-directed mutagenesis for residues C60, M67, F78 and N249 was also studied to further understand the biosynthesis of **17** and the results confirmed the importance of the Phe residue located three positions upstream

of the Asp-rich motif and of the Asn residue located nine positions downstream of the NSE triad.



**Scheme 22.** Divergent cyclisation mechanisms from FPP to **17**. A) Mechanism for the geosmin synthase SCO6073 from *S. coelicolor*, B) mechanism for DpTPS9 from *D. purpureum*. Bold dots indicate  $^{13}\text{C}$ -labelled carbons,  $H_R$  or  $H_S$  were substituted with deuterium.

In this work, I was responsible for the isolation, purification and structure elucidation of **17** together with conducting all the labelling experiments to determine the absolute configuration and to identify the biosynthetic pathway of **17**. For the site-directed mutagenesis study, I conducted all the enzymatic reactions with the mutants for the statistics of their activities.

## Chapter 7

### Catalytic Role of Carbonyl Oxygens and Water in Selinadiene Synthase

Yong-Heng Wang,<sup>[a],[b],#</sup> Houchao Xu,<sup>[c],#</sup> Jian Zou,<sup>[a],#</sup> Xian-Bo Chen,<sup>[a]</sup> Yu-Qing Zhuang,<sup>[a]</sup> Wei-Liang Liu,<sup>[a]</sup> Guo-Dong Chen,<sup>[a]</sup> Ersin Celik,<sup>[c]</sup> Dan Hu,<sup>[a]</sup> Hao Gao,<sup>[a]</sup> Ruibo Wu,<sup>[b],\*</sup> Ping-Hua Sun<sup>[a],\*</sup> and Jeroen S. Dickschat<sup>[c],\*</sup>

[a] College of Pharmacy, Jinan University, Guangzhou 510632, P. R. China. E-mail: pinghuasunny@163.com.

[b] School of Pharmaceutical Sciences, Sun Yat-sen University, Guangzhou 510006, P. R. China. E-mail: wurb3@mail.sysu.edu.cn.

[c] Kekulé-Institute for Organic Chemistry and Biochemistry, University of Bonn, Gerhard-Domagk-Straße 1, 53121 Bonn, Germany. E-mail: dickschat@uni-bonn.de.

# These authors contributed equally to this work.

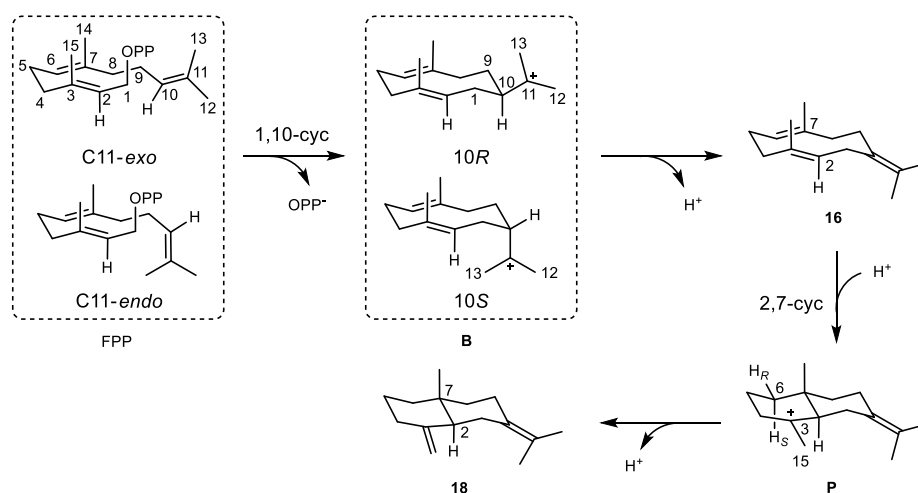
Reprinted from *Nat. Catal.* **2022**, *5*, 128 with kind permission from Springer Nature

The publication “Catalytic Role of Carbonyl Oxygens and Water in Selinadiene Synthase” can be found in Appendix F.



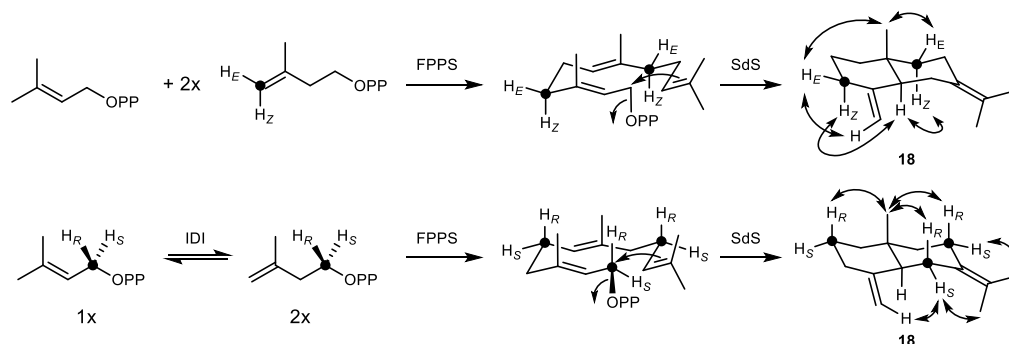


Terpene synthases (TSs) catalyse a complex cyclisation from simple, acyclic and achiral oligoprenyl diphosphates to provide diverse carbon skeletons with multiple rings and stereogenic centres. The formation of diverse terpene structures is attributed to the pluripotent reactivity of carbocationic intermediates, which can proceed with ring closures by intramolecular attack of an olefin to a cationic centre, hydride or proton shifts and Wagner–Meerwein rearrangements, and then undergo a terminal deprotonation or nucleophilic attack of water to give neutral products. As discussed in the previous chapters for germacrene A and hedycaryol, these neutral products can be reprotonated to participate in further cascaded reactions to give more complex polycyclic compounds.<sup>[131]</sup> Deprotonation-reprotonation steps happen frequently in the terpene biosynthesis, which are proposed as an alternative to direct intramolecular proton transfers.<sup>[132]</sup> Selina-4(15),7(11)-diene (**18**, Scheme 23) is another interesting case whose biosynthetic pathway contains a deprotonation-reprotonation step through the intermediate germacrene B. Since the crystal structure of selina-4(15),7(11)-diene synthase (SdS) was previously reported,<sup>[84]</sup> the deprotonation-reprotonation mechanism in the biosynthesis of **18** by SdS from *Streptomyces pristinaespiralis* was hereby investigated using QM/MM molecular dynamics (MD) simulations and isotopic labelling experiments



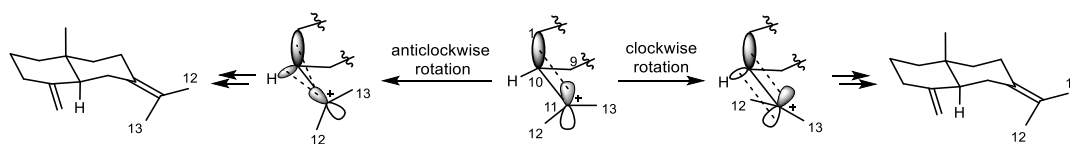
**Scheme 23.** Possible mechanisms for SdS catalysis. Carbon numbers follow FPP numbering, carbon numbers in blue follow **16** and **18** numbering.

The precursor of **18** is farnesyl pyrophosphate (FPP) biosynthetically (Scheme 23). After abstraction of diphosphate, a 1,10-cyclisation gives **B**. Then deprotonation from C10 happens to the cationic intermediate to yield germacrene B. This neutral compound can proceed with 2,7-cyclisation initiated by reprotonation at C6 to obtain **P**. After the final deprotonation from C15, **18** is formed.<sup>[133]</sup>



**Scheme 24.** Stereoselective deuteration experiments ( $H_R$  or  $H_S = {}^2\text{H}$ , black dots represent  ${}^{13}\text{C}$  labels) for assignment of the absolute configuration of **18**. Double headed arrows indicate key NOESY correlations.

Meanwhile, the absolute configuration of **18** was also identified using stereoselectively deuterated substrates, including (*R*)- or (*S*)-(1- ${}^{13}\text{C}$ ,1- ${}^2\text{H}$ )IPP and (*E*)- or (*Z*)-(4- ${}^{13}\text{C}$ ,4- ${}^2\text{H}$ )IPP (Scheme 24). The HSQC analyses of the obtained products by the enzymatic conversions of these probes indicated that the absolute configuration was (2*S*,7*R*)-**18**.

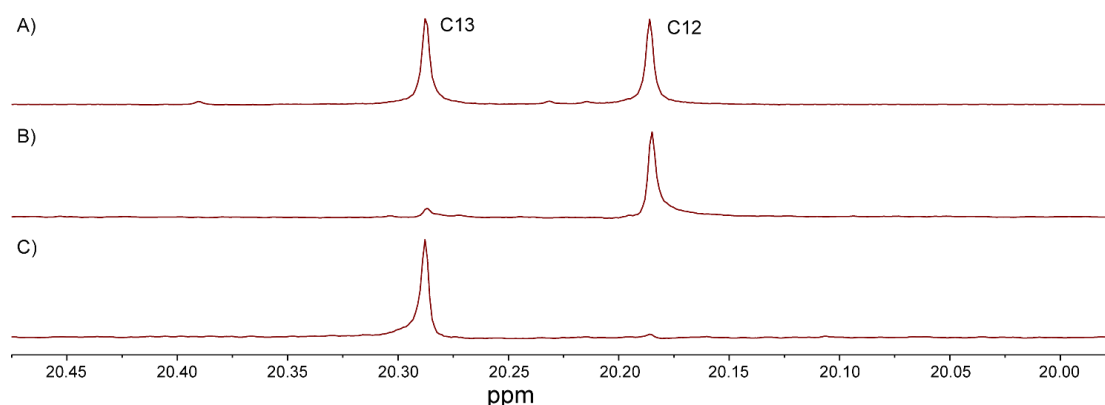


**Scheme 25.** Molecular orbital interactions during the rotation of C12–C11(+)-C13 around the C10–C11 bond in (10*R*)-**A**<sup>+</sup>. Dashed lines represent hyperconjugations, numbers at carbons follow FPP numbering.

It was reported that both conformations, i.e. C11-*exo* and C11-*endo* of FPP can be observed in the homotetrameric crystal structure of SdS.<sup>[84]</sup> Therefore, (10*R*)- and (10*S*)-**B** can be formed from C11-*exo*-FPP and C11-*endo*-FPP

respectively, either of which can lead to **18** theoretically (Scheme 23). Then these two conformations were investigated by docking them to the active site of SdS crystal structure. The result revealed that the C11-*exo*-FPP is recognized by the carbonyl oxygen of Gly182 that later has a critical function in stabilization of (10*R*)-**B** through electrostatic interaction.

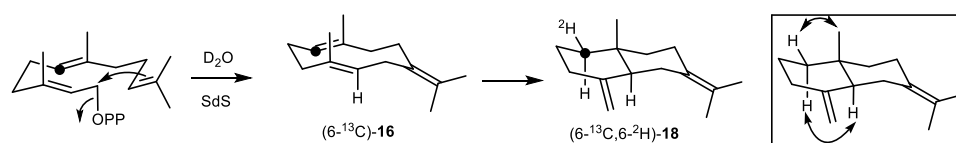
After the cyclisation of C11-*exo*-FPP to (10*R*)-**B**, either clockwise or anticlockwise rotation of C12–C11(+)-C13 around the C10–C11 bond is required to complete the subsequent deprotonation (Scheme 25). The PES scans showed that clockwise rotation is preferred and a spontaneous deprotonation of C10 by Gly182 happened upon dihedral rotation indicating that the carbonyl oxygen of Gly182 serves as the base in the deprotonation of (10*R*)-**B**. The theoretically favoured clockwise rotation of C12–C11(+)-C13 was also confirmed by isotopic labelling experiments using (12-<sup>13</sup>C)FPP and (13-<sup>13</sup>C)FPP (Figure 10).



**Figure 10.** The stereochemical fate of the geminal methyl groups C12 and C13 of FPP in the cyclisation to **18**. Enzymatic conversions with SdS of A) FPP, B) (12-<sup>13</sup>C)FPP, and C) (13-<sup>13</sup>C)FPP.

After the deprotonation of **B** to **16**, the cyclisation to **18** requires reprotonation of **16** at C6. Notably, the protonated carbonyl of Gly182 was considered to be further involved in this reprotonation step in conjunction with a molecule of water. This hypothesis is supported by QM/MM MD simulation. The reprotonation step was also confirmed experimentally using (6-<sup>13</sup>C)FPP in D<sub>2</sub>O buffer showing deuterium incorporation into the 6-*pro-R* position of **18** (Scheme 26). Taken together, Gly182 plays an essential role in the deprotonation–reprotonation

sequence because it first serves as the base and then as the acid mediated through one water molecule.



**Scheme 26.** Conversion of (6-<sup>13</sup>C)FPP in deuterium oxide buffer with SdS into (6-<sup>13</sup>C)-**16** and (6-<sup>13</sup>C,6-<sup>2</sup>H)-**18**. Double headed arrows at the structure in the box show key NOESY correlations observed for unlabelled **18** for the assignment of the 1-*pro-R* and 1-*pro-S* hydrogens.

My contributions to this work include isolation, structure elucidation and full NMR assignment of **18** and carrying out the isotopic labelling experiments to determine the absolute configuration of **18** and to follow the deprotonation-reprotonation sequence in **18** biosynthesis.

## Chapter 8

### Isoishwarane synthase from *Streptomyces lincolnensis*

Houchao Xu<sup>[a]</sup>, Jan Rinkel<sup>[a]</sup> and Jeroen S. Dickschat<sup>[a].\*</sup>

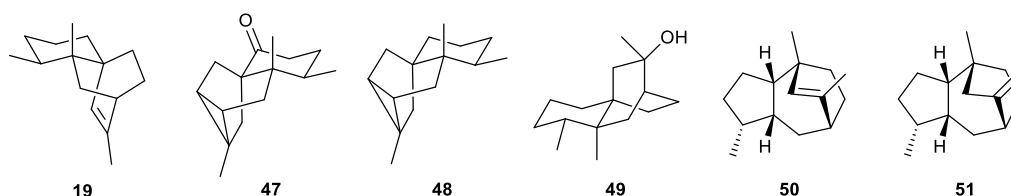
[a] Kekulé-Institute for Organic Chemistry and Biochemistry, University of Bonn, Gerhard-Domagk-Straße 1, 53121 Bonn, Germany. Email: [dickschat@uni-bonn.de](mailto:dickschat@uni-bonn.de).

Reprinted from *Org. Chem. Front.* **2021**, *8*, 1177 with kind permission from Royal Society of Chemistry

The publication “Isoishwarane synthase from *Streptomyces lincolnensis*” can be found in Appendix G.



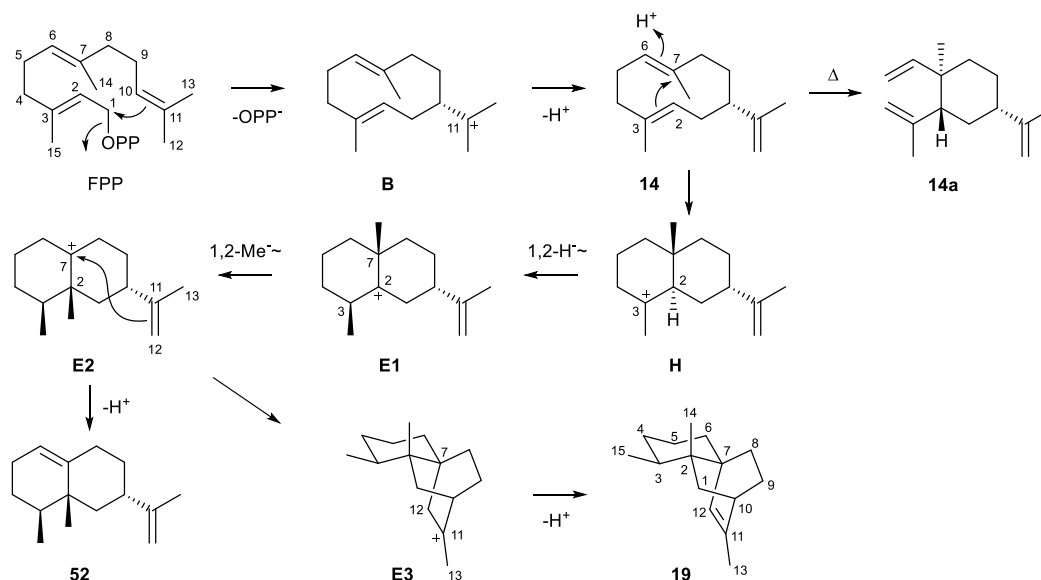
Isoishwarane (**19**, Figure 11) is a tricyclic sesquiterpene hydrocarbon that was first chemically synthesised as a racemate.<sup>[134][135]</sup> Later on, (–)-**19** was also prepared via a chemical approach from ishwarone (**47**) (Figure 11), a natural product isolated from *Aristolochia indica*.<sup>[136][137]</sup> Structurally, compared to the tricyclic patchoulol (**22**),<sup>[138]</sup> isoishwarane only has three methyl groups which is unusual since FPP, the precursor of sesquiterpenes, owns four methyl groups. It is apparent that one of the methyl groups is involved in the cyclisation and becomes incorporated into a ring. Similar natural sesquiterpenes include (+)-ishwarone (**47**) from *Aristolochia indica*,<sup>[136][137]</sup> (–)-ishwarane (**48**) from *Aristolochia indica*,<sup>[139]</sup> 8,12-seco-ishwaran-12-ol (**49**) from *Litsea amara*,<sup>[140]</sup> and rotundene (**50**) and isorotundene (**51**) from *Cyperus rotundus* (Figure 11).<sup>[141][142]</sup>



**Figure 11.** The structures of isoishwarane (**19**), (+)-ishwarone (**47**), (–)-ishwarane (**48**), 8,12-seco-ishwaran-12-ol (**49**), rotundene (**50**) and isorotundene (**51**).

*Streptomyces lincolnensis* is well known to produce lincomycin, a clinically important lincosamide antibiotic against Gram-positive bacteria.<sup>[143]</sup> The head space extraction of *Streptomyces lincolnensis* NRRL 2936 in this study using a closed loop stripping apparatus (CLSA) revealed that this strain also produces volatile terpenes including geosmin as a main component, and isoishwarane. In order to fully investigate the biosynthesis of isoishwarane generated by *S. lincolnensis* in vitro, especially to understand how one of the methyl groups is incorporated into a ring via cyclisation, its gene was cloned and expressed in *E. coli* BL 21. GC/MS analysis of the products obtained from FPP with isoishwarane synthase (IWS) showed the formation of isoishwarane together with  $\beta$ -elemene (**14a**), germacrene A (**14**) and valencene (**52**) as minor products.  $\beta$ -Elemene is regarded to be a product of germacrene A through a thermal Cope rearrangement during GC/MS analysis. Interestingly, the optical rotation of isoishwarane produced by IWS shows a positive sign, which is the opposite to

that of (–)-**19** prepared from plant-derived ishwarone. This finding is in line with the frequent observation that terpene synthases from plants and bacteria can produce the opposite enantiomers of terpenes.



**Scheme 27.** The biosynthesis of isoishwarane (**19**), germacrene A (**14**) and valencene (**52**). And the thermol Cope rearrangement to β-elemene (**14a**).

The investigations of the biosynthesis of **19** started with the enzymatic conversion of the 15 isotopomers of singly <sup>13</sup>C labelled FPP to follow all the carbons. This approach is beneficial to decipher the carbon rearrangements in the biosynthesis. It was discovered that the methyl group (C14) originally connected to C7 in FPP resides at C2 in the final product, which supports the proposed 1,2-methyl group shift from **E1** to **E2** (Scheme 27). C7 instead becomes connected to C12, a methyl group in FPP, which explains the reduced number of methyl groups in **19**. This finding is consistent with the step from **E2** to **E3** proposed in the biosynthetic pathway (Scheme 27). Interestingly, the <sup>13</sup>C NMR of **19** for the experiment with (12-<sup>13</sup>C)FPP shows a major peak for C12 together with a minor one for C13. The experiment with (13-<sup>13</sup>C)FPP showed a similar result with a high peak for C13 and a low peak for C12 observed from the <sup>13</sup>C NMR.

In the biosynthetic pathway towards **19**, germacrene A was suggested to be an intermediate. Its reprotonation was confirmed by the enzymatic conversion of



FPP in D<sub>2</sub>O buffer analysed by GC/MS showing a deuterium incorporated into **19**. A 1,2-hydride shift from **H** to **E1** was also proposed and then proved by the experiment with (2-<sup>2</sup>H,3-<sup>13</sup>C)FPP. The phenomenon described above regarding the experiments with (12-<sup>13</sup>C)FPP and (13-<sup>13</sup>C)FPP are also explainable in this pathway. The relevant step is the deprotonation of **B** to **14** because the deprotonation of **B** happens mainly from C12 with a slight involvement of C13. In addition, the formation of **52** is proposed to be obtained from **E2** by deprotonation.

The absolute configuration of **19** was determined using the isotopically labelled substrates (*E*)- or (*Z*)-(4-<sup>13</sup>C,4-<sup>13</sup>C)IPP and (*R*)- or (*S*)-(1-<sup>13</sup>C,1-<sup>2</sup>H)IPP, while the absolute configurations of **14** and **52** were also determined in this study by comparison to authentic standard samples by GC analysis on a chiral stationary phase.

Site-directed mutagenesis was also performed on IWS. The results showed that the enzyme variant IWS\_N126D has a higher production (142 ± 20%) of **19** compared to the wild type. In addition, several fragment ions of isoishwarane were also elucidated based on the 15 experiments with the singly <sup>13</sup>C labelled isotopomers of FPP.

In this study, I determined the absolute configuration of **52** by comparison to an authentic standard through GC on a chiral stationary phase. I also carried out the site-directed mutagenesis study on IWS and tested the activities of all the enzyme variants.



## Chapter 9

### The Sesquiterpene Synthase PtTPS5 Produces (1*S*,5*S*,7*R*,10*R*)-Guaia-4(15)-en-11-ol and (1*S*,7*R*,10*R*)-Guaia-4-en-11-ol in Oomycete-Infected Poplar Roots

Nathalie D. Lackus<sup>[a],#</sup>, Jennifer Morawetz<sup>[a],#</sup>, Houchao Xu<sup>[b]</sup>, Jonathan Gershenzon<sup>[a]</sup>, Jeroen S. Dickschat<sup>[b]</sup> and Tobias G. Köllner<sup>[a],\*</sup>

[a] Max Planck Institute for Chemical Ecology, Hans-Knöll-Strasse 8, D-07745 Jena, Germany. Email: koellner@ice.mpg.de.

[b] Kekulé-Institute for Organic Chemistry and Biochemistry, University of Bonn, Gerhard-Domagk-Straße 1, 53121 Bonn, Germany.

# These authors contributed equally to this work.

Reprinted from *Molecules* **2021**, *26*, 555 with kind permission from MDPI

The publication “The Sesquiterpene Synthase PtTPS5 Produces (1*S*,5*S*,7*R*,10*R*)-Guaia-4(15)-en-11-ol and (1*S*,7*R*,10*R*)-Guaia-4-en-11-ol in Oomycete-Infected Poplar Roots” can be found in Appendix H.

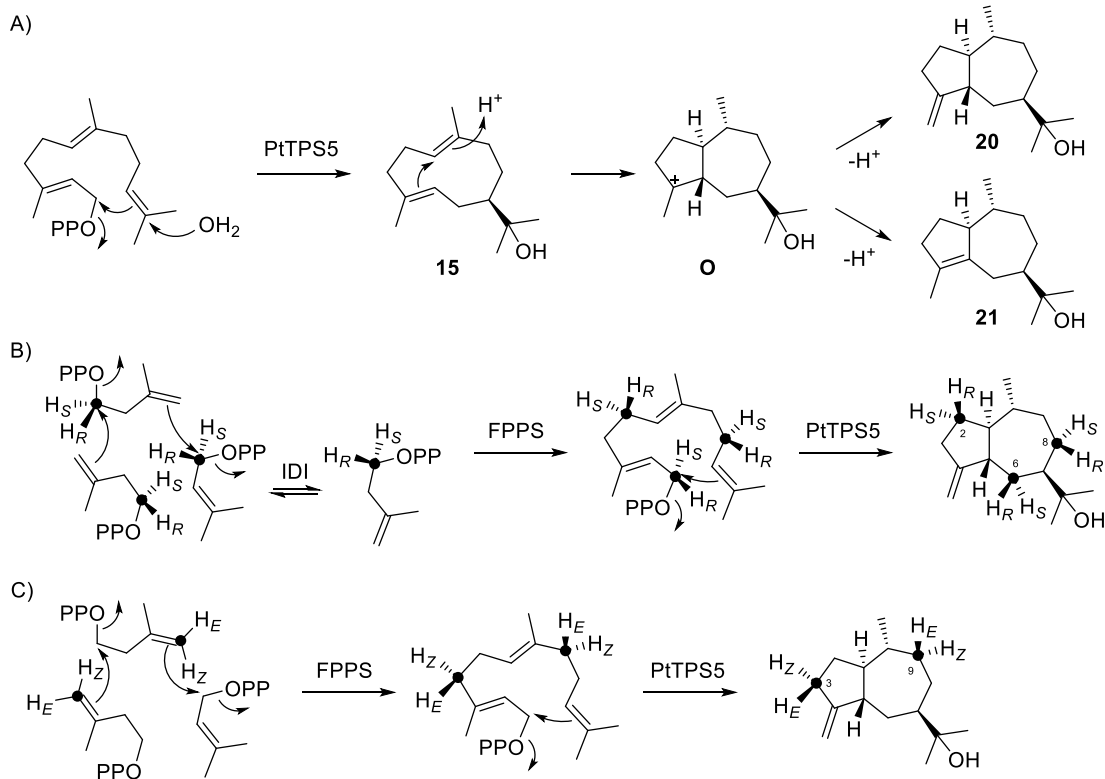


Plants are endangered by infections with various plant pathogenic microorganisms. In order to defend themselves against such infections, plants can produce secondary metabolites termed phytoalexins such as salicinoids that are phenolic glycosides usually occurring in willows and poplars.<sup>[144]</sup> As the largest class of natural products, terpenoids also function as specialised metabolites of plant defense. For example, the sesquiterpene-derived zealexins and the diterpene-derived kauralexins from maizes play essential roles for plants in response to pathogen attack.<sup>[145][146]</sup>

In recent years, the formation of defense terpenes against insect herbivores in Western balsam poplar (*Populus trichocarpa*) has been investigated and so far nineteen relevant terpene synthase genes have been cloned and characterized.<sup>[147][98][148][149][150]</sup> However, it remained unclear if these genes were also up-regulated upon pathogen attack or not. This problem was investigated and resolved in the present study.

For this purpose, the generalist oomycete *Phytophthora cactorum* was selected as a model pathogen due to its broad host specificity and economic importance. The gas chromatography-mass spectrometry (GC/MS) analysis of the hexane extracts obtained from the roots of young *P. trichocarpa* trees infected with *P. cactorum* revealed that there were two unknown sesquiterpene alcohols besides traces of some known terpenes such as elemol which is most likely a rearrangement product of hedycaryol formed during GC-MS analysis. In addition, the amount of terpenes in the *P. cactorum*-infected roots was significantly higher compared to that of the uninfected control roots. A transcriptome analysis through RT-qPCR indicated that a sesquiterpene synthase gene PtTPS5 was highly expressed in the roots of *P. trichocarpa* infected with *P. cactorum*, but not in the non-infected roots. The enzymatic reactions with recombinant PtTPS5 and farnesyl pyrophosphate (FPP) gave two major products which were identified as the terpene alcohols guaia-4(15)-en-11-ol (**20**) and guaia-4-en-11-ol (**21**) (Scheme 28), together with a minor product, hedycaryol. After comparing the retention times and mass spectra of **20** and **21** with those of the two unidentified sesquiterpene alcohols detected in *P. cactorum*-infected poplar roots, the conclusion could be drawn that PtTPS5 produces **20** and **21** which could also be detected in *P. cactorum*-infected poplar

roots. Therefore, it was proposed that PtTSP5 sesquiterpenes or their potential conversion products were involved in the defense against *P. cactorum* infection in poplar roots.



**Scheme 28.** A) Biosynthetic model for the cyclisation from FPP to the sesquiterpene alcohols. Determination of the absolute configuration of **20** by enantioselective deuteration using B) (*R*)- or (*S*)-(1-<sup>13</sup>C,1-<sup>2</sup>H)IPP with IDI, FPPS and PtTSP5, and C) (*E*)- or (*Z*)-(4-<sup>13</sup>C,4-<sup>2</sup>H)IPP with FPPS and PtTSP5.

Meanwhile, a plausible biosynthetic pathway towards **20** and **21** was proposed. According to this proposal, a 1,10-cyclisation happens to FPP, followed by attack of water to form hedycaryol. Then this neutral intermediate proceeds with a cyclisation initiated by the reprotonation at C10 to reach the guaiane skeleton. Finally, deprotonations happen from C15 or C5 to yield **20** and **21**. Thus, the biosynthesis of **20** and **21** gives another example of sesquiterpene alcohols originating from hedycaryol upon reprotonation.<sup>[151]</sup> In addition, the absolute configuration of **20** and **21** was also determined using stereoselectively deuterated substrates, including (*R*)- or (*S*)-(1-<sup>13</sup>C,1-<sup>2</sup>H)IPP<sup>[89]</sup> and (*E*)- or (*Z*)-(4-<sup>13</sup>C,4-<sup>2</sup>H)IPP<sup>[129]</sup> (Scheme 28). The HSQC analyses of the obtained products by the enzymatic conversions of these probes indicated that the absolute

configurations were proved to be (1*S*,5*S*,7*R*,10*R*)-**1** and (1*S*,7*R*,10*R*)-**2** with the former being a new compound.

Apart from terpenes, other products such as aromatic compounds and fatty acids were also detected in the hexane extracts of the infected root. Notably, almost all the fatty acids, two aromatic compounds, i.e. benzyl alcohol and 2-phenylethanol, and myristaldehyde showed a significant up-regulation upon the pathogen infection. Salicinoids was also detected in the extracts using high performance liquid chromatography (HPLC)-UV and liquid chromatography-tandem mass spectrometry (LC-MS/MS). However, the accumulation of most of the detected salicinoids was not affected by the oomycete infection with an exception that salicin-7-sulfate presented a small but significant induction.

My contribution to this study includes the isolation, purification and structure elucidation of **20** and **21** as well as conducting all the labelling experiments to determine the absolute configurations of them.





## Chapter 10

### 1,2- or 1,3-Hydride Shifts: What Controls Guaiane Biosynthesis?

Houchao Xu<sup>[a]</sup>, Bernd Goldfuss<sup>[b],\*</sup> and Jeroen S. Dickschat<sup>[a],\*</sup>

[a] Kekulé-Institute for Organic Chemistry and Biochemistry, University of Bonn, Gerhard-Domagk-Straße 1, 53121 Bonn, Germany. Email: [dickschat@uni-bonn.de](mailto:dickschat@uni-bonn.de).

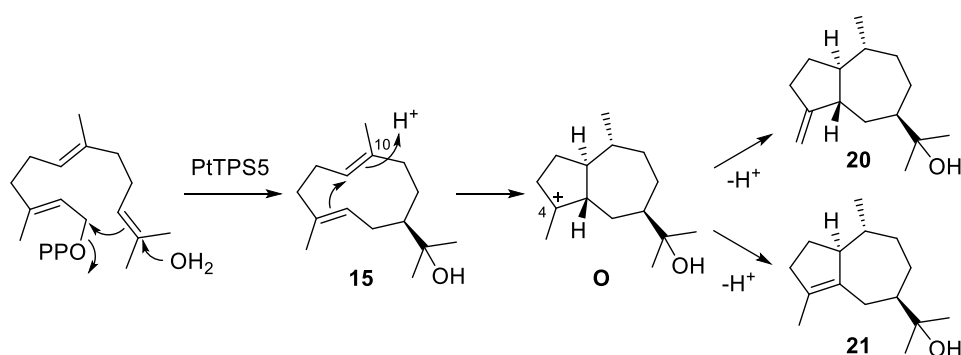
[b] Institute for Organic Chemistry, University of Cologne, Greinstraße 4, 50939 Cologne, Germany. E-mail: [goldfuss@uni-koeln.de](mailto:goldfuss@uni-koeln.de).

Reprinted from *Chem. Eur. J.* **2021**, 27, 9758 with kind permission from John Wiley & Sons

The publication “1,2- or 1,3-Hydride Shifts: What Controls Guaiane Biosynthesis?” can be found in Appendix I.



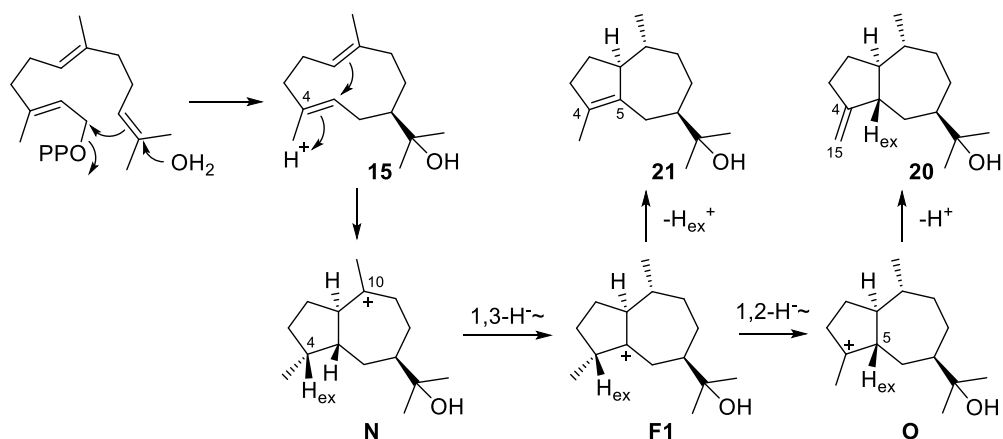
The previous research article introduced two sesquiterpenes, (1*S*,5*S*,7*R*,10*R*)-guaia-4(15)-en-11-ol (**20**) and (1*S*,7*R*,10*R*)-guaia-4-en-11-ol (**21**) (Scheme 29), generated by PtTPS5 from the roots of *Populus trichocarpa*.<sup>[119]</sup> The research showed that the expression of this enzyme generating these two products was up-regulated when the roots suffered from pathogen attack. Therefore, it was concluded that **20** and **21** or their potential conversion products were involved in the defense against the pathogen infection. In the same study, the absolute configurations of these two compounds were determined using isotopically <sup>13</sup>C labelled and deuterated substrates. The present study continues with this work on this enzyme and its enzymatic products, **20** and **21**, in the aspect of their biosynthetic pathways using an isotopic labelling strategy and extensive DFT calculations.



**Scheme 29.** The originally proposed biosynthetic pathways of **20** and **21**.

In the previous study, a biosynthetic pathway towards **20** and **21** was proposed (Scheme 29).<sup>[119]</sup> In that published mechanism, a reprotonation at C10 of hedycaryol (**15**) triggers the ring closure to yield the cationic intermediate **O**. Then alternative deprotonations of **O** give **20** and **21**. This pathway can be considered to be the most direct one from FPP towards **20** and **21**. However, when D<sub>2</sub>O was used for the enzymatic reactions in this study, it was found that the deuterium was incorporated at C4 instead of C10. This result indicated that the reprotonation of **15** happens at C4 which is contradictory to the proposed biosynthetic steps. According to the review article summarising the sesquiterpenes derived from hedycaryol (Chapter 3), the reprotonation of **15** to initiate a second cyclisation cascade can either proceed at C4 or at C10, and sesquiterpenes that are produced via both pathways have also been

discovered.<sup>[151]</sup>



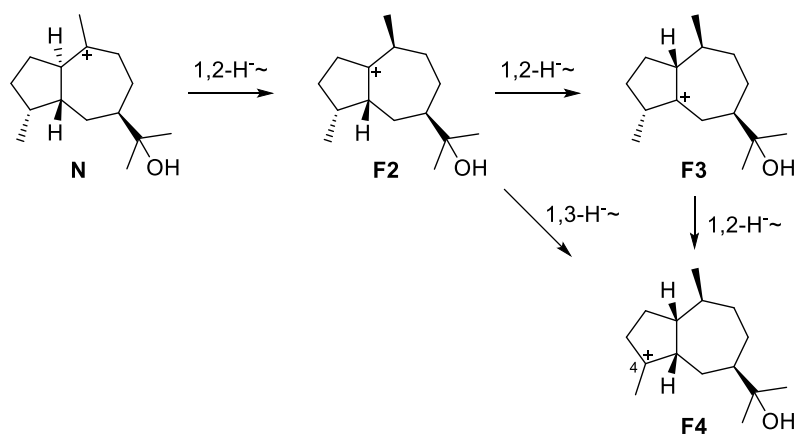
**Scheme 30.** The modified biosynthetic pathways of **20** and **21**.

Based on the experimental result of deuterium incorporation, the cyclisation mechanism was hence modified (Scheme 30). In this new mechanism, after the ring closure induced by the reprotonation of **15** at C4, the cation resides at C10 in **N**. The final compounds **20** and **21** have a double bond at C4/C15 or C4/C5, respectively, so the direct deprotonation of **N** gives neither of these two products. Therefore, hydride shifts are needed for the deprotonation at the corresponding carbons to obtain **20** and **21**. As a consequence, a 1,3-hydride shift to reach **F1** was proposed for the deprotonation to **21**. Another 1,2-hydride shift of **F1** to **O** was suggested for the deprotonation to **20**. These hypotheses were also verified by the incubation experiment of (2-<sup>2</sup>H)DMAPP<sup>[152]</sup> and (3-<sup>13</sup>C)IPP<sup>[74]</sup> with PtTPS5, isopentenyl diphosphate isomerase (IDI) from *Escherichia coli*<sup>[153]</sup> and FPP synthase (FPPS) from *Streptomyces coelicolor*.<sup>[154]</sup> Notably, 1,3-hydride shifts were rarely experimentally established for the biosynthesis of sesquiterpenes.

Nevertheless, hydride shifts occur frequently in sesquiterpene biosynthesis. Apart from the hydride shifts discussed above, in principle, **N** can also undergo three 1,2-hydride shifts (**N**–**F2**–**F3**–**F4**) or a combination of a 1,2- (**N**–**F2**) and a 1,3-hydride shift (**F2**–**F4**) to reach an intermediate in which the cation is at C4 (Scheme 31). It is worthwhile to mention that different combinations of hydride shifts often result in intermediates with different stereochemistry, because hydride shifts must proceed suprafacially and thus the stereochemistry of the

starting intermediate will be reflected in later stage intermediates derived from it. In order to investigate if hydride shifts, especially 1,3-hydride shifts, can indeed happen to any cationic intermediate of guaiane sesquiterpenes, a DFT calculation study was performed to follow the possible 1,2- or 1,3-hydride shifts of all the guaiane cationic intermediates originating from germacrene A and hedycaryol systemically. The results showed that 1,2-hydride shifts are much preferred because the energy barriers for the 1,2-hydride shift in most of the intermediates are low, while according to the computational study, 1,3-hydride shifts are only possible in a few cases of *trans*-fused guaianes.

In addition, the DFT calculation was also applied to investigate the problem how hedycaryol is reprotonated to proceed with the ring closure. The result indicated that two molecules of water may mediate the reprotonation process which shares the same mechanism as that for the biosynthesis of myrothec-15(17)-en-7-ol catalysed by its synthase from *Myrothecium gramineum*.<sup>[155]</sup>



**Scheme 31.** The possible hydride shifts happening to **N**.

In this study, I was responsible to perform all the labelling experiments to investigate the biosynthetic pathway, including the reprotonation, and the 1,2- and 1,3-hydride shifts, towards **20** and **21**.



## Chapter 11

### The enzyme mechanism of patchoulol synthase

Houchao Xu<sup>[a]</sup>, Bernd Goldfuss<sup>[b]</sup>, Gregor Schnakenburg<sup>[c]</sup> and Jeroen S. Dickschat<sup>[a].\*</sup>

[a] Kekulé-Institute for Organic Chemistry and Biochemistry, University of Bonn, Gerhard-Domagk-Straße 1, 53121 Bonn, Germany. Email: [dickschat@uni-bonn.de](mailto:dickschat@uni-bonn.de).

[b] Department of Chemistry, University of Cologne, Greinstraße 4, 50939 Cologne, Germany.

[c] Institute of Inorganic Chemistry, University of Bonn, Gerhard-Domagk-Straße 1, 53121 Bonn, Germany.

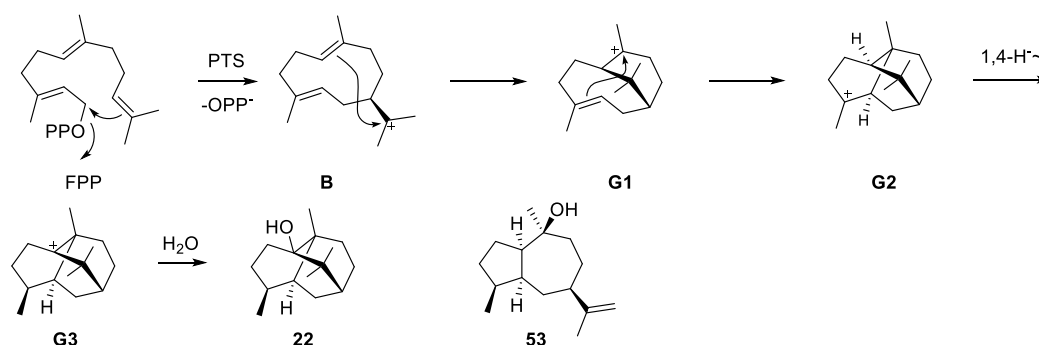
Reprinted from *Beilstein J. Org. Chem.* **2022**, *18*, 13 with kind permission from Beilstein Institute for the Advancement of Chemical Sciences

The publication “The enzyme mechanism of patchoulol synthase” can be found in Appendix J.





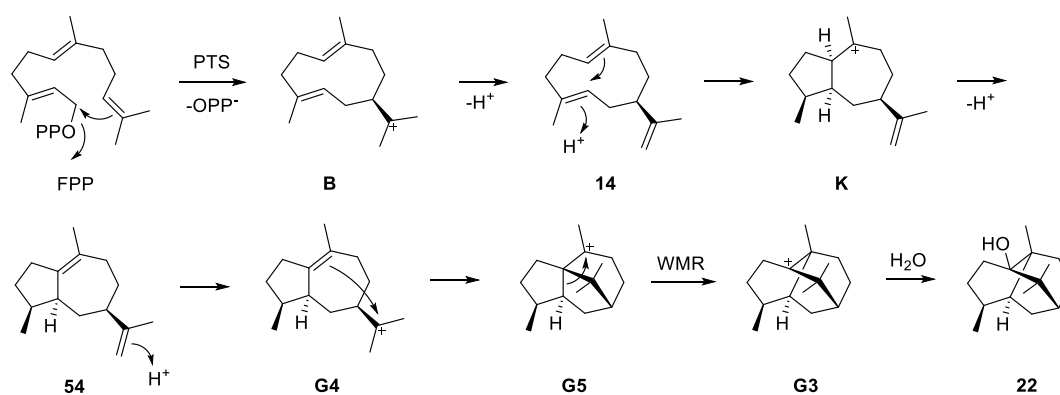
Patchoulol (**22**, Scheme 32), a sesquiterpene alcohol, is a major component in the essential oil of *Pogostemon cablin*. This compound was first reported with a wrong assignment of its planar structure.<sup>[156]</sup> After its structure was revised twice, the structure of patchoulol was finally presented as **22** with its absolute configuration still undetermined.<sup>[157]</sup> Compared to the investigated bicyclic sesquiterpene alcohols, guaia-4(15)-en-11-ol and guaia-4-en-11-ol, described in the previous chapters, patchoulol features a more complex structure because of its rearranged tricyclic skeleton.



**Scheme 32.** Cyclisation mechanism from FPP to patchoulol (**22**) as suggested by Croteau et al., and the structure of pogostol (**53**).

Several different biosynthetic pathways from FPP towards **22** have been proposed with the representative ones reported by Croteau et al. in 1987<sup>[61]</sup>, Akhila et al. in 1988<sup>[62]</sup> and Faraldos et al. in 2010.<sup>[63]</sup> These pathways were suggested based on isotopic labelling experiments, but proceed differently after the formation of the (E,E)-germacradienyl cation (**B**) from FPP. In the pathway proposed by Croteau et al. (Scheme 32), there is an unusual 1,4-hydride shift before the attack of water to reach **22**. As discussed in the previous chapter, in the biosynthesis of sesquiterpenes 1,2-hydride shifts are usually favoured over 1,3-hydride shifts,<sup>[158]</sup> while 1,4-hydride shifts have rarely been experimentally confirmed. The pathway suggested by Akhila et al. is characterised by two neutral intermediates, i. e. compounds **14** and **54** produced by the deprotonations of **B** and **K**, respectively. These two intermediates need to be reprotonated to continue with the reaction cascade (Scheme 33). The reprotonation steps were later confirmed by the enzymatic conversion with FPP in D<sub>2</sub>O buffer performed by Ekramzadeh et al..<sup>[159]</sup> The pathway proposed by

Faraldos et al. shows a 1,3-hydride shift from **G6** to **G8** instead of a sequential 1,2-hydride shift based on an experiment with (2-<sup>2</sup>H)FPP (Scheme 34). However, the enzymatic conversion of (2-<sup>2</sup>H)FPP also gave a double deuterated patchoulol with an additional deuterium atom incorporated at C15. An unusual mechanism was suggested for this phenomenon, but the reason for this finding is likely the presence of additional deuterium in the substrate introduced in the synthesis. Compared to the first discussed pathway, the second and third ones both proposed Wagner-Meerwein rearrangements (WMR) to obtain the complex skeleton of **22**.

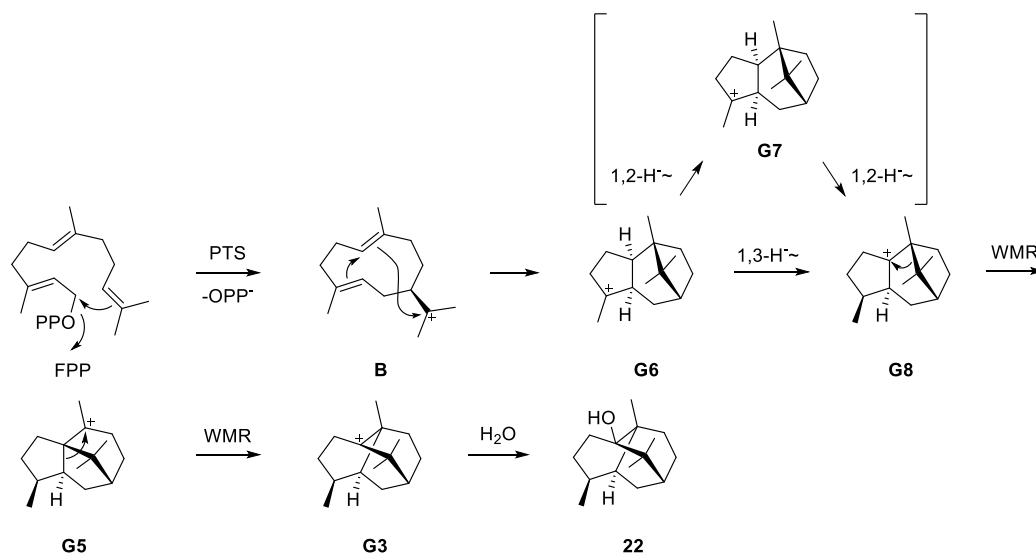


**Scheme 33.** Cyclisation mechanism from FPP to **22** as suggested by Akhila et al..

Taken together, all three biosynthetic pathways can theoretically explain the formation of **22** from FPP. However, they are likely not all practically valid to yield the same product, although some special cases indeed exist in which different pathways can lead to the same compound. For instance, as shown in Chapter 6, (4*S*,7*R*)-germacra-(1(10)*E*,5*E*)-dien-11-ol is produced by its synthases from *Dictyostelium purpureum* and *Streptomyces coelicolor* through different biosynthetic pathways,<sup>[160]</sup> but here different mechanisms occur for different enzymes. Patchoulol synthase may likely act through one clearly defined mechanism. All the hydride shifts actually were also not followed directly by labelling experiments. Therefore, it still remains unclear which pathway is the correct one.

The purpose of this study is to clarify the confusing situation described above using isotopically labelled substrates, especially to follow the reprotonations and hydride shifts precisely. Notably, a combination of <sup>13</sup>C and deuterium

labelling strategy was used in this study to follow the reprotonations and hydride shifts. For instance, the usage of D<sub>2</sub>O buffer and (3-<sup>13</sup>C)FPP for the enzymatic conversion gave the evidence that the deuterium from the buffer was incorporated at C3 in **22** since the <sup>13</sup>C NMR showed a shifted triplet (<sup>1</sup>J<sub>C,D</sub> = 19.4 Hz, Δδ = -0.45 ppm). Then (12-<sup>13</sup>C)FPP in D<sub>2</sub>O was used to indicate the other deuterium was introduced at C12. These two experiments demonstrated that there should be two reprotonations in the biosynthesis of **22**, which excludes the pathways proposed by Croteau et al. and by Faraldos et al., but supports the one by Akhila et al.. By using additional labelled substrates, all the hydride shifts proposed for the mechanism by Akhila et al. were also confirmed. DFT calculations carried out in this study also demonstrated the rationality of this pathway.



**Scheme 34.** Cyclisation mechanism from FPP to **22** as suggested by Faraldos et al..

In addition, compound **22** isolated from patchouli oil was crystallised and analysed by X-ray, which allowed the determination of its absolute configuration. The absolute configuration of pogostol, a side product of the patchoulol synthase was also determined in this study using the isotopically labelled probes, (*E*)- or (*Z*)-(4-<sup>13</sup>C,4-<sup>13</sup>C)IPP and (*R*)- or (*S*)-(1-<sup>13</sup>C,1-<sup>2</sup>H)IPP.

In this study, I isolated patchoulol, pogostol and a new compound, guaia-1,11-dien-1-ol, from patchouli oil. Then I conducted all the labelling experiments to distinguish unequivocally between the three proposed pathways and to determine to absolute configurations of **22** and **53**.



## Chapter 12

Stereochemical characterisation of the non-canonical  $\alpha$ -humulene synthase from *Acremonium strictum*

Houchao Xu<sup>[a]</sup>, Carsten Schotte<sup>[b]</sup>, Russell J. Cox<sup>[b]</sup> and Jeroen S. Dickschat<sup>[a,\*]</sup>

[a] Kekulé-Institute for Organic Chemistry and Biochemistry, University of Bonn, Gerhard-Domagk-Straße 1, 53121 Bonn, Germany. E-mail: dickschat@uni-bonn.de.

[b] Institute of Organic Chemistry, University of Hannover, Schneiderberg 38, 30167 Hannover, Germany.

Reprinted from *Org. Biomol. Chem.* **2021**, *19*, 8482 with kind permission from Royal Society of Chemistry

The publication “Stereochemical characterisation of the non-canonical  $\alpha$ -humulene synthase from *Acremonium strictum*” can be found in Appendix K.

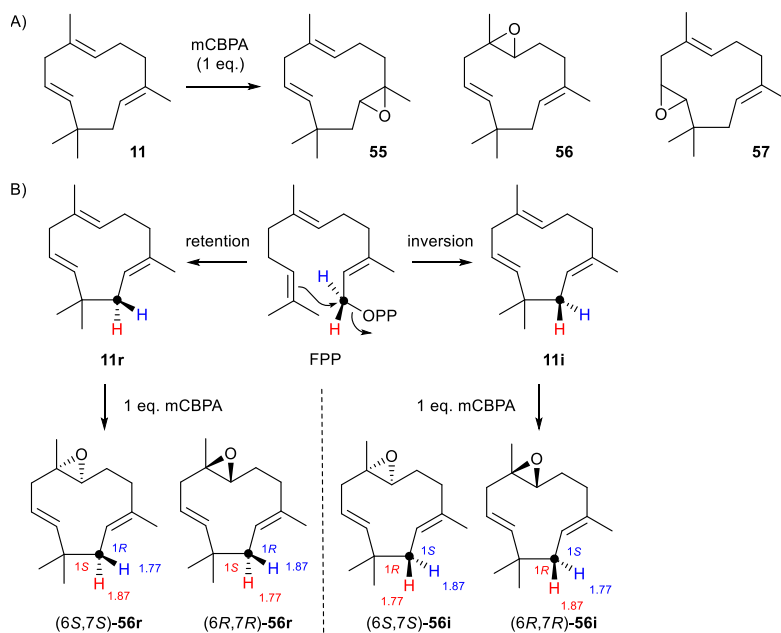


$\alpha$ -Humulene (**11**) (Scheme 35), one of the rare known achiral terpenes, was first isolated from *Humulus lupulus* in 1895<sup>[161]</sup> and its structure was correctly assigned in 1961.<sup>[162]</sup> So far  $\alpha$ -humulene synthases have been reported from plants, e.g. *Zingiber zerumbet*<sup>[163]</sup> and *Aquilaria crassna*,<sup>[164]</sup> and fungi, e.g. AsR6 from *Acremonium strictum*.<sup>[165]</sup> Biosynthetically the formation of  $\alpha$ -humulene from farnesyl pyrophosphate (FPP) only needs a simple 1,11-cyclisation followed by deprotonation. Despite the simplicity of this biosynthetic pathway there were still three challenging aspects due to the  $C_s$  symmetry of this compound which needed to be investigated, i.e. (1) inversion or retention of configuration at C1 during cyclisation, (2) attack at C-11 from the *Si* or the *Re* face for the same step, and (3) the abstraction of the 9-*pro-R* or 9-*pro-S* hydrogen in the final deprotonation. In this work, these problems were addressed using isotopic labelling experiments in conjunction with a desymmetrisation strategy.

Firstly, two enantioselectively deuterated substrates (*R*)- and (*S*)-(1-<sup>13</sup>C,1-<sup>2</sup>H)FPP were used to solve the inversion versus retention problem. These two substrates were enzymatically converted into (1-<sup>13</sup>C,1-<sup>2</sup>H)-**11** with incorporation of deuterium into enantiotopic positions, which cannot be distinguished by NMR spectroscopy. However, this problem could be solved once **11** was epoxidized by one equivalent of *meta*-chloroperbenzoic acid (*m*CPBA) (Scheme 35). The epoxidation products, i.e. humulene epoxides I, II and III (**55–57**) of enantioselectively deuterated (1-<sup>13</sup>C,1-<sup>2</sup>H)-**11** are respectively two diastereomers that could be separated by HPLC on a chiral stationary phase and identified based on their optical rotations. This allowed to follow the incorporations of deuterium into the diastereotopic positions of the epoxides by NMR spectroscopy. Eventually, the results showed inversion of configuration at C1 during the cyclisation of FPP to **11** by AsR6.

Secondly, the face selectivity at C11 during FPP cyclisation regarding *Re* or *Si* face attack was investigated by a similar approach (Scheme 36). In this study, (12-<sup>13</sup>C)FPP and (13-<sup>13</sup>C)FPP were incubated with AsR6 to give (12-<sup>13</sup>C)-**11** and (13-<sup>13</sup>C)-**11**. After mixing with unlabelled-**11** and epoxidation with *m*CPBA, the major products, humulene epoxides II (**56**), were separated by HPLC on a chiral stationary phase, yielding the diastereoisomers of labelled [12-<sup>13</sup>C]-

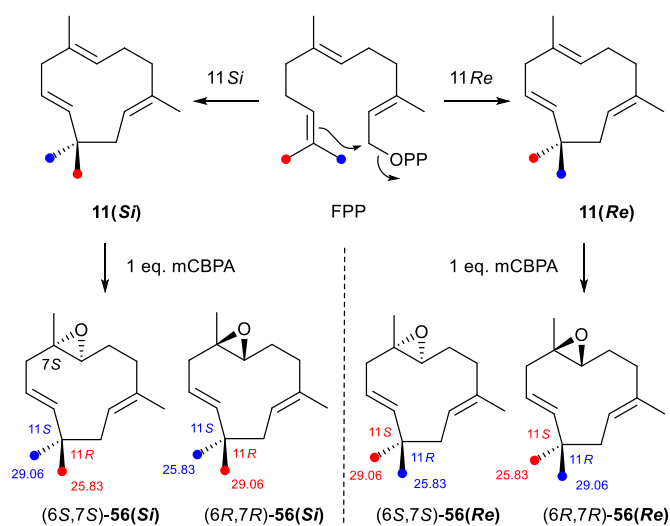
(6*R*,7*R*)-(-)-**56** and [13-<sup>13</sup>C]-(6*S*,7*S*)-(+)-**56**. The <sup>13</sup>C-NMR spectroscopy revealed that (11*R*)-configuration for both stereoisomers of labelled **56** was obtained from (12-<sup>13</sup>C)FPP, and (11*S*)-configuration for both stereoisomers of labelled **56** was obtained from (13-<sup>13</sup>C)FPP, which was in agreement with *Re* face attack at C11 in the FPP cyclisation by AsR6.



**Scheme 35.** A). Epoxidation of **11** with mCPBA yielding racemic humulene epoxides **55–57**. B) Enzymatic conversion of (*R*)-(1-<sup>13</sup>C,1-<sup>2</sup>H)FPP (blue H = <sup>2</sup>H) and (*S*)-(1-<sup>13</sup>C,1-<sup>2</sup>H)FPP (red H = <sup>2</sup>H) into **11** with retention or inversion of configuration at C1 and subsequent desymmetrisation by conversion into humulene epoxide II (**56**) with mCPBA. Black dots indicate <sup>13</sup>C-labelled carbons. The stereochemical descriptors for C1 shown in blue are for the labelled compounds from (*R*)-(1-<sup>13</sup>C,1-<sup>2</sup>H)FPP and those in red are for the labelled compounds from (*S*)-(1-<sup>13</sup>C,1-<sup>2</sup>H)FPP. Data at the labelled hydrogens indicate <sup>1</sup>H-NMR shifts in ppm.

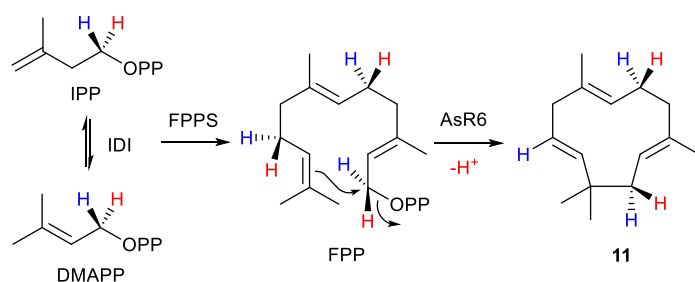
Moreover, the final deprotonation step was also investigated for the stereochemical course using the substrates (*R*)- and (*S*)-(1-<sup>2</sup>H)isopentenyl diphosphate (IPP) which were enzymatically converted by isopentenyl diphosphate isomerase (IDI) from *Escherichia coli*, followed by FPP synthase (FPPS) from *Streptomyces coelicolor* and AsR6 (Scheme 37). GC/MS analysis indicated the loss of deuterium specifically from position 9-*pro-S* of FPP.





**Scheme 36.** The stereochemical course of the cyclisation of FPP to **11** by AsR6 for C11. Enzymatic conversion of ( $12\text{-}^{13}\text{C}$ )FPP (blue dots =  $^{13}\text{C}$ ) and ( $13\text{-}^{13}\text{C}$ )FPP (red dots =  $^{13}\text{C}$ ) into **11** with  $11Si$  or  $11Re$  attack. The stereochemical descriptors for C11 shown in blue are for the labelled compounds from ( $12\text{-}^{13}\text{C}$ )FPP and those in red are for the labelled compounds from ( $13\text{-}^{13}\text{C}$ )FPP. Data at the labelled carbons indicate  $^{13}\text{C}$ -NMR shifts in ppm.

In this work, I was responsible for conducting all the experiments including the chemical synthesis of humulene epoxides and the enzymatic reactions with labelled substrates mentioned above.



**Scheme 37.** The stereochemical course of the cyclisation of FPP to **11** by AsR6 for the deprotonation from C9. Enzymatic conversion of ( $R$ )-( $1\text{-}^2\text{H}$ )IPP (blue H =  $^2\text{H}$ ) and ( $S$ )-( $1\text{-}^2\text{H}$ )IPP (red H =  $^2\text{H}$ ) with IDI, FPPS and AsR6.



## Chapter 13

### Biosynthesis of the Sesquiterpene Kitaviridene through Skeletal Rearrangement with Formation of a Methyl Group Equivalent

Houchao Xu<sup>[a]</sup>, Bernd Goldfuss<sup>[b]</sup> and Jeroen S. Dickschat<sup>[a],\*</sup>

[a] Kekulé-Institute for Organic Chemistry and Biochemistry, University of Bonn, Gerhard-Domagk-Straße 1, 53121 Bonn, Germany. Email: [dickschat@uni-bonn.de](mailto:dickschat@uni-bonn.de).

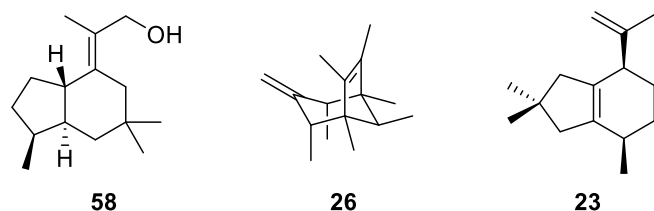
[b] Institute for Organic Chemistry, University of Cologne, Greinstraße 4, 50939 Cologne, Germany.

Reprinted from *Org. Lett.* **2023**, 25, 3330 with kind permission from American Chemical Society

The publication “Biosynthesis of the Sesquiterpene Kitaviridene through Skeletal Rearrangement with Formation of a Methyl Group Equivalent” can be found in Appendix L.



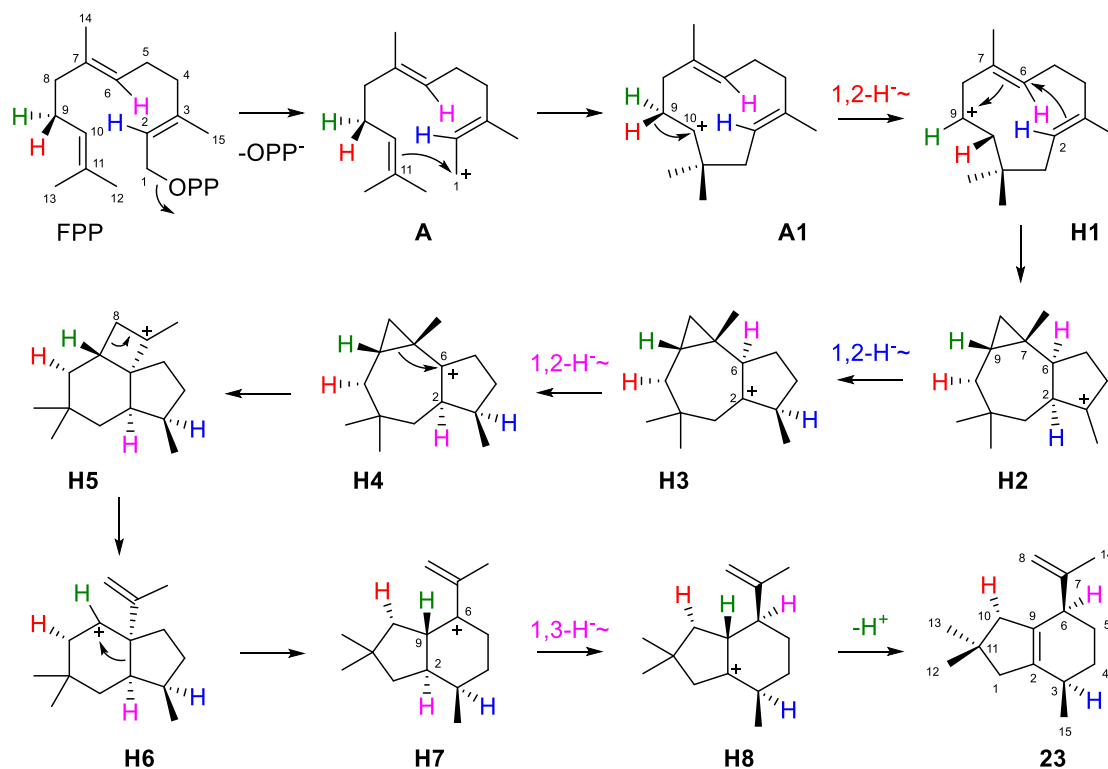
In the previous Chapter 8, isoishwarane was described as a special sesquiterpene that only has three methyl groups, while its precursor FPP has four methyl groups. A plausible cyclisation mechanism was proposed and then experimentally proved that can explain the reduced number of methyl groups. Interestingly, there are also some unusual sesquiterpenes that possess more than four methyl groups, such as trichobrasilenol (**58**, Figure 12) featuring four methyl groups and one hydroxymethyl group regarded as a methyl group equivalent,<sup>[166]</sup> and sodorifen (**26**, Figure 12) characterised by the presence of seven methyl groups originating from FPP.<sup>[67]</sup> In this study, a new terpene synthase was obtained from *Kitasatospora viridis*, an actinomycete isolated in 2005 from soil.<sup>[167]</sup> This new synthase can convert FPP into a sesquiterpene, but exhibits no effect on GPP, GGPP or GFPP. The large scale enzymatic reaction allowed the isolation of this natural product. The structure elucidation based on the 1D and 2D NMR data indicates that this compound is a 5/6-bicyclic sesquiterpene hydrocarbon, which was designated as kitaviridene (**23**, Figure 12). Its synthase was accordingly named as kitaviridene synthase (KvKS). Notably, kitaviridene owns four methyl groups and one terminal olefin that theoretically also equals one methyl group.



**Figure 12.** Structures of trichobrasilenol (**58**), sodorifen (**26**) and kitaviridene (**23**).

A biosynthetic pathway towards kitaviridene was proposed in this study (Scheme 38). This pathway starts with the departure of the diphosphate from FPP to give the reactive farnesyl cation **A**. A 1,11-cyclisation of **A** reaches the (*E,E*)-humulyl cation **A1**. After a 1,2-hydride shift from C9 to C10 (**A1**–**H1**), a 7,9- followed by a 2,6-cyclisation of **H1** was suggested to yield **H2**. Then two sequential 1,2-hydride shifts give the cationic intermediate **H4**. A Wagner–Meerwein rearrangement of **H4** leads to the formation of **H5**, an intermediate with a 4/6/5-fused ring system. Later on, the four-membered ring of **H5** is

opened to give **H6**, which converts the tricyclic ring system into a bicyclic ring system with the generation of a terminal double bond, namely a methyl group equivalent formed here. Further, a Wagner–Meerwein rearrangement was proposed to accomplish the conversion of a 6/5-membered skeleton into a 5/6-membered skeleton (**H6** to **H7**). Then a 1,3-hydride shift of **H7** followed by a final deprotonation gives kitaviridene.



**Scheme 38.** The biosynthetic pathway towards **23**.

The whole biosynthetic pathway was then investigated using isotopically  $^{13}\text{C}$  labelled and/or deuterated substrates. In order to gain experimental evidence for the 1,2-hydride shift from **A1** to **H1**, (1,1- $^2\text{H}_2$ ,2- $^{13}\text{C}$ )DMAPP and IPP were utilised for the enzymatic reaction with FPPS and KvKS. In this reaction, FPPS first biosynthesised (9,9- $^2\text{H}_2$ ,10- $^{13}\text{C}$ )FPP using one molecule of (1,1- $^2\text{H}_2$ ,2- $^{13}\text{C}$ )DMAPP and two molecules of IPP. Then KvKS cyclised the product from FPPS to the  $^{13}\text{C}$  labelled and deuterated kitaviridene. The  $^{13}\text{C}$  NMR of the extracts from the reaction mixture gave an enhanced up-field shifted triplet for C10, meaning that one of the deuterium atoms from C9 migrates to C10 during the cyclisation cascade. The HSQC spectrum of the sample only showed an

enhanced cross peak for H10 $\beta$ , while the signal of H10 $\alpha$  was vanished, meaning the position for H10 $\alpha$  was occupied by the migrated deuterium. Moreover, the GC/MS analyses of the samples prepared from the enzymatic conversions with (*R*)- or (*S*)-(1-<sup>2</sup>H,1-<sup>13</sup>C)IPP by IDI, FPPS and KvKS indicated that it is H9 $\beta$  that shifts from C9 to C10. These experimental results hereby confirmed the proposed 1,2-hydride shift from **A1** to **H1**. The same strategy was used to follow all other hydride shifts with different labelled substrates. Interestingly, the labelling experiments showed that the proton which migrates from C6 to C2 via a 1,2-hydride shift (**H3–H4**) shifts back to C6 via a 1,3-hydride shift (**H7–H8**). The formation of the terminal olefin (**H4–H5–H6**) was established by the enzymatic conversion of (8-<sup>13</sup>C)FPP by KvKS. DFT calculations were also applied in this research to support the biosynthesis of kitaviridene.

In addition, the absolute configuration of kitaviridene was determined in this study using the isotopically labelled probes (*E*)- or (*Z*)-(4-<sup>2</sup>H,4-<sup>13</sup>C)IPP and (*R*)- or (*S*)-(1-<sup>2</sup>H,1-<sup>13</sup>C)IPP.

In this work, I cultivated the strain *Kitasatospora viridis*, isolated the genomic DNA from the organism, cloned the gene of KvKS using PCR technology and expressed it in *E.coli* BL21. Then I performed the enzymatic conversions with GPP, FPP, GGPP and GFPP by KvKS to identify it as a sesquiterpene synthase. Besides that, I isolated kitaviridene by conducting large scale enzymatic reaction and elucidated its structure. I was also responsible for carrying out all the labelling experiments to follow the hydride shifts, skeleton rearrangements and deprotonation in the biosynthetic pathway, and to determine the absolute configuration of kitaviridene.





## Chapter 14

### Fragmentation and [4 + 3] cycloaddition in sodorifen biosynthesis

Houchao Xu<sup>[a]</sup>, Lukas Lauterbach<sup>[a]</sup>, Bernd Goldfuss<sup>[b]</sup>, Gregor Schnakenburg<sup>[c]</sup> and Jeroen S. Dickschat<sup>[a],\*</sup>

[a] Kekulé-Institute for Organic Chemistry and Biochemistry, University of Bonn, Gerhard-Domagk-Straße 1, 53121 Bonn, Germany. Email: [dickschat@uni-bonn.de](mailto:dickschat@uni-bonn.de).

[b] Institute for Organic Chemistry, University of Cologne, Greinstraße 4, 50939 Cologne, Germany.

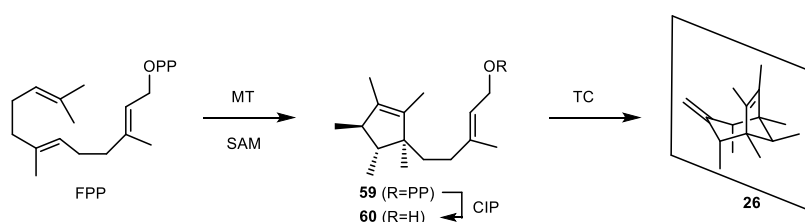
[c] Institute of Inorganic Chemistry, University of Bonn, Gerhard-Domagk-Straße 1, 53121 Bonn, Germany.

Reprinted from *Nat. Chem.* **2023**, *15*, accepted with kind permission from Springer Nature

The publication “Fragmentation and [4 + 3] cycloaddition in sodorifen biosynthesis” can be found in Appendix M.

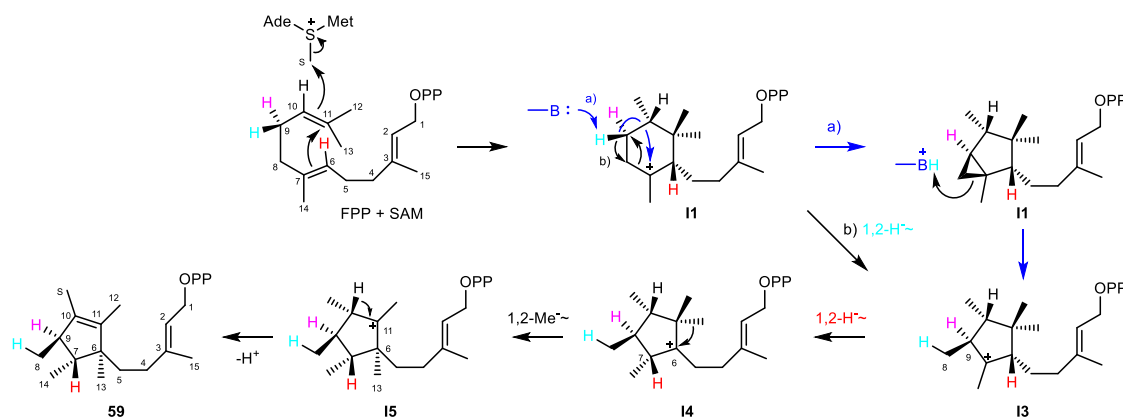


Sodorifen (**26**) is a highly methylated bicyclic sesquiterpene that was discovered from *Serratia odorifera* in 2010.<sup>[67]</sup> Distinguished from all other sesquiterpenes discussed in the previous chapters, this compound has 16 carbons with a symmetrical backbone so that only 10 signals can be observed from its <sup>13</sup>C NMR spectrum. In 2018, it was found that two enzymes are involved in the biosynthesis of sodorifen.<sup>[168]</sup> The first one, a methyl transferase (MT), induces an *S*-adenosyl methionine (SAM) dependent methylation of FPP that initiates a cyclisation cascade to presodorifen diphosphate (**59**), while the second enzyme, a terpene cyclase (TC), converts presodorifen pyrophosphate into sodorifen (Scheme 39).



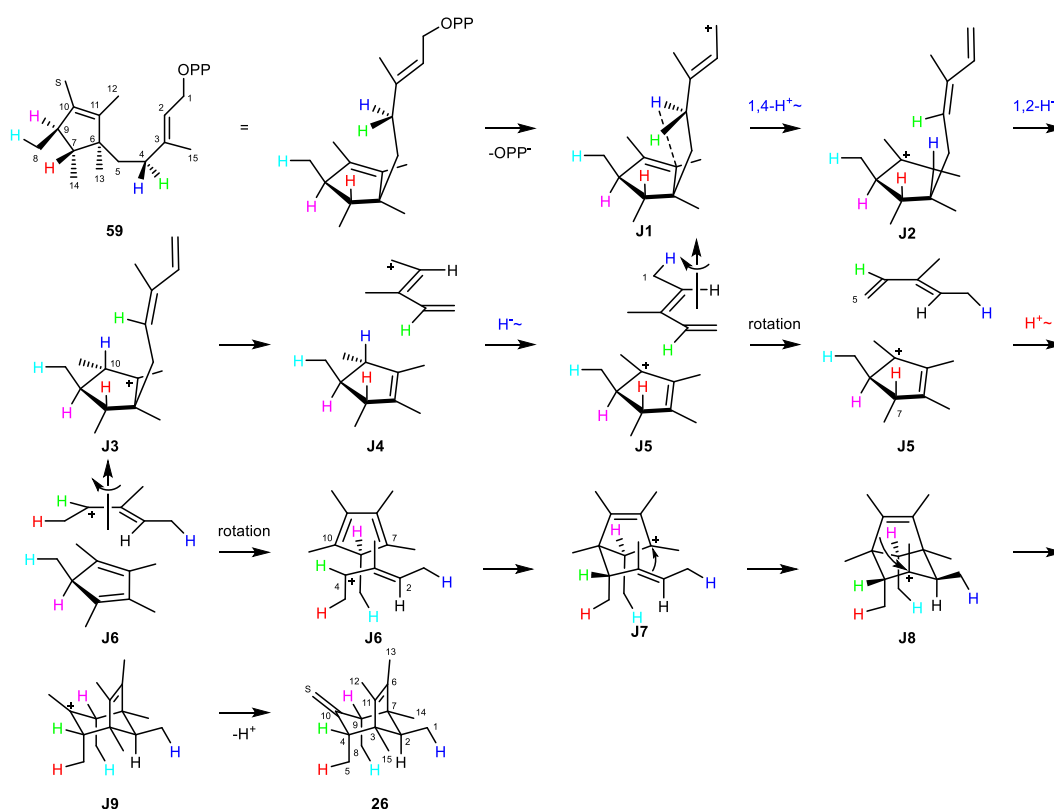
**Scheme 39.** Biosynthesis of sodorifen (**26**).

The proposed biosynthesis of presodorifen pyrophosphate is shown in Scheme 40. In this pathway, MT catalyses the methylation, cyclisation, hydride shifts, ring contraction, methyl group migration and final deprotonation. Labelling experiments were carried out in this study to support the cyclisation mechanism. The experimental results indicated that the additional methyl group is introduced to C10. Notably, H9 $\beta$  was proved to shift to C8, and C8 becomes a methyl group after the ring contraction.



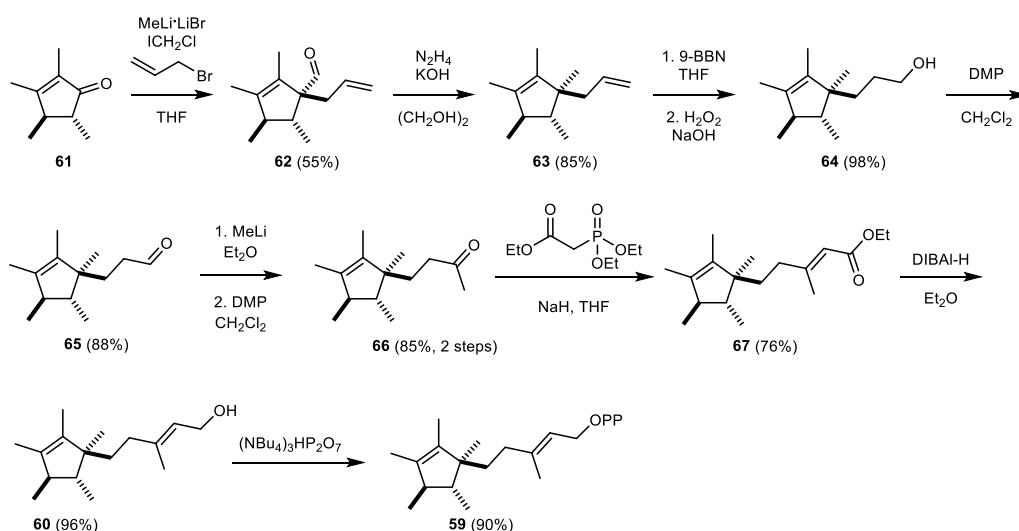
**Scheme 40.** Cyclisation mechanism from FPP to **59**.

In order to elucidate the biosynthesis of sodorifen, all 16 singly  $^{13}\text{C}$  labelled isotopomers were prepared to track the final location for each carbon in **26**. In addition, doubly or even multiply  $^{13}\text{C}$  labelled FPP isotopomers were synthesised chemically and enzymatically to figure out which carbons are connected to one another. Apart from the investigations on the carbons, the hydrogens were also followed based on several enzymatic conversions with different  $^{13}\text{C}$ -labelled and deuterated substrates. The results showed that the *pro-S* hydrogen at C4, the hydrogen at C6 and the *pro-R* hydrogen at C9 migrate to C1, C5 and C8, respectively. The hydrogen at C10 is lost in the cyclisation cascade. Apart from those, all other hydrogens remain at their original carbons. After all the carbons and hydrogens were disclosed for their locations in sodorifen, a biosynthetic pathway starting from presodorifen was hereby suggested (Scheme 41). It is worthwhile to mention that in this cyclisation mechanism, the fragmentation and [4 + 3] cycloaddition were for the first time proposed for the biosynthesis of terpenes. In addition, the hydride shifts, proton shifts, fragment rotation and final deprotonation were also proposed for this unusually complex reaction cascade to sodorifen.



**Scheme 41.** Cyclisation mechanism from **59** to **26**.

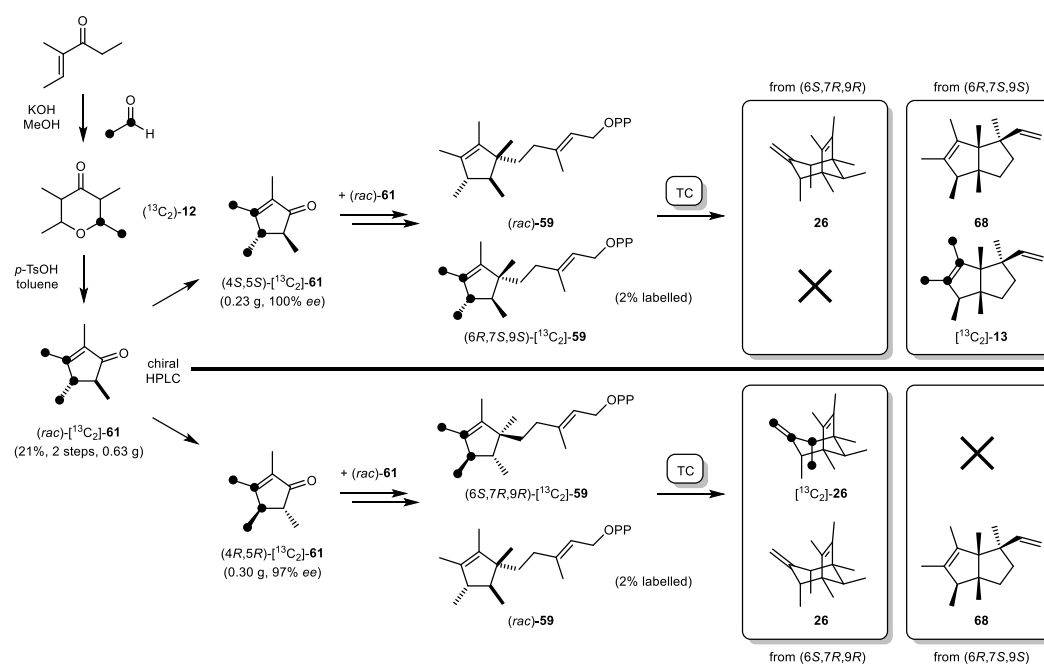
For a deeper understanding of the biosynthesis of presodorifen pyrophosphate and sodorifen, density functional theory computations were also performed in this study. All the computational results showed the rationality of these proposed pathways.



**Scheme 42.** Total synthesis of (*rac*)-**59**.

Although the biosynthesis of sodorifen was deciphered, the absolute configuration of **59** in our biosynthetic hypothesis was the opposite of the reported one which was determined based on a comparison of the measured to a calculated ECD spectrum.<sup>[169]</sup> Therefore, it seemed doubtful if the reported absolute configuration was the correct one. In order to solve this problem, a total synthesis of **59** was conducted (Scheme 42). For this purpose, (*rac*)-**59** was synthesised from the starting material (*4R*\*,*5R*\*)-2,3,4,5-tetramethylcyclopent-2-en-1-one (**61**). Unfortunately, the final product together with all the synthetic intermediates could not be separated into the two enantiomers using chiral HPLC, except compound **61** which could be separated in a limited amount at a cost of a long preparation time. Their chemical derivatisation as 2,4-dinitrophenylhydrazones, crystallisation and X-ray analysis provided the possibility to determine their absolute configurations as (*4S*,*5S*)-(-)-**61** and (*4R*,*5R*)-(+)-**61**. In addition, the separation of (*rac*)-[<sup>13</sup>C<sub>2</sub>]-**61**, a racemate synthesized from (<sup>13</sup>C<sub>2</sub>)-acetaldehyde (Scheme 43), was also achieved. Notably, (*4S*,*5S*)-[<sup>13</sup>C<sub>2</sub>]-**61** and (*4R*,*5R*)-[<sup>13</sup>C<sub>2</sub>]-**61** were used in this study to determine the absolute configuration of the natural precursor to

sodorifen as (6*S*,7*R*,9*R*)-**59** (Scheme 43), which confirmed the absolute configuration of presodorifen pyrophosphate as proposed in this study for the biosynthesis of **26**. Surprisingly, it was observed that TC can also convert the non-natural enantiomer, namely (6*R*,7*S*,9*S*)-**59**, into a new compound which was designated as enantiofen (**68**). The biosynthetic pathway of **68** was also discussed in this research.



**Scheme 43.** Labelling strategy to investigate the absolute configuration of natural **59**. Black dots represent <sup>13</sup>C-labelled carbons. For *rac*-[<sup>13</sup>C<sub>2</sub>]-**61** and materials derived therefrom, the <sup>13</sup>C-labelling was distributed over the two C2 units connected by red bonds.

In this work, I first optimised the synthetic route to (*rac*)-**59** and then synthesised (*rac*)-**61**, (*rac*)-[<sup>13</sup>C<sub>2</sub>]-**61**, (*rac*)-**59** with (6*R*,7*S*,9*S*)-[<sup>13</sup>C<sub>2</sub>]-**59** (2%, w/w) and (*rac*)-**59** with (6*R*,7*S*,9*S*)-[<sup>13</sup>C<sub>2</sub>]-**59** (2%, w/w). I also carried out the large scale enzymatic reactions to isolate **26**, [<sup>13</sup>C<sub>2</sub>]-**26**, **68** and [<sup>13</sup>C<sub>2</sub>]-**68** to determine the absolute configuration of **59**, and elucidated the structure of **68**.

## Chapter 15

### A Detailed View on Geosmin Biosynthesis

Houchao Xu<sup>[a]</sup>, and Jeroen S. Dickschat<sup>[a],\*</sup>

[a] Kekulé-Institute for Organic Chemistry and Biochemistry, University of Bonn, Gerhard-Domagk-Straße 1, 53121 Bonn, Germany. Email: [dickschat@uni-bonn.de](mailto:dickschat@uni-bonn.de).

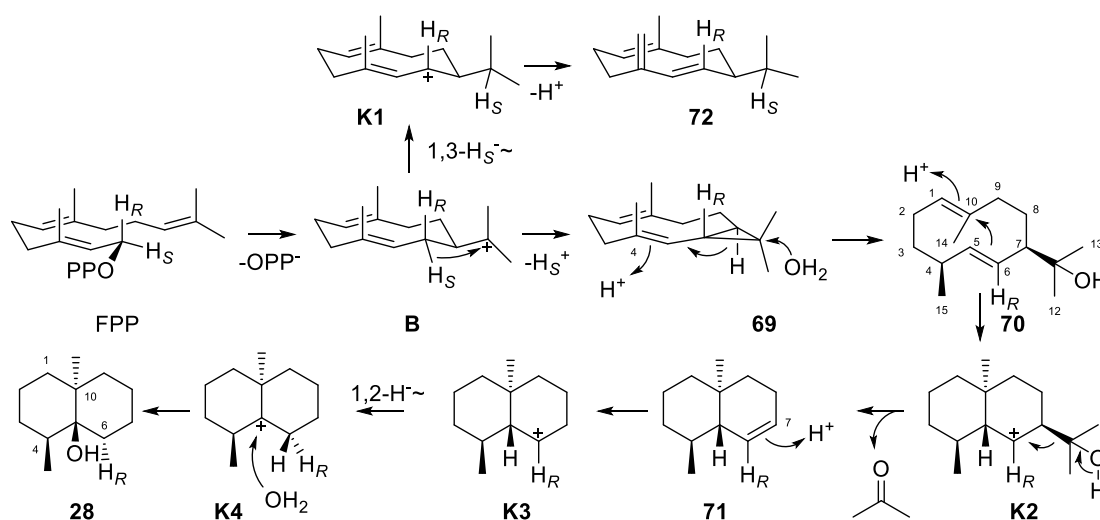
Reprinted from *ChemBioChem* **2023**, *24*, e202300101 with kind permission from John Wiley & Sons

The publication “A Detailed View on Geosmin Biosynthesis” can be found in Appendix N.





Geosmin (**28**, Scheme 44) is a terpene alcohol with a typical earthy smell.<sup>[170]</sup> This compound is widespread in nature since it is a secondary metabolite of many organisms including bacteria, especially actinomycetes,<sup>[171]</sup> fungi<sup>[172]</sup> and plants.<sup>[173][174]</sup> This natural product is also closely linked with our daily life due to its existence in foodstuff such as beetroot,<sup>[175]</sup> buckwheat<sup>[176]</sup> and maize.<sup>[177]</sup> It is worth mentioning that the special flavour of beetroot attributes to geosmin. In addition, the typical smell of a plowed earth or a field after a rain also derives from geosmin.



**Scheme 44.** The biosynthetic pathway towards geosmin (**28**) and germacrene D (**72**).

Geosmin originates from FPP biosynthetically, but it is a bicyclic compound with only 12 carbons. A plausible biosynthetic pathway was proposed based on feeding experiments with deuterated substrates (Scheme 44).<sup>[128]</sup> In this pathway, three neutral intermediates, i.e. isolepidozene (**69**), (1(10)*E*,5*E*)-germacradien-11-ol (**70**) and (8*S*,9*S*,10*S*)-8,10-dimethyl-1-octalin (**71**), were suggested. To be more specific, **69** is formed by the deprotonation of the cationic intermediate **B**. The reprotonation of **69** at C4 followed by the attack of water reaches **70**. Compound **71** is obtained by a retro-Prins reaction of intermediate **K2** and this is the step that causes a loss of three carbons in the form of a molecule of acetone. In order to characterise the biosynthesis of geosmin more comprehensively, especially these reprotonation steps of the neutral intermediates, a geosmin synthase (AaGS) from *Allokutzneria albata* was cloned and expressed in *E.coli* BL21. The enzymatic conversion of FPP by

AaGS led to the production of **28**, **70**, **71** and germacrene D (**72**). A preparative scale enzymatic reaction allowed the isolation of products **28**, **70** and **72**. The full NMR assignments for these compounds were accordingly achieved.

Notably, compound **70** instead of **28** is the major product of the enzymatic conversion, indicating that the formation of **71** from **70** via intermediate **K2** is inefficient. Biosynthetically, compound **72** can be yielded if **B** undergoes a 1,3-hydride shift followed by deprotonation. Labelling experiments were performed in this study to follow the reaction cascade towards **28** from FPP. For the reprotonation steps, the corresponding  $^{13}\text{C}$  labelled FPP was converted by AaGS in  $\text{D}_2\text{O}$  buffer. The successful incorporation of deuterium at a  $^{13}\text{C}$  labelled carbon could be observed from the  $^{13}\text{C}$  NMR by showing an up field-shifted triplet. All the reprotonation steps in the biosynthetic pathway were finally confirmed by this approach. In addition, the labelling experiment was also conducted to follow the 1,2-hydride shift from **K3** to **K4** using (1- $^{13}\text{C}$ ,2- $^2\text{H}$ )FPP. Notably, the proton which undergoes the 1,3-hydride shift from **B** to **K1** is the same one which is lost during the deprotonation from **B** to **69**.<sup>[178][179]</sup>

It has been reported that the geosmin synthase from *Streptomyces coelicolor* contains two domains. The N-terminal domain is responsible for the biosynthesis of **70** from FPP, while the C-terminal domain converts **70** into **28**.<sup>[70]</sup> However, it remains unclear how these two domains collaborate with each other in the reaction cascade. One of the hypotheses could be that there is a channel for conveying **70** from the N-terminal domain to the active site of the C-terminal domain. As an alternative, compound **70** could also be released by the N-terminal domain to the incubation buffer and captured by the C-terminal domain to proceed with the reaction cascade. As a result, the latter one was experimentally verified in this research using  $\beta$ -cyclodextrin, a reagent reported to form host-guest complexes with terpenes.<sup>[180]</sup>

It was also found by Dionigi that the production of geosmin in *Streptomyces albidoflavus* could be increased during cultivation in the presence of highly concentrated copper sulfate.<sup>[181]</sup> However, Schrader and Blevins reported that the addition of divalent zinc, iron or copper to the growth medium inhibits the production in *Streptomyces halstedii*.<sup>[182]</sup> In this study, different metal ions, i.e.  $\text{Mg}^{2+}$ ,  $\text{Mn}^{2+}$ ,  $\text{Ca}^{2+}$ ,  $\text{Zn}^{2+}$ ,  $\text{Co}^{2+}$ ,  $\text{Cu}^{2+}$ ,  $\text{Fe}^{2+}$  and  $\text{Ni}^{2+}$  were used to investigate their

influence on the production of geosmin by AaGS in vitro. The experimental results indicated that  $Mg^{2+}$  plays an essential role for the catalytic activity of geosmin synthase since the absence of  $Mg^{2+}$  resulted in no production of geosmin. In addition,  $Ca^{2+}$ ,  $Zn^{2+}$ ,  $Co^{2+}$ ,  $Cu^{2+}$ ,  $Fe^{2+}$  and  $Ni^{2+}$  hardly affect the generation of geosmin. Interestingly,  $Mn^{2+}$  could also facilitate the production of geosmin, but its efficiency was much lower than that of  $Mg^{2+}$ .

In this study, I isolated geosmin, (1(10)*E*,5*E*)-germacradien-11-ol and germacrene D. Apart from that, I performed all the incubation experiments to track the steps of reprotonation, hydride shift and deprotonation, and to study the effects of  $\beta$ -cyclodextrin and various metal ions on the productivity of AaGS.



## Chapter 16

### Revision of the Cyclisation Mechanism for the Diterpene Spiroviolene and Investigations on its Mass Spectrometric Fragmentation

Houchao Xu<sup>[a]</sup> and Jeroen S. Dickschat<sup>[a],\*</sup>

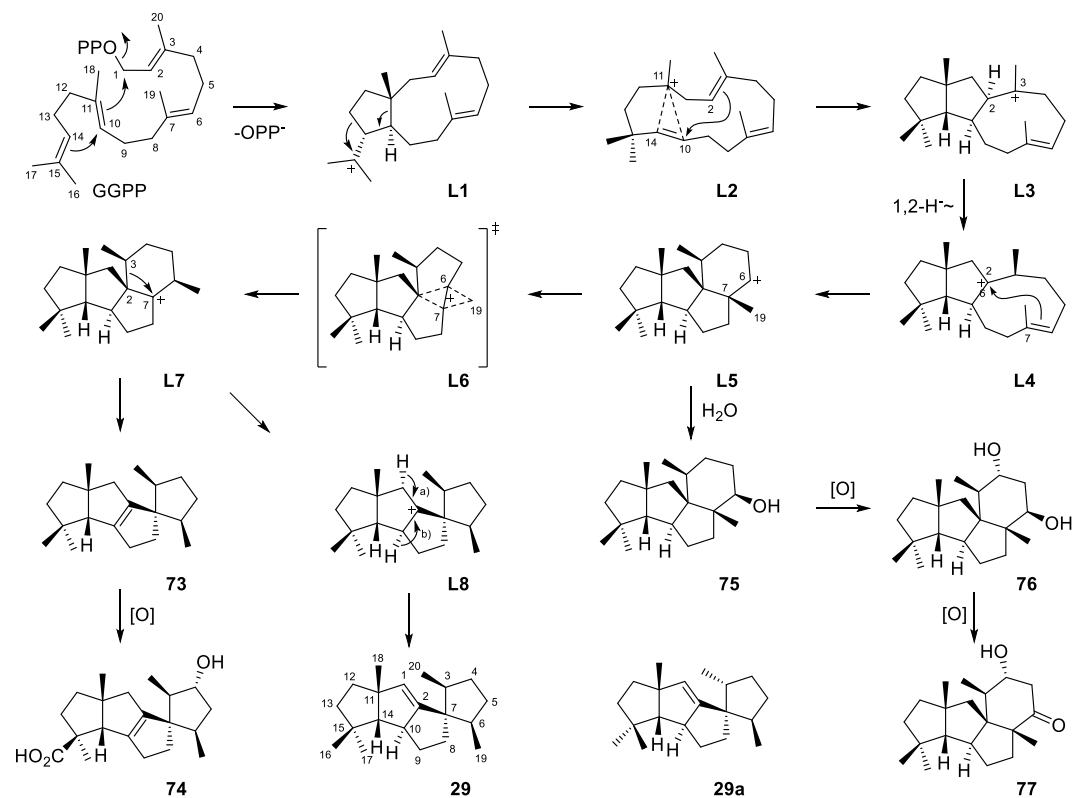
[a] Kekulé-Institute for Organic Chemistry and Biochemistry, University of Bonn, Gerhard-Domagk-Straße 1, 53121 Bonn, Germany, E-mail: [dickschat@uni-bonn.de](mailto:dickschat@uni-bonn.de).

Reprinted from *ChemBioChem* **2021**, 22, 850 with kind permission from John Wiley & Sons

The publication “Revision of the Cyclisation Mechanism for the Diterpene Spiroviolene and Investigations on its Mass Spectrometric Fragmentation” can be found in Appendix O.

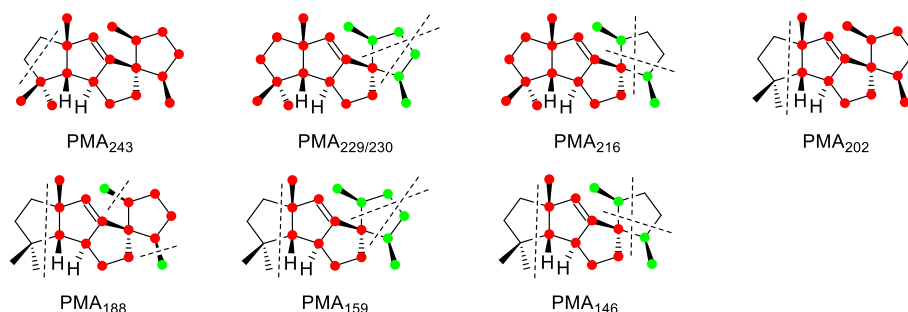


Spiroviolene (**29**, Scheme 45) is a spirocyclic triquinane diterpene which is produced by spiroviolene synthase (SvS) from the actinomycete *Streptomyces violens* with the originally reported structure of **29a** (Scheme 45).<sup>[74]</sup> Notably, the structure of **29a** differs with respect to the configuration of the stereogenic center at C3 from spirograterpene A (**74**) from *Penicillium granulatum* which possesses the same skeleton,<sup>[183]</sup> and from other similar natural products including compound **75** produced by a bifunctional cyclopiane-type diterpene synthase from *Penicillium chrysogenum*<sup>[184]</sup> together with the fungal cyclopiane-type diterpenes conidiogenol (**76**) and conidiogenone (**77**) from *Penicillium cyclopium*.<sup>[185]</sup> Later Snyder and coworkers conducted a total synthesis of **29**,<sup>[75]</sup> indicating that the stereogenic center at C3 of **29** has the same configuration as for the fungal compounds with the consequence that our initially proposed cyclisation mechanism<sup>[74]</sup> cannot be correct. In order to clarify this situation, the absolute configuration and the cyclisation mechanism of **29** were reinvestigated using further labelling experiments in addition to those in the original report.



**Scheme 45.** Revised cyclisation mechanism from GGPP to **29** and biosynthetic links to related fungal compounds.

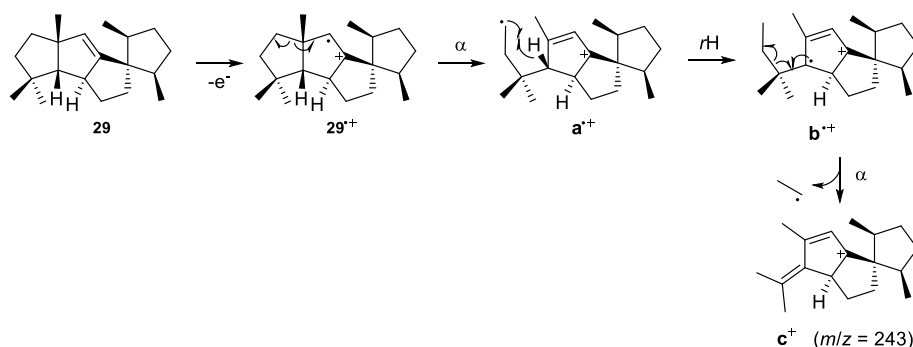
In the original report, substrates (*R*)- and (*S*)-(1-<sup>13</sup>C,1-<sup>2</sup>H)farnesyl pyrophosphate (FPP) in conjunction with (*R*)- and (*S*)-(1-<sup>13</sup>C,1-<sup>2</sup>H)geranyl pyrophosphate (GPP) were used to resolve the stereochemistry of spiroviolene<sup>[74]</sup>, which was proved not to be comprehensive enough. Therefore in this study, stereoselectively deuterated substrates, including (*R*)- and (*S*)-(1-<sup>13</sup>C,1-<sup>2</sup>H)isopentenyl pyrophosphate (IPP)<sup>[89]</sup> and (*E*)- and (*Z*)-(4-<sup>13</sup>C,4-<sup>2</sup>H)IPP<sup>[129]</sup> were used to further determine the absolute configuration of **29**. The HSQC analyses of the obtained products by the enzymatic conversions of these probes indicated that the stereochemistry was in line with what is shown in **29**. Based on the revised structure, a modified biosynthetic pathway (Scheme 45) was developed. Firstly, a 1,11-10,14-cyclisation of GGPP reaches **L1**. Then the expansion of the cyclopentane ring of **L1**, followed by a transformation that was described as “highly asynchronous ring-opening/ring-closing process that accomplishes the same net result as a 1,2-alkyl shift<sup>[186]</sup> from C10 to C14, gives **L2**. **L3** is achieved by this reaction together with a 2,10-cyclisation. The cationic intermediate **L3** then undergoes a 1,2-hydride shift from C2 to C3 to yield **L4**. A 2,7-cyclisation of **L4** leads to the secondary cation **L5** which can be attacked by water to give **75**,<sup>[184]</sup> a compound that may be oxidised via **76** to **77**.<sup>[185]</sup> In addition, **L5** may also proceed with a Me19 shift from C7 to C6 via the non-classical cation **L6** to yield **L7**. The spirocenter of **L8** is formed by ring contraction of **L7**. Then the deprotonations of cation **L8** leads to spiroviolene (**29**) or the hypothetical natural product **73** which is probably the precursor of spirograterpene A (**74**).<sup>[183]</sup>



**Figure 13.** Position-specific mass shift analysis ( $PMA_{m/z}$ ) for main EI fragment ions  $m/z$  of **29**. Red dots indicate carbons that contribute fully, green dots indicate carbons that contribute partially to the formation of a fragment ion. Dotted lines indicate relevant carbon-carbon bond cleavages.



Moreover, the EI-MS fragmentation mechanisms were also investigated by a  $^{13}\text{C}$  isotopic labelling strategy. For this study, twenty isotopomers of ( $^{13}\text{C}$ )-**29** were enzymatically prepared from labelled terpene substrates with SvS. By comparing their mass spectra to the mass spectrum of unlabelled **29**, a position-specific mass shift analysis ( $\text{PMA}_{m/z}$ ) indicated for a studied fragment ion ( $m/z$ ) which carbons contribute to its formation (Figure 13). The plausible mechanisms for the formation of these fragments were also described in detail. Taking  $\text{PMA}_{243}$  as an example, the PMA for  $m/z = 243$  ( $[\text{M}-\text{C}_2\text{H}_5]^+$ ) reveals the specific formation of this fragment ion by cleavage of C12 and C13. The electron impact ionisation of **29** leads to the radical cation **29 $^{+\bullet}$**  which can proceed with an  $\alpha$ -cleavage to **a $^{+\bullet}$**  and hydrogen rearrangement to the conjugated butadienyl cation **b $^{+\bullet}$** . Then **c $^{+\bullet}$**  can be eventually formed after another  $\alpha$ -fragmentation happens to **b $^{+\bullet}$**  (Scheme 46).



**Scheme 46.** Fragmentation mechanisms for fragment ion  $m/z = 243$  of **29**.  $\alpha$ :  $\alpha$ -cleavage,  $r\text{H}$ : hydrogen rearrangement.

In this work, I was responsible for conducting all the labelling experiments to re-determine the absolute configuration, to identify the modified biosynthetic pathway and to study the fragmentation mechanisms of **29**.



## Chapter 17

### Mechanistic Characterisation of the Bacterial Sesterviridene Synthase from *Kitasatospora viridis*

Houchao Xu<sup>[a]</sup>, Gregor Schnakenburg<sup>[b]</sup>, Bernd Goldfuss<sup>[c]</sup>, and Jeroen S. Dickschat<sup>[a],\*</sup>

[a] Kekulé-Institute for Organic Chemistry and Biochemistry, University of Bonn, Gerhard-Domagk-Straße 1, 53121 Bonn, Germany, E-mail: dickschat@uni-bonn.de.

[b] Institute for Inorganic Chemistry, University of Bonn, Gerhard-Domagk-Straße 1, 53121 Bonn, Germany.

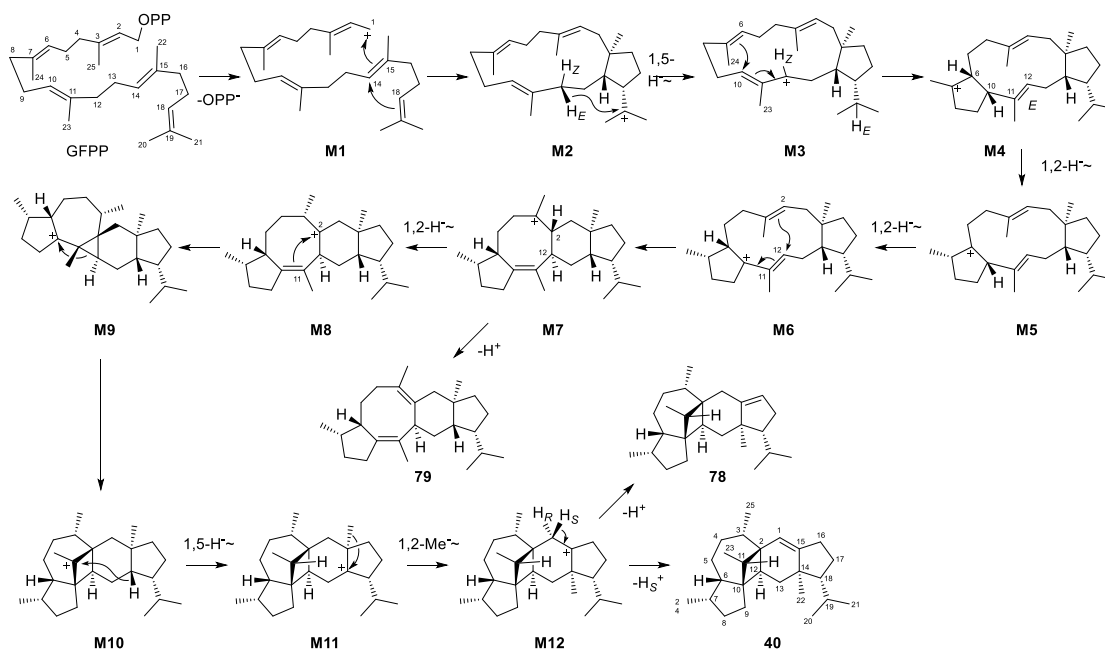
[c] Department for Chemistry, University of Cologne, Greinstraße 4, 50939 Cologne, Germany.

Reprinted from *Angew. Chem. Int. Ed.* **2023**, 62, e202306429 and *Angew. Chem.* **2023**, 135, e202306429 with kind permission from John Wiley & Sons

The publication “Mechanistic Characterisation of the Bacterial Sesterviridene Synthase from *Kitasatospora viridis*” can be found in Appendix P.



Besides several sesquiterpene synthases and one diterpene synthase, a sesterterpene synthase has also been studied for this thesis. The further genome mining of *Kitasatospora viridis* from which kitaviridene synthase was obtained led to the discovery of a gene coding a new sesterterpene synthase. After gene cloning and expression in *E. coli* B 21, this terpene synthase was tested for its function with GPP, FPP, GGPP and GFPP. GCMS analyses of the extracts from enzymatic reactions with different substrates indicated that the enzyme is able to catalyse the conversions of GGPP and GFPP into cembrene A and one major together with two minor unknown sesterterpene hydrocarbons, respectively. The amount of the major product, namely sesterviridene A (**40**) (Scheme 47), was accumulated by repeated large-scale enzymatic reactions that also allowed the isolation of these two minor products designated as sesterviridenes B (**78**) and C (**79**) (Scheme 47). This terpenes synthase represents one of the few bacterial type I sesterterpene synthases. The planar structures of these sesterterpenes were later on characterised through comprehensive 1D and 2D NMR data. Notably, the structures of sesterviridene A and B share one unprecedented carbon skeleton with a 5/7/4/6/5-membered pentacyclic ring system.



**Scheme 47.** The biosynthesis pathway towards sesterviridenes **A–C** (**40**, **78** and **79**).

For sesterviridenes A and B, however, their relative configurations remained doubtful due to the almost identical chemical shifts of H6 and one of the hydrogen at C5. Chemical derivations for **40** were therefore performed aiming to obtain derivatives with these two hydrogens being distinguishable. Although the problem still occurred in most of the synthetic products, the treatment with O<sub>3</sub> gave a product that finally contributed to the settlement of the relative configuration. The determination of the absolute configuration for sesterviridenes A attributed to the successful crystallisation and CuK $\alpha$  X-ray analysis.

In addition, a biosynthetic pathway towards sesterviridenes A–C was also suggested in this study. The proposed pathway starts with the departure of diphosphate in GFPP followed by a 1,14- and 14,18-cyclisation to reach the cationic intermediate **M2**. **M2** then undergoes a 1,5-hydride shift to give **M3** with an allyl cation at C12. The intermediate **M3** then proceeds with a 6,10-cyclisation, two sequential 1,2-hydride shifts and 2,12-cyclisation to afford **M7** that on the one hand can be deprotonated to yield sesterviridene C, on the other hand undergoes a 1,2-hydride transfer and a 2,11-cyclisation to obtain intermediate **M9** that possesses a three-membered ring. Then ring expansion through Wagner–Meerwein rearrangement happens to this intermediate to give **M10** in which the particular four-membered ring is suggested to be formed. Further 1,5-hydride followed by a 1,2-methyl group migration of **M10** gives intermediate **M12** with cation at C15 that can be deprotonated to C1 to produce the major product sesterviridene A, or to C16 for the formation of sesterviridene B.

Isotopic labelling experiments have been carried out to follow the biosynthesis of **40** especially for the skeletal rearrangements, hydride shifts and terminal deprotonation proposed in the route. First of all, 25 experiments with singly <sup>13</sup>C labelled GFPP that were obtained enzymatically have been performed to track down all the carbons. These experiments provided evidence for the Wagner–Meerwein rearrangement step from **M9** to **M10** and the 1,2-methyl group migration step from **M11** to **M12**. For those hydride shifts, different <sup>13</sup>C labelled and deuterated substrates were used. Notably, in order to investigate which hydrogen at C12 undergoes the 1,5-hydride shift, substrates (*E*)- or (*Z*)-(4-

$^{13}\text{C},4\text{-}^2\text{H}$ )IPP in conjunction with  $(7\text{-}^{13}\text{C})\text{GPP}$  were used. For the purpose of following the 1,2-hydride shift from **M7** to **M8**, substrate  $(2\text{-}^{13}\text{C}, 3\text{-}^2\text{H})\text{GFPP}$  was synthesised in this study. Concerning the stereochemical course of the final deprotonation, experimental results based on *(R)*- and *(S)*- $(1\text{-}^{13}\text{C},1\text{-}^2\text{H})\text{IPP}$  indicated that the specific 1-*pro-S* hydrogen is lost. DFT calculations were also applied in this study to prove the rationality of the biosynthetic pathway.

Although the absolute configuration of sesterviridenes **A** is determined based on X-ray analysis which is a widely used technology, wrong determination can still happen. In order to confirm the correctness of the absolute configuration of sesterviridenes A, isotopically labelled probes *(E)*- or *(Z)*- $(4\text{-}^2\text{H},4\text{-}^{13}\text{C})\text{IPP}$  and *(R)*- or *(S)*- $(1\text{-}^2\text{H},1\text{-}^{13}\text{C})\text{IPP}$  were utilised.

Since the direct precursor of these compounds is GFPP, the syntheses of all these  $^{13}\text{C}$  labelled and/or deuterated GFPP substrates will be challenging and time consuming. Therefore the majority of the labelled substrates used in this study were not directly the labelled GFPP substrates, but shorter substrates such as labelled DMAPP, GPP and FPP that can be elongated by geranylgeranyl pyrophosphate synthase (GGPPS) with IPP. In this study, almost all the enzymatic reactions required the involvement of two or three enzymes, which shows an efficient collaboration between enzymes to fulfil the complicated reaction cascade.

In this study, I first isolated these three compounds and elucidated their structures. Then I crystallised sesterviridene A and synthesised  $(2\text{-}^{13}\text{C},3\text{-}^2\text{H})\text{GFPP}$ . I also performed all the labelling experiments to follow the whole biosynthetic pathway and to confirm the absolute configuration of sesterviridene A.





## Chapter 18

### Structural Insights into Three Sesquiterpene Synthases for the Biosynthesis of Tricyclic Sesquiterpenes and Chemical Space Expansion by Structure-Based Mutagenesis

Tingting Lou,<sup>[a],#</sup> Annan Li,<sup>[a],#</sup> Houchao Xu,<sup>[b],#</sup> Jingfeng Pan,<sup>[c],#</sup> Baiying Xing<sup>[a]</sup>, Ruibo Wu,<sup>[c],\*</sup> Jeroen S. Dickschat,<sup>[b],\*</sup> Donghui Yang,<sup>[a],\*</sup> and Ming Ma<sup>[a],\*</sup>

[a] State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University, Beijing 100191, China. E-mail: ydhui@bjmu.edu.cn; mma@bjmu.edu.cn.

[b] Kekulé-Institute for Organic Chemistry and Biochemistry, University of Bonn, 53121 Bonn, Germany. E-mail: dickschat@uni-bonn.de.

[c] School of Pharmaceutical Sciences, Sun Yat-sen University, Guangzhou 510006, China. E-mail: wurb3@mail.sysu.edu.cn

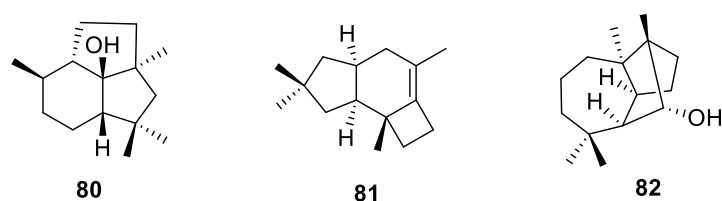
# These authors contributed equally to this work.

Reprinted from *J. Am. Chem. Soc.* **2023**, *145*, 8474 with kind permission from American Chemical Society

The publication “Structural Insights into Three Sesquiterpene Synthases for the Biosynthesis of Tricyclic Sesquiterpenes and Chemical Space Expansion by Structure-Based Mutagenesis” can be found in Appendix Q.



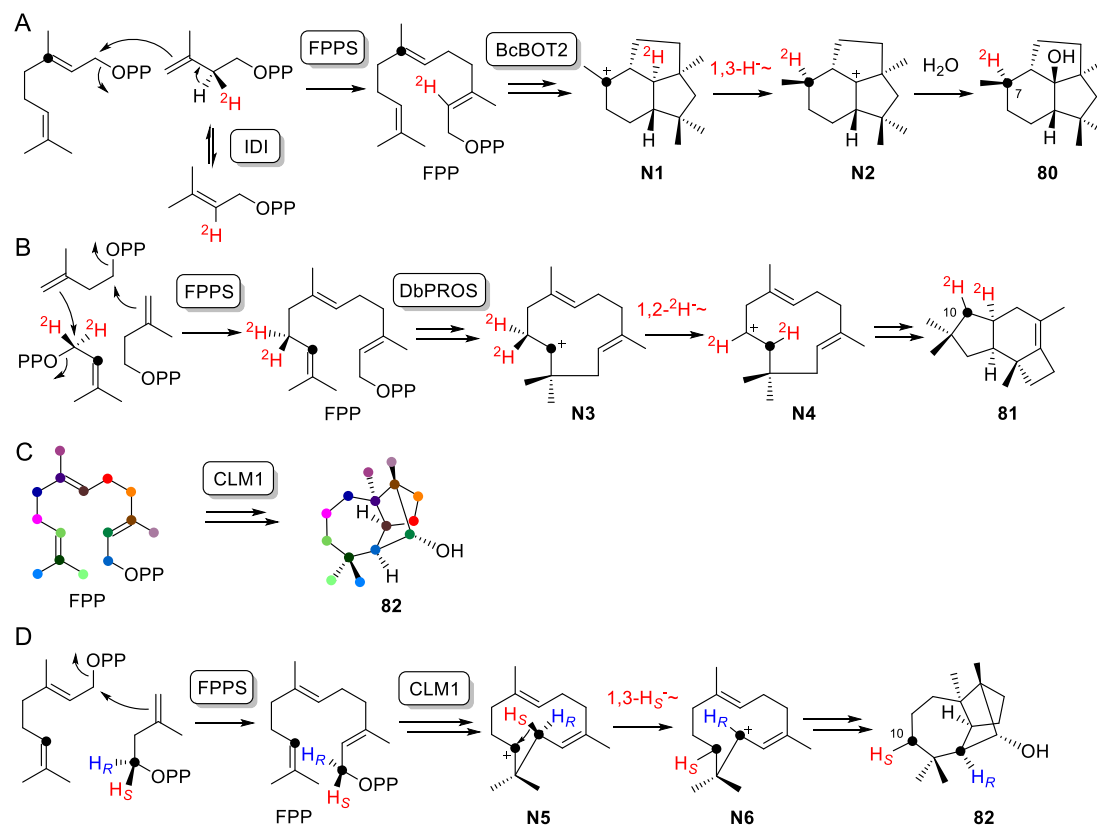
As shown in the previous chapters, sesquiterpene synthases can produce monocyclic compounds such as germacrene A, hedycaryol and  $\alpha$ -humulene. It was also demonstrated that these compounds can be reprotonated to initiate another round of cyclisation events towards bi- and even polycyclic sesquiterpenes. Biosynthetic processes leading to polycyclic compounds hypothetically require a much tighter enzyme control as compared to those cyclisation reactions that result in only monocyclic products. Several crystal structures of type I sesquiterpene synthases have been reported to date, including the selina-4(15),7(11)-diene synthase from *Streptomyces pristinaespiralis*,<sup>[84]</sup> trichodiene synthase from *Fusarium sporotrichioides*,<sup>[187]</sup>  $\delta$ -cadinene synthase from *Gossypium hirsutum*,<sup>[188]</sup> and *epi*-aristolochene synthase from *Nicotiana tabacum*,<sup>[189]</sup> but only a few sesquiterpene synthases for tricyclic compounds have been structurally studied. The representative sesquiterpene synthases are pentalenene synthase from *Streptomyces exfoliatus*,<sup>[56]</sup> *epi*-isozizaene synthase from *Streptomyces coelicolor*,<sup>[190]</sup> cucumene synthase from *Streptomyces clavuligerus*,<sup>[191]</sup>  $\alpha$ -santalene synthase from *Santalum album*,<sup>[192]</sup> and 10-*epi*-cubebol synthase from *Sorangium cellulosum*.<sup>[193]</sup>



**Figure 14.** The three tricyclic sesquiterpenes presilphiperfolan-8 $\beta$ -ol (**80**), protoillud-6-ene (**81**) and longiborneol (**82**) produced by BcBOT, DbPROS and CLM1 respectively.

In the present work, three more sesquiterpene synthases which produce tricyclic compounds were crystallised, including BcBOT2 for presilphiperfolan-8 $\beta$ -ol (**80**, Figure 14) from *Botrytis cinerea*, DbPROS for protoillud-6-ene (**81**, Figure 14) from *Dendrothele bispora*, and CLM1 for longiborneol (**82**, Figure 14) from *Fusarium graminearum*, and their structures were determined through X-ray analysis. These observed crystal structures were all complexed with benzyltriethylammonium cation (BTAC) and Mg<sup>2+</sup>. In addition, BcBOT2 and

CLM1 were also complexed with diphosphate (PPi). Notably, both DbPROS and BcBOT2 structure complexes show a highly similar  $\alpha$ -helical fold, while a different number of  $\alpha$ -helices was observed in CLM1.



**Scheme 48.** The biosynthetic studies on compounds **80–82** produced by BcBOT, DbPROS and CLM1 using isotopic labelling experiments. (A) The 1,3-hydride shift in the biosynthesis of **80** was indicated by a triplet signal for C-7 in the  $^{13}\text{C}$ -NMR using a selective deuteration and  $^{13}\text{C}$ -labelling. (B) The 1,2-hydride shift in the biosynthesis of **81** was confirmed using a same strategy but different labelled substrates. (C) The origin of each carbon from FPP in **83** was followed by single  $^{13}\text{C}$ -labelling experiments. Colored dots represent  $^{13}\text{C}$ -labelled carbons. (D) In the biosynthesis of **83**, the stereochemical course of the 1,3-hydride shift was disclosed using a stereoselective deuteration together with  $^{13}\text{C}$ -labelling approach.

Furthermore, computational studies using quantum mechanics/molecular mechanics based molecular dynamics (QM/MM MD) were conducted to follow the interactions between enzymes and cationic intermediates in the biosynthetic pathways. With these computational modellings, the important

residues in the active sites were identified. Based on the discovery of these key residues, a site-directed mutagenesis study with generation of 37 enzyme variants was also performed to further support the calculation results.

Interestingly, some of the enzyme variants generated by site-directed mutagenesis could convert farnesyl pyrophosphate (FPP) not only into the original sesquiterpene products, but also new ones with different carbon backbones. Plausible cyclisation mechanisms to these newly generated compounds were proposed and then further confirmed by isotopic labelling experiments. By using the specific  $^{13}\text{C}$  labelled and/or deuterated substrates, several hydride shifts, a methyl group migration and the terminal deprotonations to all products were followed unambiguously. In addition, the key steps of cyclisation mechanisms in the biosynthetic pathways of the originally produced compounds **80–82** were also investigated using a stereoselective deuteration together with  $^{13}\text{C}$ -labelling approach. Eventually, the absolute configurations of **80**, **81** and **82** were determined using stereoselectively deuterated substrates, including (*R*)- or (*S*)-(1- $^{13}\text{C}$ ,1- $^2\text{H}$ )IPP and (*E*)- or (*Z*)-(4- $^{13}\text{C}$ ,4- $^2\text{H}$ )IPP in conjunction with HSQC detection.

My contributions to this work include the synthesis of (3- $^{13}\text{C}$ )GPP for one specific labelling experiment in the biosynthesis study of **80**. In addition, I also carried out all the isotopic labelling experiments to investigate the biosynthetic pathways of **80–82** and the compounds generated by the enzyme variants and to establish the absolute configurations of **80–82**.



## Chapter 19

### Crystal Structure Based Mutagenesis of Cattleylene Synthase Leads to the Generation of Rearranged Polycyclic Diterpenes

Baiying Xing,<sup>[a],#</sup> Houchao Xu,<sup>[b],#</sup> Annan Li,<sup>[a]</sup> Tingting Lou,<sup>[a]</sup> Meng Xu,<sup>[a]</sup> Kaibiao Wang,<sup>[a]</sup> Zhengren Xu,<sup>[a]</sup> Jeroen S. Dickschat,<sup>[b],\*</sup> Donghui Yang,<sup>[a],\*</sup> Ming Ma<sup>[a],\*</sup>

[a] State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University, 38 Xueyuan Road, Haidian District, Beijing 100191, China. E-mail: ydhui@bjmu.edu.cn; mma@bjmu.edu.cn.

[b] Kekulé-Institute for Organic Chemistry and Biochemistry, University of Bonn, Gerhard-Domagk-Straße 1, 53121 Bonn, Germany. E-mail: dickschat@uni-bonn.de.

# These authors contributed equally to this work.

Reprinted from *Angew. Chem. Int. Ed.* **2022**, *61*, e202209785 and *Angew. Chem.* **2022**, *134*, e202209785 with kind permission from John Wiley & Sons

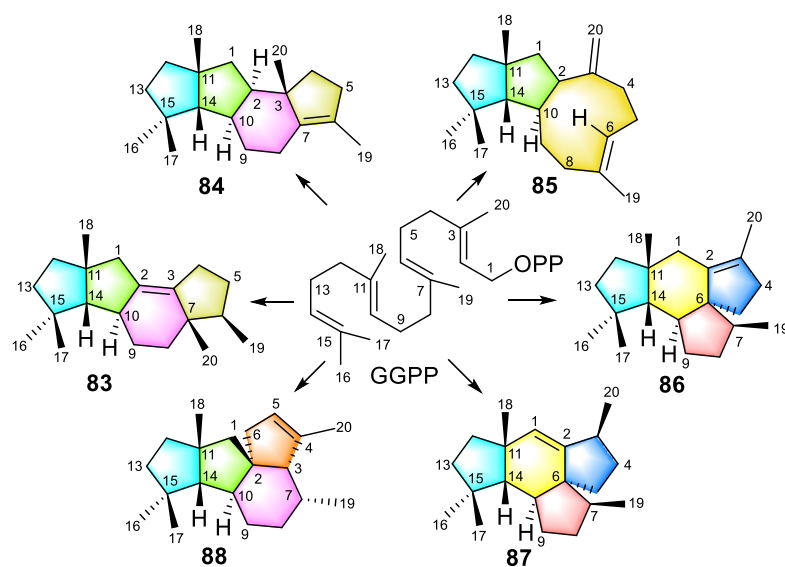
The publication “Crystal Structure Based Mutagenesis of Cattleylene Synthase Leads to the Generation of Rearranged Polycyclic Diterpenes” can be found in Appendix R.





Apart from sesquiterpenes derived from farnesyl diphosphate (FPP) discussed so far in this thesis, diterpenes are another large class in the terpene family which are derived from geranylgeranyl pyrophosphate (GGPP). Structurally, GGPP has one more isoprene unit than FPP. Although GGPP and FPP usually undergo the same cyclisation reactions, e.g. ring closures, hydride or proton shifts, Wagner-Meerwein rearrangements, and deprotonation, due to more reactive sites in the natural diterpene precursor GGPP than in the sesquiterpene precursor FPP, the cyclisation cascades of diterpenes usually can be more complex than those of sesquiterpenes, which leads to more rearranged structures with highly complex ring systems.

Cattleyene (**83**, Figure 15) is a tetracyclic diterpene produced by cattleyene synthase (CyS) from *Streptomyces cattleya*. The skeleton of this compound features an unusual 5/5/6/5 ring system. Its biosynthetic pathway was fully elucidated and its absolute configuration was also determined unambiguously using isotopic labelling experiments.<sup>[194]</sup>

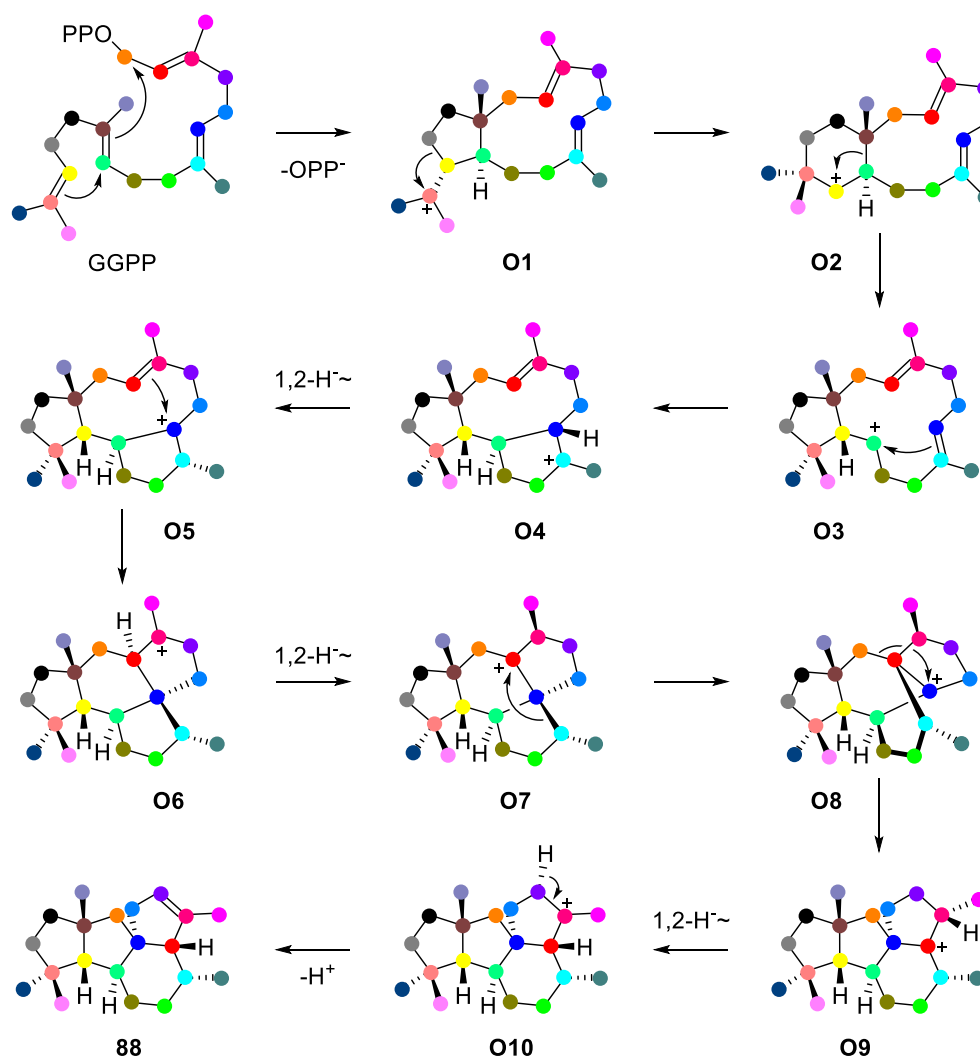


**Figure 15.** The diterpenes **83–88** produced in this study by the diterpene synthase CyS and its enzyme variants. Different rings in these compounds are indicated with different colours.

In order to investigate this synthase more deeply, the structures of apo-Cys, and Cys complexed with GGPP and magnesium ions (CyS-GGPP-Mg<sup>+</sup>) were crystallised and then resolved using X-ray analysis in this study. So far only a

few X-ray structures of class I DTSs have been reported, including taxadiene synthase from *Taxus brevifolia*,<sup>[195]</sup> spiroviolene synthase (SvS) from *Streptomyces violens*,<sup>[196]</sup> cyclooctat-9-en-7-ol synthase (CotB2) from *S. melanosporofaciens*,<sup>[197][198][199]</sup> *ent*-kaurene synthase from *Bradyrhizobium japonicum*,<sup>[200]</sup> and isopimarane synthases Sat1646 from *Salinispora* sp. and Stt4548 from *Streptomyces* sp.<sup>[201]</sup> Notably, Cys-GGPP-Mg<sup>+</sup> represents the first case of a crystal structure that is complexed with the native substrate GGPP instead of a non-reactive GGPP analogue. These crystal structures allowed modellings of biosynthetic intermediates into the active site to investigate how CyS catalyses the conversion of GGPP into the highly complex compound **83**. The key residues shown in these modellings were further studied by site-directed mutagenesis experiments using an engineered *E. coli* strain. In this engineered strain, the isopentenol utilisation pathway (IUP) is established to produce isopentenyl (IPP) and dimethylallyl pyrophosphate (DMAPP). The crystallographic analysis demonstrated the essential roles of active residues F62, W81A, W318 in the biosynthesis of **83**. Furthermore, the enzyme variants of mutants C59A, F86A, W160A, A190G and A191G produced altered product profiles with C59A presenting the most efficient production of compounds **84–88** (Figure 1). All the five compounds possess different carbon skeletons and the skeleton of **88** is novel. In order to understand why the C59A variant (Cys<sup>C59A</sup>) has an increased, but less selective production, the crystal structure of Cys<sup>C59A</sup> was also resolved. It was found that the crystal structure of Cys<sup>C59A</sup> turned out to be highly similar to the apo-CyS and CyS-GGPP-Mg<sup>2+</sup> structures of the wild-type enzyme apart from position F86 moving slightly away from the active site. This movement enlarges the active site cavity which may lead to an improved uptake of GGPP to produce more compounds with different skeletons. Since compound **88** is a new natural product with a novel carbon backbone, its biosynthesis pathway was then fully elucidated using isotopic labelling experiments. Firstly, 20 experiments using all 20 isotopomers of (<sup>13</sup>C)GGPP indicated the skeleton of **88** is highly rearranged biosynthetically. Based on this experimental result, a remarkable biosynthetic pathway was revealed (Scheme 49). In this pathway, three 1,2 hydride shifts and one deprotonation were proposed. The enzymatic reaction using (3-<sup>13</sup>C,2-<sup>2</sup>H)FPP and IPP with GGPP

synthase (GGPPS) and  $\text{CyS}^{\text{C59A}}$  supports the 1,2-hydride shift from intermediate **O4** to **O5**. Interestingly, the 1,2-hydride shifts from **O6** to **O7** and from **O9** to **O10** happen to the same hydrogen. This hypothesis was confirmed using (3- $^{13}\text{C}$ ,2- $^2\text{H}$ )GGPP with  $\text{CyS}^{\text{C59A}}$ . The stereoselectivity of the terminal deprotonation was investigated by conversion of DMAPP and (*E*)- or (*Z*)-(4- $^2\text{H}$ , 4- $^{13}\text{C}$ )IPP with GGPPS and  $\text{CyS}^{\text{C59A}}$ . The results showed that the  $\alpha$ -oriented proton in **O10** is lost.



**Scheme 49.** The cyclization mechanism from GGPP to **88** based on  $^{13}\text{C}$ -labelling experiments. The  $^{13}\text{C}$  labellings are indicated with coloured dots.

In this work, I was responsible for conducting all the isotopic labelling experiments to follow these skeletal rearrangements, 1,2 hydride shifts and the stereochemical course of the deprotonation for the biosynthesis of **88**.



## Chapter 20

### Mechanistic Investigations on Microbial Type I Terpene Synthases through Site-Directed Mutagenesis

Houchao Xu<sup>[a]</sup>, and Jeroen S. Dickschat<sup>[a],\*</sup>

[a] Kekulé-Institute for Organic Chemistry and Biochemistry, University of Bonn, Gerhard-Domagk-Straße 1, 53121 Bonn, Germany. Email: [dickschat@uni-bonn.de](mailto:dickschat@uni-bonn.de).

Reprinted from *Synthesis* **2022**, 54, 1551 with kind permission from Georg Thieme Verlag KG and Copyright Clearance Center

The publication “Mechanistic Investigations on Microbial Type I Terpene Synthases through Site-Directed Mutagenesis” can be found in Appendix S.

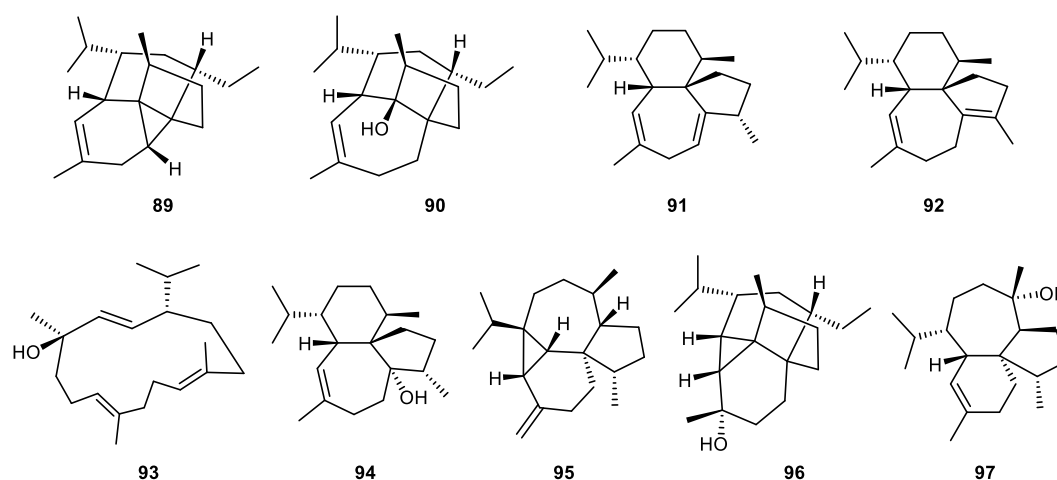


It is well known that terpenoids dominate the kingdom of natural products in nature with around 100,000 compounds having been discovered up to present. The isolation of these compounds was performed originally on macroorganisms, e.g. plants. Later on, microorganisms, including bacteria, were also found to be an indispensable resource. With the development of biological technology, specific enzymes that produce terpenes from macro- or microorganisms can be obtained. This allows the discovery of new terpenes that cannot be extracted from those organisms. Recently, more and more terpene synthases together with their products have been reported. It is remarkable that terpenes, especially the highly cyclised ones, can be produced by a single enzyme from acyclic substrates. The technologies of crystallisation and X-ray analysis allow further investigation on the three dimensional structures of these terpene synthases. These structures provide a direct and clear view in the aspect of amino acids to observe the conserved and other essential residues in the polypeptide chains and to understand how these complex cascade reactions happen in the active sites. The crystal structure of selina-4(15),7(11)-diene synthase (SdS) from *Streptomyces pristinaespiralis* is a typical example and arguably one of the most insightful crystal structures of a type I terpene synthase (Figure 9).<sup>[84]</sup> The determination of its structure made it accessible for the investigation on the deprotonation and reprotonation mechanism in the biosynthesis of selina-4(15),7(11)-diene using QM/MM molecular dynamics (MD) simulations and isotopic labelling experiments.<sup>[202]</sup> Also due to the access of these enzyme structures, site-directed mutagenesis (SDM) aiming to alter the function of terpene synthases, including increasing the production, generating other compounds or accepting other terpene precursors, is utilised more and more frequently.

Although the amino acid sequences of terpene synthases differ strongly from one another, there are still some characteristics in common. First of all, they all show an  $\alpha$ -helical fold. In addition, they all have a set of highly conserved motifs and residues for different functions in their active sites. For instance, the aspartate-rich motif DDXX(X)D serves as a  $Mg^{2+}$  cofactor-binding site.<sup>[203]</sup> The NSE triad ND(L,I,V)XSXX(R,K)E is also responsible for the binding of the  $Mg^{2+}$  cofactor.<sup>[204]</sup> A single Arg located upstream of the NSE triad is the

pyrophosphate sensor that forms hydrogen bridge to the substrate and is thus involved in substrate recognition.<sup>[84]</sup> In addition, the RY pair can recognise the substrate through hydrogen bridge.<sup>[187]</sup> Other important residues have a structural function such as the Arg-Glu salt bridge between helices F and G, and the conserved Trp upstream of the RY pair.

In this review article, the accumulated knowledge of the site-directed mutagenesis study on the highly conserved or other essential residues of microbial type I terpene synthases is presented in detail.



**Figure 16.** Structures of diterpene products obtained from wild-type and mutant polytrichastrene A synthase (CpPS).

It is concluded that the site-directed mutagenesis of the highly conserved residues for substrate binding and catalysis usually results in a dysfunctional protein or significantly reduced terpene production. This further proves that these residues are absolutely necessary for all the functional terpene synthases. The existence of the residues for sustaining the space structures shows the same priority since any site-directed mutagenesis for these residues also leads to an inefficient enzyme. These findings provide an effective way to identify if a newly discovered terpene synthase is functionally normal or not. Once these highly conserved residues are incomplete in their polypeptide chains, it very probably means that the enzyme is inactive but not always, e.g. DDXXD in sestermobaraene synthase (SmTS1) from *Streptomyces mobaraensis* is modified to N<sup>86</sup>DLTV. The NSE triad is changed to N<sup>226</sup>QRYSYFKE. The pyrophosphate sensor is missing in SmTS1 and instead a glycine is observed



in the corresponding position (G180).<sup>[205]</sup>

However, it is found that alteration on the residues in the active-site cavity by site-directed mutagenesis usually gives a promising result in achieving more compounds and a higher yield. One of the interesting examples is CpPS\_I66F, a variant of CpPS. CpPS was reported to be a diterpene synthase from *Cryptosporangium polytrichastri*. It can convert geranylgeranyl diphosphate (GGPP) into multiple polycyclic diterpenes, i.e. polytrichastrene A (**89**), polytrichastrol A (**90**), wanju-2,5-diene(**91**), wanju-2,6-diene (**92**) and thunbergol (**93**), while its variant CpPS\_I66F can not only produce **89**, **91** and **92** much more efficiently compared to the wild type, but also generate several new compounds **94–97** with different complex skeletons (Figure 16).<sup>[206]</sup> Notably, I66 is one of the important hydrophobic residues residing in front of the Asp-rich motif in the active-site cavity. Taken together, it is suggested in this review that the residues in the active-site cavity are sometimes changeable to explore the potential functions of terpene synthase such as increasing the production and broadening the diversity of compounds.

My contributions to this article include analysing the biosynthetic pathways of the compounds and presenting them in the review.



## Chapter 21

### Summary and outlook

The studies for this doctoral dissertation concentrate on the biosynthesis of terpenes with 4 reviews and 15 research articles included. For these reviews, three of them summarise sesquiterpenes that are biosynthetically related. The fourth one presents the accumulated knowledge of mechanistic investigations on microbial type I terpene synthases through site-directed mutagenesis. Among the 15 research articles, 12 publications focus on sesquiterpene biosynthesis and the other three are about the biosynthesis of diterpenes (2 articles) and sesterterpene (1 article).

In the review articles about sesquiterpenes, compounds biosynthetically originating from germacrene A (**14**), hedycaryol (**15**) and germacrene B (**16**) are presented, respectively. They exhibit not only the structural diversity derived from these three neutral intermediates, but also the various reaction cascades proposed for their biosynthesis. It is also interesting to find that so many sesquiterpenes are to a certain degree related in the aspect of their structures or biosynthesis. The cyclisation mechanisms introduced in the reviews also can be utilised for proposing the biosynthesis pathways towards structurally more complex terpenes such as diterpenes, sesterterpenes and triterpenes.

The biosynthetic studies of terpenes require several critical factors. First of all, characterisation of a functional enzyme, i.e. of a terpene synthase is the primary step. The compound isolation and structure elucidation then can be realised. Based on the structure with its determined absolute configuration, a reasonable biosynthetic pathway is hereby proposed. At the end the suggested pathway should be validated experimentally. However, nowadays numerous terpenes are identified but much fewer biosynthetic pathways towards them have been deciphered, indicating the importance of exploring terpene biosynthesis in much more depth and detail. For the 12 research articles about sesquiterpenes, the majority (9 articles) reported the work with compounds that have been known for decades or even more than one century. For instance,  $\alpha$ -humulene (**11**) was first isolated in 1895<sup>[161]</sup>, but the stereochemical course regarding its biosynthesis had never been reported—a challenging problem for this achiral

compound. On the contrary, three biosynthetic pathways towards patchoulol (**22**), a sesquiterpene alcohol discovered in 1869<sup>[207]</sup>, have been proposed in the literature which resulted in a confusing situation.<sup>[61][62][63]</sup> Although a series of studies on geosmin (**28**), a compound isolated in 1965 had provided evidence for a rational biosynthetic pathway towards it, the cyclisation mechanism was not fully confirmed experimentally.

Similar cases also include selina-4(15),7(11)-diene (**18**), and sodorifen (**26**) of which the reported biosynthetic pathways lack sufficient experimental support.<sup>[84][68]</sup> The problems of these products have been resolved, which are all described in this thesis. Apart from those known sesquiterpenes, investigations on unprecedented terpene synthases for new compounds, i.e. isoishwarane (**19**), (1*S*,5*S*,7*R*,10*R*)-guaia-4(15)-en-11-ol (**20**), kitaviridene (**23**) and sesterviridene (**40**) were also performed. Since these are novel natural products, their biosynthetic pathways are suggested for the first time which are also confirmed experimentally.

All in all, no matter if they were known or new compounds, attention was always paid in the study to their biosynthesis, namely the cyclisation mechanism from the corresponding precursors. Usually the cyclisation mechanisms are studied for carbon skeleton rearrangements, hydrogen shifts, deprotonations and reprotonations. These mechanistic steps are followed using an isotopic labelling strategy in conjunction with NMR and GC-MS detections. The investigations on biosynthetic mechanisms require different <sup>13</sup>C labelled and/or deuterated substrates that were obtained through chemical approaches. Analytical-scale enzymatic conversions with labelled substrates by terpene synthases lead to the formation of trace products with specific carbons being <sup>13</sup>C labelled and/or hydrogens being deuterated. By measuring the <sup>13</sup>C NMR and/or HSQC spectra for these minor labelled samples and comparing them with those of unlabelled samples, the cyclisation mechanisms can be tracked. In addition, the isotopically labelled substrates can also be used to determine the absolute configurations for terpenes as introduced in the previous chapters. The labelling strategy has been proved to be a very efficient and convincing method to investigate the biosynthesis of terpenes. The difficulty of this strategy may arise from the syntheses of diverse labelled substrates. High cost of the

labelled starting materials, highly volatile synthetic intermediates, risk of losing deuterium during synthesis and the difficulty of obtaining pure  $^{13}\text{C}$  NMR spectra due to the easily detectable  $^{13}\text{C}$  signals in other minor side products can all make the syntheses challenging. Enzymatic conversions catalysed by terpene synthase are a fast process in active cavities so that no experimental evidence for cationic intermediate e.g. by use of spectroscopic methods can be obtained. The active cavities in terpene synthases act as black boxes to which labelled substrates enter and the corresponding labelled products are subsequently released from. The biosynthesis pathways are proposed according to the introduction, loss or migration of labelled atoms in products compared with substrates. Sometimes more than one biosynthetic route can be suggested based on the labelling experiments.

Another method to study the biosynthesis of terpenes is rooted in DFT calculation. By using this method, every single step via a transient state in the biosynthesis can be calculated for their energy barriers. This energy barrier can reflect how smoothly the reaction cascade proceeds and how rational the whole biosynthesis route is. However, the disadvantage of this technology alone is that sometimes two or more pathways are calculated to be reasonable, or the “true” mechanism may be overlooked and remains uncalculated. Considering the strengths and shortcomings of both methods described above, the solution would be to use both methods together for elucidating the biosynthesis of terpenes. This idea was realised in several studies of this thesis, as exemplified by kitaviridene and sesterviridene.

Apart from isotopic labelling strategies and DFT calculations, the structures of terpene synthases can also provide evidence for the biosynthesis. The access of the structure for selina-4(15),7(11)-diene synthase leads to the discovery of Gly182, a residue that serves with its main chain carbonyl oxygen as a base and acid to fulfil the deprotonation and reprotonation sequence of the intermediate germacrene A in selina-4(15),7(11)-diene biosynthesis. This important new finding may be of relevance also for other terpene biosynthetic pathways involving deprotonation-reprotonation sequences.

The structures of terpene synthase are also helpful in other aspects such as site-directed mutagenesis based on enzyme structures. The mutagenesis

studies on the one hand contribute to the identification of functional residues and motifs, on the other hand provide a possibility to explore the potential enzyme functions by reshaping the active sites. The accumulated knowledge from previous studies are summarised in one of the review articles. Related research was also performed for three sesquiterpene synthases and a diterpene synthase in this thesis. Interestingly, the enzyme variant C59A of cattleyene synthase (CyS) is capable to convert GGPP into five more new compounds with different skeletons. The product with a novel skeleton was studied through isotopic labelling experiments for its biosynthesis. The small difference between crystal structures of the CyS<sup>C59A</sup> and the CyS lies in the movement of F86 leading to a slightly enlarged active site cavity, which explains the formation of multiple products in the C59A enzyme variant.

Structures of terpene synthases allow a deeper comprehension of terpene biosynthesis chemically and biologically. However, up to present, only a limited number of structures of terpene synthases has been resolved. Sometimes homology modelling based on known crystal structures are utilised for the terpene biosynthesis studies. Nevertheless, it would be ideal to obtain the structures for some special cases such as geosmin synthase that contains two domains for the biosynthesis and sodorifen synthase that catalyses an incredible reaction cascade including fragmentation and [4+3] cycloaddition.

After investigation of several terpenes for their biosynthetic pathways, some conclusions and outlooks are hereby summarised. First of all, discovery of new terpenes especially the ones with unprecedented carbon skeletons enriches the diversity of the terpene family. Their unique biosynthetic pathways demonstrate what reactions a single terpene synthase is able to catalyse. In addition, we also paid attention to the biosynthesis of the known compounds e.g.  $\alpha$ -humulene, patchoulol and geosmin that can be easily overlooked. Although they were reported long time ago, there are still some interesting aspects to study or confusing points to sort out. These investigations enable us to understand these compounds more comprehensively.

All the terpenes that were investigated for this thesis are either terpene hydrocarbons or alcohols directly generated by terpene synthase. However, often oxidised terpenoids are produced by terpene synthases in conjunction

with cytochromes P450 that come into the focus of our biosynthesis study. In some cases next to a terpene synthase gene resides a gene coding for a P450, exemplified by the isoishwarane synthase. In such cases, the terpene hydrocarbons and alcohol are very probably not the final products for this gene cluster but serve as intermediates to other oxidised terpenoids. It would be interesting to follow the whole post modification process and test the potential bioactivities of the final products.

As discussed above, DFT calculations are a widely used approach to study terpene biosynthesis. Usually the calculation is conducted only based on the substrate or intermediate structure itself without considering the influence of the enzyme structure. But actually the structure of the active cavity very likely affects a lot on the enzymatic conversion including which substrate the enzyme can accept, how the intermediates are shaped for the next reaction and where the reaction cascade ends. Therefore, the computation of a biosynthetic route inside the active site is able to more precisely imitate the reaction cascade in the enzyme. The typical example is the biosynthesis study on selina-4(15),7(11)-diene using QM(DFT)/MM MD simulations. This method is capable to simulate and record every single reaction step happening in the active cavity for terpene biosynthesis. This method ideally requires a crystal structure of the terpene synthase. However, nowadays AlphaFold can predict an enzyme structure from its amino acid sequence. Since this method can give deep insights into enzyme catalysis, it may become the trend for the future to investigate the biosynthetic pathways proposed based on isotopically labelling experiments.

Compared to polyketides that are formed by the collaboration of different domains, terpenes are usually produced by one single enzyme. Thus the difficulty level of intervening a terpene synthase by altering certain amino acids to obtain relevant intermediates or to extend to reaction cascade is much higher than that of polyketide synthase (PKS). For the modular PKS, a deep understanding of the enzymology is available that allows the prediction of the structure of a product from the gene sequence. But with more and more enzyme structures being available and artificial intelligence (AI) technology more frequently applied in protein research, such prediction may come true for

terpene synthases. In that case, the biosynthesis study on a terpene will become more and more fascinating. With the technology development of enzyme engineering, de novo synthesis of a terpene gene to give a super-efficient enzyme that can produce a designed target compound for pharmaceutical use may also be realised.



## References

- [1] D. A. Dias, S. Urban, U. Roessner, *Metabolites* **2012**, 2, 303.
- [2] F. Wöhler, *Ann. Phys.* **1828**, 88, 253.
- [3] E. Kinne-Saffran, R. K. H. Kinne, *Am. J. Nephrol.* **1999**, 19, 290.
- [4] A. Fleming, *Br. J. Exp. Pathol.* **1929**, 10, 226.
- [5] J. M. Gulland, R. Robinson, *J. Chem. Soc., Trans.* **1923**, 123, 980.
- [6] R. N. Moore, G. Bigam, J. K. Chan, A. M. Hogg, T. T. Nakashima, J. C. Vederas, *J. Am. Chem. Soc.* **1985**, 107, 3694.
- [7] M. W. Saville, J. Lietzau, J. M. Pluda, W. H. Wilson, R. W. Humphrey, E. Feigel, S. M. Steinberg, S. Broder, R. Yarchoan, J. Odom, I. Feuerstein, *Lancet* **1995**, 346, 26.
- [8] M. C. Wani, H. L. Taylor, M. E. Wall, P. Coggon, A. T. McPhail, *J. Am. Chem. Soc.* **1971**, 93, 2325.
- [9] J. Wang, C. Xu, Y. K. Wong, Y. Li, F. Liao, T. Jiang, Y. Tu, *Engineering* **2019**, 5, 32.
- [10] G. Bergstroem, *Proc. Phytochem. Soc. Eur.* **1991**, 31, 287.
- [11] M. Moriya, *J. Chem. Soc., Trans.* **1881**, 39, 77.
- [12] R. Eccles, *J. Pharm. Pharmacol.* **1994**, 46, 618.
- [13] H. Erdtman, Y. Hirose, *Acta Chem. Scand.* **1962**, 16, 1311.
- [14] M. Furusawa, T. Hashimoto, Y. Noma, Y. Asakawa, *Chem. Pharm. Bull.* **2005**, 53, 1513.
- [15] A. Ortega, J. F. Blount, P. S. Manchand, *J. Chem. Soc., Perkin Trans. 1* **1982**, 10, 2505.
- [16] E. R. Butelman, M. J. Kreek, *Front. Pharmacol.* **2015**, 6, 190.
- [17] K. A. MacLean, M. W. Johnson, C. J. Reissig, T. E. Prisinzano, R. R. Griffiths, *Psychopharmacology* **2013**, 226, 381.
- [18] M. Nakamura, K. Ishibashi, *Nippon Nogei Kagaku Kaishi* **1958**, 32, 739.
- [19] R. Morrison, C. Gardiner, A. Evidente, R. Kiss, H. Townley, *Pharm. Res.* **2014**, 31, 2904.
- [20] Z. Zerbst, *Chemisches Zentralblatt* **1833**, 4, 411.
- [21] N. Kočevár Glavač, S. Kreft, *Food Chem.* **2012**, 131, 305.
- [22] O. Wallach O, *Eur. J. Org. Chem.* **1887**, 238, 78.
- [23] J. S. Dickschat, *Nat. Prod. Rep.* **2011**, 28, 1917.
- [23a] J. W. Cornforth, R. H. Cornforth, G. Popjak, L. Yengoyan, *J. Biol. Chem.* **1966**, 241, 3970.
- [24] M. Rohmer, *Nat Prod Rep.* **1999**, 16, 565.
- [25] E. Beytia, A. A. Qureshi, J. W. Porter, *J. Biol. Chem.* **1973**, 248, 1856.
- [26] H. Tao, L. Lauterbach, G. Bian, R. Chen, A. Hou, T. Mori, S. Cheng, B. Hu, L. Lu, X. Mu, M. Li, N. Adachi, M. Kawasaki, T. Moriya, T. Senda, X. Wang, Z. Deng, I. Abe, J. S. Dickschat, T. Liu, *Nature* **2022**, 606, 414.
- [27] P.-H. Liang, T.-P. Ko, A. H.-J. Wang, *Eur. J. Biochem.* **2002**, 269, 3339.
- [28] M. N. Ashby, P. A. Edwards, *J. Biol. Chem.* **1990**, 265, 13157.
- [29] N. Shimizu, T. Koyama, K. Ogura, *J. Bacteriol.* **1998**, 180, 1578.

- [30] K. -I. Asai, S. Fujisaki, Y. Nishimura, T. Nishino, K. Okada, T. Nakagawa, M. Kawamukai, H. Matsuda, *Biochem. Biophys. Res. Commun.* **1994**, 202, 340.
- [31] K. Okada, K. Suzuki, Y. Kamiya, X. Zhu, S. Fujisaki, Y. Nishimura, T. Nishino, T. Nakagawa, M. Kawamukai, H. Matsuda, *Biochim. Biophys. Acta* **1996**, 1302, 217.
- [32] K. Okada, M. Minehira, X. Zhu, K. Suzuki, T. Nakagawa, H. Matsuda, M. Kawamukai, *J. Bacteriol.* **1997**, 179, 3058.
- [33] S. -I. Ohnuma, T. Koyama, K. Ogura, *J. Biochem. (Tokyo)* **1992**, 112, 743.
- [34] S. -I. Ohnuma, T. Koyama, K. Ogura, *J. Biol. Chem.* **1991**, 266, 23706.
- [35] H. Teclebrhan, J. Olsson, E. Swiezewska, G. Dallner, *J. Biol. Chem.* **1993**, 268, 23081.
- [36] K. Suzuki, K. Okada, Y. Kamiya, X. F. Zhu, T. Nakagawa, M. Kawamukai, H. Matsuda, *J. Biochem. (Tokyo)* **1997**, 121, 496.
- [37] M. V. Keenan, C. M. J. Allen, *Arch. Biochem. Biophys.* **1974**, 161, 375.
- [38] C. M. Allen, *Meth. Enzymol.* **1985**, 110, 281.
- [39] S. Nishikawa, A. Nakano, *Proc. Natl. Acad. Sci. USA* **1993**, 90, 8179.
- [40] M. Sato, K. Sato, S. -I. Nishikawa, A. Hirata, J. -I Kato, A. Nakano. *Mol. Cell. Biol.* **1999**, 19, 471.
- [41] S. K. Oh, K. H. Han, S. B. Ryu, H. Kang, *J. Biol. Chem.* **2000**, 275, 18482.
- [42] M. C. Schulbach, P. J. Brennan, D. C. Crick, *J. Biol. Chem.* **2000**, 275, 22876.
- [43] J. S. Dickschat, *Nat. Prod. Rep.* **2016**, 33, 87.
- [44] D. W. Christianson, *Chem. Rev.* **2017**, 117, 11570.
- [45] X. Pan, W. Du, X. Zhang, X. Lin, F.-R. Li, Q. Yang, H. Wang, J. D. Rudolf, B. Zhang, L.-B. Dong, *J. Am. Chem. Soc.* **2022**, 144, 22067.
- [46] H. Seto, H. Yonehara, *J. Antibiot.* **1980**, 33, 92.
- [47] G. D. Annis, L. A. Paquette, *J. Am. Chem. Soc.* **1982**, 104, 4504.
- [48] G. Mehta, K. S. Rao, *J. Chem. Soc., Chem. Commun.* **1985**, 21, 1464
- [49] D. H. Hua, *J. Am. Chem. Soc.* **1986**, 108, 3835.
- [50] M. K. Pallerla, J. M. Fox, *Org. Lett.* **2007**, 9, 5625.
- [51] D. E. Cane, A. M. Tillman, *J. Am. Chem. Soc.* **1983**, 105, 122.
- [52] D. E. Cane, C. Abell, A. M. Tillman, *Bioorg. Chem.* **1984**, 12, 312.
- [53] D. E. Cane, C. Abell, R. Lattman, C. T. Kane, B. R. Hubbard, P. H. M. Harrison, *J. Am. Chem. Soc.* **1988**, 110, 4081.
- [54] P. H. M. Harrison, J. S. Oliver, D. E. Cane, *J. Am. Chem. Soc.* **1988**, 110, 5922.
- [55] M. Seemann, G. Zhai, K. Umezawa, D. E. Cane, *J. Am. Chem. Soc.* **1999**, 121, 591.
- [56] C. A. Lesburg, G. Zhai, D. E. Cane, D. W. Christianson, *Science* **1997**, 277, 1820.
- [56a] C. M. Starks, K. Back, J. Chappell, J. P. Noel, *Science* **1997**, 277, 1815.

- [57] M. Seemann, G. Zhai, J.-W. de Kraker, C. M. Paschall, D. W. Christianson, D. E. Cane, *J. Am. Chem. Soc.* **2002**, *124*, 7681.
- [58] P. Gutta, D. J. Tantillo, *J. Am. Chem. Soc.* **2006**, *128*, 6172.
- [59] L. Zu, M. Xu, M. W. Lodewyk, D. E. Cane, R. J. Peters, D. J. Tantillo, *J. Am. Chem. Soc.* **2012**, *134*, 11369.
- [60] J. A. Faraldos, S. Wu, J. Chappell, R. M. Coates, *Tetrahedron* **2007**, *63*, 7733.
- [61] R. Croteau, S. L. Munck, C. C. Akoh, H. J. Fisk, D. M. Satterwhite, *Arch. Biochem. Biophys.* **1987**, *256*, 56.
- [62] A. Akhila, P. K. Sharma, R. S. Thakur, *Phytochemistry* **1988**, *27*, 2105.
- [63] J. A. Faraldos, S. Wu, J. Chappell, R. M. Coates, *J. Am. Chem. Soc.* **2010**, *132*, 2998.
- [64] A.-K. Aschenbrenner, M. Kwon, J. Conrad, D.-K. Ro, O. Spring, *Phytochemistry* **2016**, *124*, 29.
- [65] P. Rabe, J. Rinkel, T. A. Klapschinski, L. Barra, J. S. Dickschat, *Org. Biomol. Chem.* **2016**, *14*, 158.
- [66] L. Lauterbach, A. Hou, J. S. Dickschat, *Chem. Eur. J.* **2021**, *27*, 7923.
- [67] S. H. von Reuss, M. Kai, B. Piechulla, W. Francke, *Angew. Chem. Int. Ed.* **2010**, *49*, 2009.
- [68] S. H. von Reuss, D. Domik, M. C. Lemfack, N. Magnus, M. Kai, T. Weise, B. Piechulla, *J. Am. Chem. Soc.* **2018**, *140*, 11855.
- [69] N. Magnus, S. H. von Reuss, F. Braack, C. Zhang, K. Baer, A. Koch, P. L. Hampe, S. Sutour, F. Chen, B. Piechulla, *Angew. Chem. Int. Ed.* **2023**, e202303692.
- [70] J. Jiang, X. He, D. E. Cane, *Nat. Chem. Biol.* **2007**, *3*, 711.
- [71] T. Dairi, Y. Hamano, T. Kuzuyama, N. Itoh, K. Furihata, H. Seto, *J. Bacteriol.* **2001**, *183*, 6085.
- [72] Y. Hamano, T. Kuzuyama, N. Itoh, K. Furihata, H. Seto, T. Dairi, *J. Biol. Chem.* **2002**, *277*, 37098.
- [73] J. S. Dickschat, *Angew. Chem. Int. Ed.* **2019**, *58*, 15964.
- [74] P. Rabe, J. Rinkel, E. Dolja, T. Schmitz, B. Nubbemeyer, T. H. Luu, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2017**, *56*, 2776.
- [75] H. M. Chi, C. J. F. Cole, P. Hu, C. A. Taylor, S. A. Snyder, *Chem. Sci.* **2020**, *11*, 10939.
- [76] J. D. Rudolf, L.-B. Dong, K. Manoogian, B. Shen, *J. Am. Chem. Soc.* **2016**, *138*, 16711.
- [77] R. Chiba, A. Minami, K. Gomi, H. Oikawa, *Org. Lett.* **2013**, *15*, 594.
- [78] J. Shao, Q.-W. Chen, H.-J. Lv, J. He, Z.-F. Liu, Y.-N. Lu, H.-L. Liu, G.-D. Wang, Y. Wang, *Org. Lett.* **2017**, *19*, 1816.
- [79] Y. Yang, Y. Zhang, S. Zhang, Q. Chen, K. Ma, L. Bao, Y. Tao, W. Yin, G. Wang, H. Liu, *J. Nat. Prod.* **2018**, *81*, 1089.
- [80] A. Hou, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2020**, *59*, 19961.

- [81] B. Gu, B. Goldfuss, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2023**, *62*, e202215688.
- [82] S.-H. Kim, W. Lu, M. K. Ahmadi, D. Montiel, M. A. Ternei, S. F. Brady, *ACS Synth. Biol.* **2019**, *8*, 109.
- [83] D. W. Christianson, *Chem. Rev.* **2018**, *118*, 11795.
- [83a] L. C. Tarshis, M. Yan, C. D. Poulter, J. C. Sacchettini, *Biochemistry* **1994**, *33*, 10871.
- [84] P. Baer, P. Rabe, K. Fischer, C. A. Citron, T. A. Klapschinski, M. Groll, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2014**, *53*, 7652.
- [85] J. Rinkel, L. Lauterbach, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2017**, *56*, 16385.
- [86] A. J. Weinheimer, W. W. Youngblood, P. H. Washecheck, T. K. B. Karns, L. S. Ciereszko, *Tetrahedron Lett.* **1970**, *26*, 497.
- [87] V. Sykora, V. Herout, F. Sorm, *Coll. Czech. Chem. Commun.* **1956**, *21*, 267.
- [87a] A. D. Wagh, S. K. Paknikar, S. C. Bhattacharyya, *Tetrahedron* **1964**, *20*, 2647.
- [88] J. A. Faraldos, Y. Zhao, P. E. O'Maille, J. P. Noel, R. M. Coates, *ChemBioChem* **2007**, *8*, 1826.
- [89] J. Rinkel, J. S. Dickschat, *Org. Lett.* **2019**, *21*, 2426.
- [90] R. O. Hellyer, *Aust. J. Chem.* **1962**, *15*, 157.
- [91] I. A. Southwell, *Phytochemistry* **1970**, *9*, 2243.
- [92] I. A. Southwell, *Aust. J. Chem.* **1978**, *31*, 2527.
- [93] S. Hasegawa, Y. Hirose, *Phytochemistry* **1981**, *20*, 508.
- [94] E. Stahl, *Planta Med.* **1984**, *50*, 157.
- [95] S. Nagahama, M. Tazaki, H. Kobayashi, M. Sumimoto, *Phytochemistry* **1993**, *33*, 879.
- [96] T. Hieda, M. Tazaki, Y. Morishita, T. Aoki, S. Nagahama, *Phytochemistry* **1996**, *42*, 159.
- [97] H. Xu, N. Lackus, T. G. Köllner, J. S. Dickschat, *Org. Lett.* **2022**, *24*, 587.
- [98] S. Irmsch, Y. Jiang, F. Chen, J. Gershenzon, T. G. Köllner, *BMC Plant Biol.* **2014**, *14*, 270.
- [99] J. Hattan, K. Shindo, T. Ito, Y. Shibuya, A. Watanabe, C. Tagaki, F. Ohno, T. Sasaki, J. Ishii, A. Kondo, N. Misawa, *Planta* **2016**, *243*, 959.
- [100] L. Chuang, C.-S. Wen, Y.-R. Lee, Y.-L. Lin, L.-R. Hsu, S.-Y. Wang, F.-H. Chu, *J. Nat. Prod.* **2018**, *81*, 1162.
- [101] Z. Li, Y. Jiang, X. Zhang, Y. Chang, S. Li, X. Zhang, S. Zheng, C. Geng, P. Men, L. Ma, Y. Yang, Z. Gao, Y.-J. Tang, S. Li, *ACS Catal.* **2020**, *10*, 5846.
- [102] R. D. Hartley, C. H. Fawcett, *Phytochemistry* **1969**, *8*, 1793.
- [103] K. Nishimura, N. Shinoda, Y. Hirose, *Tetrahedron Lett.* **1969**, 3097.
- [104] E. D. Brown, M. D. Solomon, J. K. Sutherland, A. Torre, *Chem. Comm.* **1967**, 111.

- [105] G. Rücker, G. A. A. Brasil e Silva, L. Bauer, M. Schikarski, *Planta Med.* **1977**, *31*, 322.
- [106] B. C. Clark, T. S. Chamblee, G. A. Lacobucci, *J. Agric. Food Chem.* **1987**, *35*, 514.
- [107] A. M. Adio, C. Paul, H. Tesso, P. Kloth, W. A. König, *Tetrahedron: Asymmetr.* **2004**, *15*, 1631.
- [108] P. M. Bleeker, E. A. Spyropoulou, P. J. Diergaarde, H. Volpin, M. T. J. De Both, P. Zerbe, J. Bohlmann, V. Falara, Y. Matsuba, E. Pichersky, M. A. Haring, R. C. Schuurink, *Plant Mol. Biol.* **2011**, *77*, 323.
- [109] J. J. Zager, I. Lange, N. Srividya, A. Smith, B. M. Lange, *Plant Physiol.* **2019**, *180*, 1877.
- [110] S. M. Colby, J. Crock, B. Dowdle-Rizzo, P. G. Lemaux, R. Croteau, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 2216.
- [111] S. Picaud, M. E. Olsson, M. Brodelius, P. E. Brodelius, *Arch. Biochem. Biophys.* **2006**, *452*, 17.
- [112] W. K. W. Chou, I. Fanizza, T. Uchiyama, M. Komatsu, H. Ikeda, D. E. Cane, *J. Am. Chem. Soc.* **2010**, *132*, 8850.
- [113] B. W. Pyle, H. T. Tran, B. Pickel, T. M. Haslam, Z. Gao, G. MacNevin, J. C. Vederas, S. -U. Kim, D. -K. Ro, *FEBS J.* **2012**, *279*, 3136.
- [114] A. M. Adio, C. Paul, H. Tesso, P. Kloth, W. A. König, *Tetrahedron: Asymmetr.* **2004**, *15*, 1631.
- [115] K. Takeda, *Tetrahedron* **1974**, *30*, 1525.
- [116] R. V. H. Jones, M. D. Sutherland, *Chem. Commun.* **1968**, 1229.
- [117] M. Kodama, S. Yokoo, Y. Matsuki, S. Ito, *Tetrahedron Lett.* **1979**, *20*, 1687.
- [118] P. Baer, P. Rabe, C. A. Citron, C. C. de Oliveira Mann, N. Kaufmann, M. Groll, J. S. Dickschat, *ChemBioChem* **2014**, *15*, 213.
- [119] N. D. Lackus, J. Morawetz, H. Xu, J. Gershenzon, J. S. Dickschat, T. G. Köllner, *Molecules* **2021**, *26*, 555.
- [120] L. Ding, H. Goerls, K. Dornblut, W. Lin, A. Maier, H. -H. Fiebig, C. Hertweck, *J. Nat. Prod.* **2015**, *78*, 2963.
- [121] P. Rabe, T. Schmitz, J. S. Dickschat, *Beilstein J. Org. Chem.* **2016**, *12*, 1839.
- [122] S. Schulz, J. Fuhlendorff, H. Reichenbach, *Tetrahedron* **2004**, *60*, 3863.
- [123] J. S. Dickschat, S. C. Wenzel, H. B. Bode, R. Müller, S. Schulz, *ChemBioChem* **2004**, *5*, 778.
- [124] J. S. Dickschat, T. Martens, T. Brinkhoff, M. Simon, S. Schulz, *Chem. Biodivers.* **2005**, *2*, 837.
- [125] P. Rabe, C. A. Citron, J. S. Dickschat, *ChemBioChem*, **2013**, *14*, 2345.
- [126] C. Höckelmann, P. G. Becher, S. H. von Reuss, F. Jüttner, *Z. Naturforsch.* **2009**, *64C*, 49.
- [127] L. Barra, P. Barac, G. M. König, M. Crüsemann, J. S. Dickschat, *Org. Biomol. Chem.* **2017**, *15*, 7411.

- [128] J. S. Dickschat, H. B. Bode, T. Mahmud, R. Müller, S. Schulz, *J. Org. Chem.* **2005**, *70*, 5174.
- [129] L. Lauterbach, J. Rinkel, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2018**, *57*, 8280.
- [130] C. O. Schmidt, H. J. Bouwmeester, J. de Kraker, W. A. König, *Angew. Chem. Int. Ed.* **1998**, *37*, 1400.
- [131] D. J. Tantillo, *J. Nat. Prod. Rep.* **2011**, *28*, 1035.
- [132] Y. J. Hong, D. J. Tantillo, *J. Am. Chem. Soc.* **2015**, *137*, 4134.
- [133] H. Xu, J. S. Dickschat, *Beilstein J. Org. Chem.* **2023**, *19*, 186.
- [134] R. B. Kelly, J. Zamecnik, *J. Chem. Soc., Chem. Commun.* **1970**, 1102.
- [135] R. B. Kelly, J. Zamecnik, B. A. Becket, *Can. J. Chem.* **1972**, *50*, 3455.
- [136] H. Fuhrer, A. K. Ganguly, K. W. Gopinath, T. R. Govindachari, K. Nagarajan, B. R. Pai, P. C. Parthasarathy, *Tetrahedron* **1970**, *26*, 2371.
- [137] U. S. K. Rao, B. L. Manjunath, K. N. Menon, *J. Indian Chem. Soc.* **1935**, *12*, 494
- [138] H. Xu, B. Goldfuss, G. Schnakenburg, J. S. Dickschat, *Beilstein J. Org. Chem.* **2022**, *18*, 13.
- [139] T. R. Govindachari, P. A. Mohamed, P. C. Parthasarathy, *Tetrahedron* **1970**, *26*, 615.
- [140] S. A. Achmad, E. L. Ghisalberti, E. H. Hakim, L. Makmur, M. Manurung, *Phytochemistry* **1992**, *31*, 2153.
- [141] S. K. Paknikar, O. Motl, K. K. Chakravarti, *Tetrahedron Lett.* **1977**, *18*, 2121.
- [142] M. M. Sonwa, W. A. König, *Phytochemistry* **2001**, *58*, 799.
- [143] D. J. Mason, A. Dietz, C. De Boer, *Antimicrob. Agents Chemother.* **1963**, *1962*, 554.
- [144] G. A. Boeckler, J. Gershenzon, S. B. Unsicker, *Phytochemistry* **2011**, *72*, 1497.
- [145] E. A. Schmelz, F. Kaplan, A. Huffaker, N. J. Dafoe, M. M. Vaughan, X. Ni, J. R. Rocca, H. T. Alborn, P. E. Teal, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 5455.
- [146] A. Huffaker, F. Kaplan, M. M. Vaughan, N. J. Dafoe, X. Ni, J. R. Rocca, H. T. Alborn, P. E. Teal, E. A. Schmelz, *Plant Physiol.* **2011**, *156*, 2082.
- [147] H. Danner, G. A. Boeckler, S. Irmisch, J. S. Yuan, F. Chen, J. Gershenzon, S. B. Unsicker, T. G. Köllner, *Phytochemistry* **2011**, *72*, 897.
- [148] S. Irmisch, A. T. Müller, L. Schmidt, J. Günther, J. Gershenzon, T. G. Köllner, *BMC Plant Biol.* **2015**, *15*, 262.
- [149] N. D. Lackus, S. Lackner, J. Gershenzon, S. B. Unsicker, T. G. Köllner, *Sci. Rep.* **2018**, *8*, 17936.
- [150] N. D. Lackus, N. P. Petersen, R. Nagel, A. Schmidt, S. Irmisch, J. Gershenzon, T. G. Köllner, *Molecules* **2019**, *24*, 2408.
- [151] H. Xu, J. S. Dickschat, *Chem. Eur. J.* **2022**, *28*, e202200405.

- [152] J. Rinkel, P. Rabe, X. Chen, T. G. Köllner, F. Chen, J. S. Dickschat, *Chem. Eur. J.* **2017**, *23*, 10501.
- [153] F. M. Hahn, A. P. Hurlburt, C. D. Poulter, *J. Bacteriol.* **1999**, *181*, 4499.
- [154] P. Rabe, J. Rinkel, B. Nubbemeyer, T. G. Köllner, F. Chen, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2016**, *55*, 15420.
- [155] F. L. Lin, L. Lauterbach, J. Zhou, Y. H. Wang, J. M. Lv, G. D. Chen, D. Hu, H. Gao, X. S. Yao, J. S. Dickschat, *ACS Catal.* **2020**, *10*, 4306.
- [156] W. Treibs, *Justus Liebigs Ann. Chem.* **1949**, *564*, 141.
- [157] M. Dobler, J. D. Dunitz, B. Gubler, H. P. Weber, G. Büchi, O. J. Padilla, *Proc. Chem. Soc., London* **1963**, 383.
- [158] H. Xu, B. Goldfuss, J. S. Dickschat, *Chem. Eur. J.* **2021**, *27*, 9758.
- [159] K. Ekramzadeh, C. Brämer, T. Frister, J. Fohrer, A. Kirschning, T. Scheper, S. Beutel, *Biotechnol. Prog.* **2020**, *36*, e2935.
- [160] H. Xu, J. Rinkel, X. Chen, T. G. Köllner, F. Chen, J. S. Dickschat, *Org. Biomol. Chem.* **2021**, *19*, 370.
- [161] A. C. Chapman, *J. Chem. Soc., Trans.* **1895**, *67*, 54.
- [162] M. D. Sutherland, O. J. Waters, *Aust. J. Chem.* **1961**, *14*, 596.
- [163] F. Yu, S. Okamoto, K. Nakasone, K. Adachi, S. Matsuda, H. Harada, N. Misawa, R. Utsumi, *Planta* **2008**, *227*, 1291.
- [164] Y. Kumeta, M. Ito, *J. Nat. Med.* **2016**, *70*, 452.
- [165] S. Schor, C. Schotte, D. Wibberg, J. Kalinowski, R. J. Cox, *Nat. Commun.* **2018**, *9*, 1963.
- [166] K. Murai, L. Lauterbach, K. Teramoto, Z. Quan, L. Barra, T. Yamamoto, K. Nonaka, K. Shiomi, M. Nishiyama, T. Kuzuyama, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2019**, *58*, 15046.
- [167] Z. Liu, C. Rodriguez, L. Wang, Q. Cui, Y. Huang, E. T. Quintana, M. Goodfellow, *Int. J. Syst. Evol. Microbiol.* **2005**, *55*, 707.
- [168] S. von Reuss, D. Domik, M. C. Lemfack, N. Magnus, M. Kai, T. Weise, B. Piechulla, *J. Am. Chem. Soc.* **2018**, *140*, 37.
- [169] M. C. Lemfack, W. Brandt, K. Krueger, A. Gurowietz, J. Djifack, J.-P. Jung, M. Hopf, H. Noack, B. Junker, S. von Reuss, B. Piechulla, *Sci. Rep.* **2021**, *11*, 3182.
- [170] R. G. Buttery, J. A. Garibaldi, *J. Agric. Food Chem.* **1976**, *24*, 1246.
- [171] N. N. Gerber, *Biotechnol. Bioeng.* **1967**, *9*, 321.
- [172] T. Kikuchi, S. Kadota, H. Suehara, A. Nishi, K. Tsubaki, *Chem. Pharm. Bull.* **1981**, *29*, 1782.
- [173] J. Spoerle, H. Becker, N. S. Allen, M. P. Gupta, *Z. Naturforsch.* **1991**, *46C*, 183.
- [174] Y. Saritas, M. M. Sonwa, H. Iznaguen, W. A. König, H. Muhle, R. Mues, *Phytochemistry* **2001**, *57*, 443.
- [175] K. E. Murray, P. A. Bannister, R. G. Buttery, *Chem. Ind.* **1975**, 973.
- [176] J. Shi, G. Tong, Q. Yang, M. Huang, H. Ye, Y. Liu, J. Wu, J. Zhang, X. Sun, D. Zhao, *J. Agric. Food Chem.* **2021**, *69*, 11361.

- [177] R. A. Flath, R. R. Forrey, J. O. John, B. G. Chan, *J. Agric. Food Chem.* **1978**, *26*, 1290.
- [178] X. He, D. E. Cane, *J. Am. Chem. Soc.* **2004**, *126*, 2678.
- [179] J. Jiang, X. He, D. E. Cane, *J. Am. Chem. Soc.* **2006**, *128*, 8128.
- [180] N. Ajisaka, K. Hara, K. Mikuni, K. Hara, H. Hashimoto, *Biosci. Biotechnol. Biochem.* **2000**, *64*, 731.
- [181] C. P. Dionigi, *Water Sci. Technol.* **1995**, *31*, 135.
- [182] K. K. Schrader, W. T. Blevins, *J. Ind. Microbiol. Biotechnol.* **2001**, *26*, 241.
- [183] S. Niu, Z.-W. Fan, C.-L. Xie, Q. Liu, Z.-H. Luo, G. Liu, X.-W. Yang, *J. Nat. Prod.* **2017**, *80*, 2174.
- [184] T. Mitsuhashi, T. Kikuchi, S. Hoshino, M. Ozeki, T. Awakawa, S.-P. Shi, M. Fujita, I. Abe, *Org. Lett.* **2018**, *20*, 5606.
- [185] T. Roncal, S. Cordobés, U. Ugalde, Y. He, O. Sterner, *Tetrahedron Lett.* **2002**, *43*, 6799.
- [186] Y. J. Hong, D. J. Tantillo, *Aust. J. Chem.* **2017**, *70*, 362.
- [187] M. J. Rynkiewicz, D. E. Cane, D. W. Christianson, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 13543.
- [188] H. A. Gennadios, V. Gonzalez, L. Di Costanzo, A. Li, F. Yu, D. J. Miller, R. K. Allemann, D. W. Christianson, *Biochemistry* **2009**, *48*, 6175.
- [189] C. M. Starks, K. Back, J. Chappell, J. P. Noel, *Science*, **1997**, *277*, 1815.
- [190] R. Li, W. K. Chou, J. A. Himmelberger, K. M. Litwin, G. G. Harris, D. E. Cane, D. W. Christianson, *Biochemistry* **2014**, *53*, 1155.
- [191] P. N. Blank, T. A. Pemberton, J. Y. Chow, C. D. Poulter, D. W. Christianson, *Biochemistry* **2018**, *57*, 6326.
- [192] X. Han, W. D. Liu, C. C. Chen, R. T. Guo, *unpublished* (data obtained from RCSB Protein Data Bank).
- [193] J. N. Whitehead, N. G. H. Leferink, G. K. Reddy, C. W. Levy, S. Hay, E. Takano, N. S. Scrutton. *ACS Catal.* **2022**, *12*, 12123.
- [194] J. Rinkel, S. T. Steiner, J. S. Dickschat, *Angew. Chem. Int. Ed. Engl.* **2019**, *58*, 9230.
- [195] M. Köksal, Y. Jin, R. M. Coates, R. Croteau, D. W. Christianson, *Nature* **2011**, *469*, 116.
- [196] K. Schriever, P. Saenz-Mendez, R. S. Rudraraju, N. M. Hendrikse, E. P. Hudson, A. Biundo, R. Schnell, P. O. Syrén, *J. Am. Chem. Soc.* **2021**, *143*, 3794.
- [197] R. Janke, C. Görner, M. Hirte, T. Brück, B. Loll, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2014**, *70*, 1528.
- [198] T. Tomita, S.-Y. Kim, K. Teramoto, A. Meguro, T. Ozaki, A. Yoshida, Y. Motoyoshi, N. Mori, K. Ishigami, H. Watanabe, M. Nishiyama, T. Kuzuyama, *ACS Chem. Biol.* **2017**, *12*, 1621.
- [199] W. Liu, X. Feng, Y. Zheng, C. H. Huang, C. Nakano, T. Hoshino, S. Bogue, T. P. Ko, C. C. Chen, Y. Cui, J. Li, I. Wang, S. T. Hsu, E. Oldfield, R. T. Guo, *Sci. Rep.* **2014**, *4*, 6214.



- [200] B. Xing, J. Yu, C. Chi, X. Ma, Q. Xu, A. Li, Y. Ge, Z. Wang, T. Liu, H. Jia, F. Yin, J. Guo, L. Huang, D. Yang, M. Ma, *Commun. Chem.* **2021**, *4*, 140.
- [201] K. Raz, R. Driller, N. Dimos, M. Ringel, T. Brück, B. Loll, D. T. Major, *J. Am. Chem. Soc.* **2020**, *142*, 21562.
- [202] Y.-H. Wang, H. Xu, J. Zou, X.-B. Chen, Y.-Q. Zhuang, W.-L. Liu, E. Celik, G.-D. Chen, D. Hu, H. Gao, R. Wu, P.-H. Sun, J. S. Dickschat, *Nat. Catal.* **2022**, *5*, 128.
- [203] K. U. Wendt, G. E. Schulz, *Structure* **1998**, *6*, 127.
- [204] D. E. Cane, I. Kang, *Arch. Biochem. Biophys.* **2000**, *376*, 354.
- [205] A. Hou, J. S. Dickschat, *Beilstein J. Org. Chem.* **2021**, *17*, 2441.
- [206] A. Hou, B. Goldfuss, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2021**, *60*, 20781.
- [207] H. Gal, *Justus Liebigs Ann. Chem.* **1869**, *150*, 374.
- [208] R. G. Buttery, R. E. Lundin, L. Ling, *Chem. Ind.* **1966**, *28*, 1225.



## Appendices A–S

## Appendix A

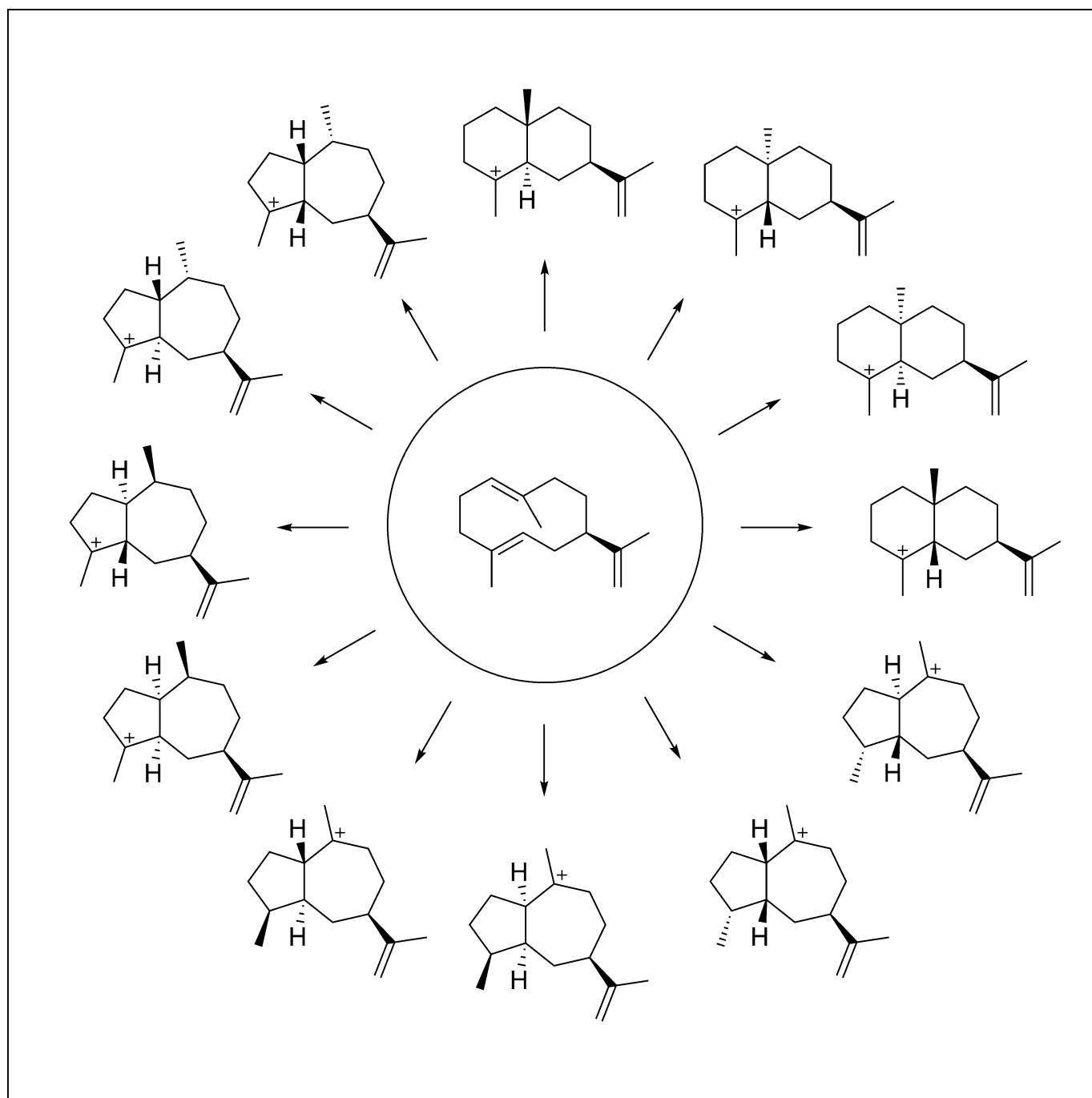
### **Germacrene A – A Central Intermediate in Sesquiterpene Biosynthesis**

*Chem. Eur. J.* **2020**, *26*, 17318

DOI: 10.1002/chem.202002163

## ■ Natural Products

## Germacrene A—A Central Intermediate in Sesquiterpene Biosynthesis

Houchao Xu and Jeroen S. Dickschat<sup>\*[a]</sup>

**Abstract:** This review summarises known sesquiterpenes whose biosyntheses proceed through the intermediate germacrene A. First, the occurrence and biosynthesis of germacrene A in Nature and its peculiar chemistry will be high-

lighted, followed by a discussion of 6–6 and 5–7 bicyclic compounds and their more complex derivatives. For each compound the absolute configuration, if it is known, and the reasoning for its assignment is presented.

## 1. Introduction

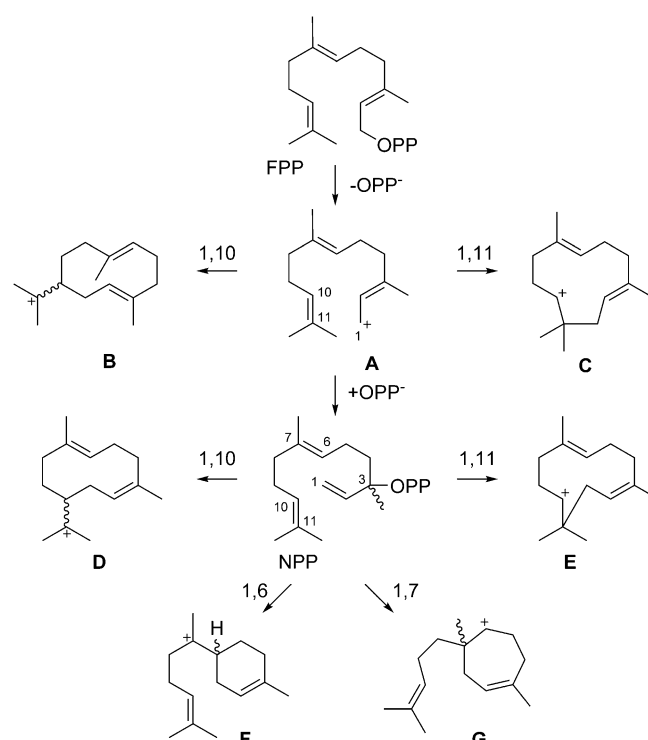
With an estimated number of over 80,000 compounds terpenes form the largest class of natural products. They are produced by all kingdoms of life and can be classified as mono- ( $C_{10}$ ), sesqui- ( $C_{15}$ ) or diterpenes ( $C_{20}$ ) etc. according to the number of incorporated isoprenoid units. During the past decades many sesquiterpene synthases have been reported<sup>[1–6]</sup> that catalyse the cyclisation of farnesyl diphosphate (FPP) through diphosphate abstraction to give the reactive farnesyl cation (**A**, Scheme 1). Attack of the  $C_{10}=C_{11}$  double bond to  $C_1$  can yield the (*E,E*)-germacradienyl cation (**B**) by 1,10- or the (*E,E*)-humulyl cation (**C**) by 1,11-cyclisation. The alternative reaction by reattack of diphosphate to  $C_3$  results in nerolidyl diphosphate (NPP). After a conformational rearrangement of the vinyl group by rotation around the  $C_2$ – $C_3$  bond, cyclisation reactions may proceed to the (*E,Z*)-germacradienyl cation (**D**), the (*E,Z*)-humulyl cation (**E**), the bisabobyl cation (**F**), or to cation **G**, with possible formation of either enantiomer for chiral intermediates. Deprotonation of **B** leads to germacrene A, a widespread natural product and central intermediate in the biosynthesis of many 1,10-cyclised sesquiterpenes. This review discusses its occurrence in Nature, its chemistry, and central importance as an intermediate towards many sesquiterpenes.

## 2. Germacrene A

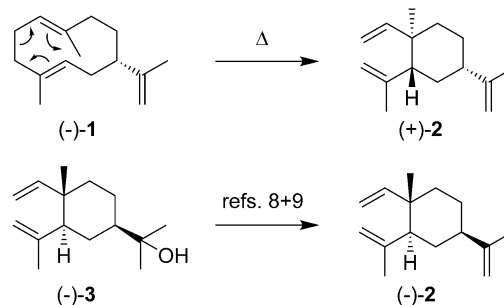
### 2.1. Occurrence in Nature

(–)-Germacrene A (**1**, Scheme 2) was first isolated in 1970 from the gorgonian *Eunicea mammosa*.<sup>[7]</sup> Its absolute configuration was established as (*S*)-(–)-**1** through its Cope rearrangement to (+)- $\beta$ -elemene (**2**) for which the configurational assignment was performed by chemical correlation of (–)-elemol (**3**) to (–)-**2**.<sup>[8,9]</sup> Compound (–)-**1** is also believed to occur in the soft coral *Lobophytum*,<sup>[10]</sup> and is the alarm pheromone of the aphid *Terioaphis maculata*.<sup>[11,12]</sup> In the course of this work it was no-

ticed that the optical rotation ( $[\alpha]_D^{25} = -26.8$ ,  $c$  1.0,  $CCl_4$ ) was significantly higher than initially reported ( $[\alpha]_D^{25} = -3.2$ ,  $c$  14.4,  $CCl_4$ ),<sup>[7]</sup> which is explainable by a partial rearrangement of purified (–)-**1** to (+)-**2**, or alternatively, **1** isolated from *E. mammosa* was not enantiomerically pure. However, the optical rotation of (+)-**2** ( $[\alpha]_D^{25} = +15.1$ , neat) reported in this initial study<sup>[7]</sup> matches the reported value for (–)-**2** ( $[\alpha]_D^{25} = -15.8$ ,  $c$  0.50,  $CHCl_3$ ) obtained by Cope rearrangement of (+)-**1**,<sup>[13]</sup> thus disfa-



**Scheme 1.** Terpene cyclisation modes for FPP.



**Scheme 2.** Structure of **1** and its absolute configuration by chemical correlation.

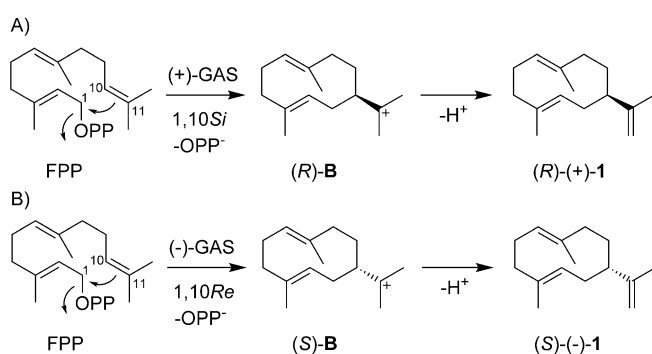
[a] H. Xu, Prof. Dr. J. S. Dickschat  
 Kekulé-Institute for Organic Chemistry and Biochemistry  
 University of Bonn  
 Gerhard-Domagk-Straße 1, 53121 Bonn (Germany)  
 E-mail: dickschat@uni-bonn.de

The ORCID identification number(s) for the author(s) of this article can be found under:  
<https://doi.org/10.1002/chem.202002163>.

© 2020 The Authors. Published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

vouring the latter hypothesis. In fact, the enantiomeric composition of a compound cannot be concluded only from the optical rotation upon its first isolation, or not with certainty if a compound is known to be unstable. Methods such as chromatographic separation on a chiral stationary phase may be more conclusive. Through this approach, König and co-workers found that **1** from various plants is a mixture of enantiomers, ranging from nearly pure (+)-**1** in *Piper nigrum* to mainly (–)-**1** in the liverwort *Barbilophozia barbata*.<sup>[14]</sup>

Germacrene A synthase (GAS) catalyses a 1,10-cyclisation of FPP to **B**, followed by deprotonation to **1** (Scheme 3). Both enantiomers of **1** are accessible through this reaction, depending on whether C10 of FPP is attacked from the *Re* or the *Si* face. Since this face selectivity may be altered by subtle conformational changes of FPP in the active sites of GASs, predictions based on amino acid sequences or phylogenetic analyses regarding the stereochemical implications may be difficult. Many plant GAS have been identified during the past two decades, including two (+)-GASs from *Cichorium intybus*<sup>[15,16]</sup> and one from *Matricaria recutita*,<sup>[17]</sup> with the absolute configuration of (+)-**1** established by chiral GC. Sometimes the absolute configuration can be rationally suggested, because **1** is transformed in the same organism into another compound such as (+)-costunolide.<sup>[18–21]</sup> Further GASs are known from many other plant species,<sup>[22–32]</sup> but the absolute configuration of **1** has frequently not been determined. While the accumulated literature shows that (+)-**1** is typical for plants, the recently characterised bacterial GAS from *Micromonospora marina* produces (–)-**1**,<sup>[33]</sup> reflecting the observation that terpenes and cationic intermediates towards them from plants and bacteria often represent different enantiomers.<sup>[34–37]</sup> The coinciding absolute configuration of (–)-**1** from *E. mammosa* may point to a biosynthesis by symbiotic bacteria in the gorgonian.<sup>[38]</sup>



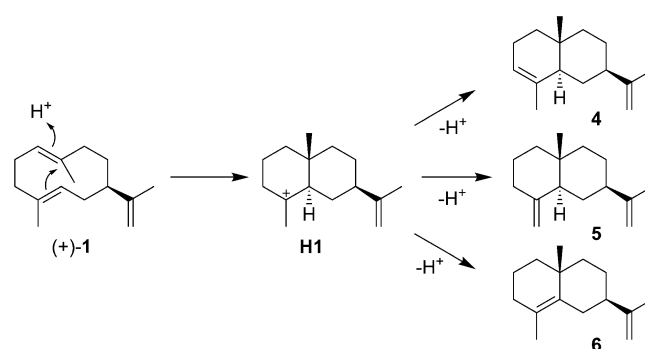
Scheme 3. Cyclisation mechanism from FPP to A) (R)-(+)-**1** and B) (S)-(–)-**1**.

## 2.2. Chemistry of germacrene A

The isolation and full structural and NMR-spectroscopic characterisation of **1** was a long-standing problem significantly hampered by its high reactivity. Its first isolation from *E. mammosa* in 1970 was done by extraction and concentration at temperatures below 35 °C to avoid the Cope rearrangement to **2** (Scheme 2).<sup>[7]</sup> Chromatographic purification on slightly acidic

silica gel induces a cyclisation through cation **H1** to  $\alpha$ -selinene (**4**),  $\beta$ -selinene (**5**), and selina-4,11-diene (**6**, Scheme 4).<sup>[7,11,15]</sup>

The skeleton of **1** is characterised by a conformationally flexible 10-membered ring that shows sufficient ring strain to prevent a fast interconversion between conformers, resulting in broadened signals and multiple signal sets in the NMR spectra. Partial <sup>1</sup>H- and <sup>13</sup>C-NMR data were first published for **1** from *T. maculata*.<sup>[12]</sup> Later studies improved the NMR data assignments for the main conformers of **1** (recorded at 25 °C), but did not allow for a completion of the data sets.<sup>[13,39]</sup> Through NOESY the conformers of **1a** (UU, Me14 and Me15 up), **1b** (UD, up-down) and **1c** (DU, down-up) in a 5:3:2 ratio were identified (Scheme 5A).<sup>[13]</sup> The NMR data sets (25 °C) for all three conformers were recently completed using a <sup>13</sup>C-labelling strategy by conversion of all 15 isotopomers of (<sup>13</sup>C)FPP<sup>[40]</sup> with GAS from *M. marina* into (–)-**1**, resulting in strongly enhanced <sup>13</sup>C-

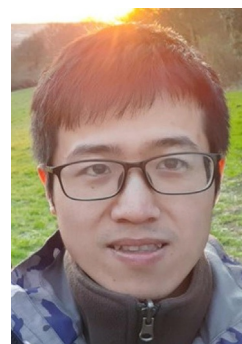


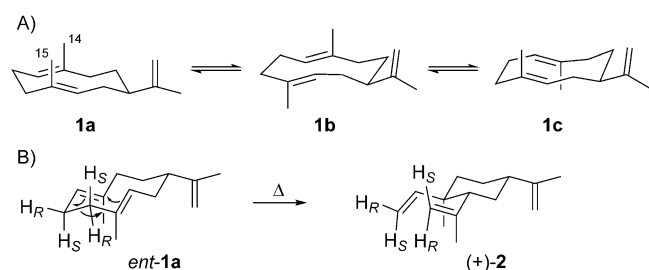
Scheme 4. Acid catalysed conversion of **1** into selinenes.

Jeroen S. Dickschat studied Chemistry at TU Braunschweig and obtained his PhD in 2004. Since 2014 he is a Professor of Organic Chemistry and Biochemistry at the University of Bonn and holds a honorary Professorship at the NIOO Wageningen (The Netherlands). He also serves as an editor for the Beilstein Journal of Organic Chemistry. His research interests include the identification, biosynthesis and synthesis of microbial natural products, with a special focus on enzyme mechanisms.



Houchao Xu obtained his B.Sc. degree in Medicinal Chemistry from the Kunming Institute of Botany, Chinese Academy of Sciences. He joined the group of Prof. Dickschat at the University of Bonn in September 2019 to pursue his PhD degree. His research focuses on the biosynthesis of bacterial terpenes and polyketides.



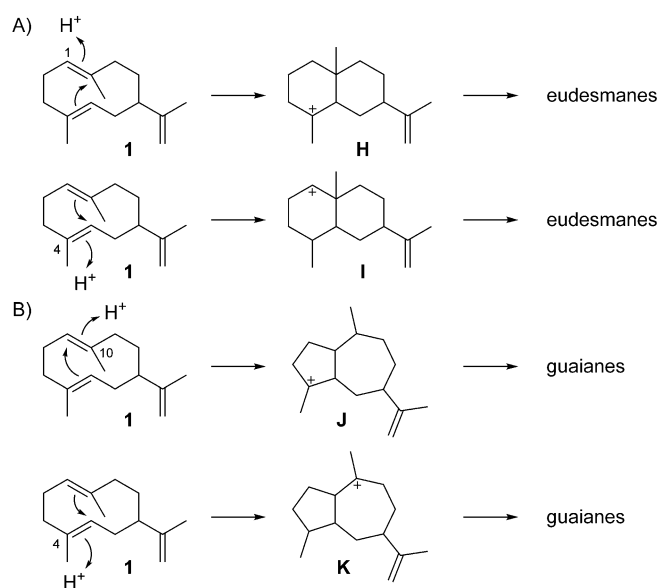


**Scheme 5.** A) Conformers of (+)-1. B) Cope rearrangement of *ent*-1a.

NMR signals for the labelled carbons. HSQC spectroscopy of enzymatically prepared stereoselectively deuterated and  $^{13}\text{C}$ -labelled **1** allowed the NMR assignment of all hydrogens.<sup>[33,41]</sup>

The stereoselectively deuterated and  $^{13}\text{C}$ -labelled isotopomers of **1** were also used to study the stereochemical course of its Cope rearrangement (Scheme 5B). According to the Woodward–Hoffmann rules, pericyclic reactions follow a stereochemical course determined by the symmetry of frontier orbitals.<sup>[42]</sup> For the Diels–Alder reaction this has been verified by stereoselective deuteration,<sup>[43,44]</sup> while classical experiments for the Cope rearrangement have been performed with *meso*- and *rac*-3,4-dimethylhexa-1,5-diene.<sup>[45]</sup> The enzymatic access to labelled **1** allowed to follow the rearrangement to (+)-2 that proceeds from *ent*-1a through a chair-chair transition state.<sup>[33]</sup>

For many terpene synthase reactions **1** is further cyclised in a second step initiated by reprotonation. This can occur at C1 and lead to the 6–6 bicyclic system of **H** as a precursor of eudesmane sesquiterpenes (Scheme 6A). The 6–6 bicyclic system could in theory also arise by protonation at C4 leading to the secondary cation **I**, but this reaction is not preferred. Furthermore, **1** can be protonated at C10 with cyclisation to the 5–7 bicyclic skeleton of **J**, or at C4 resulting in **K**, representing the precursors to guaiane sesquiterpenes. As an alternative to the



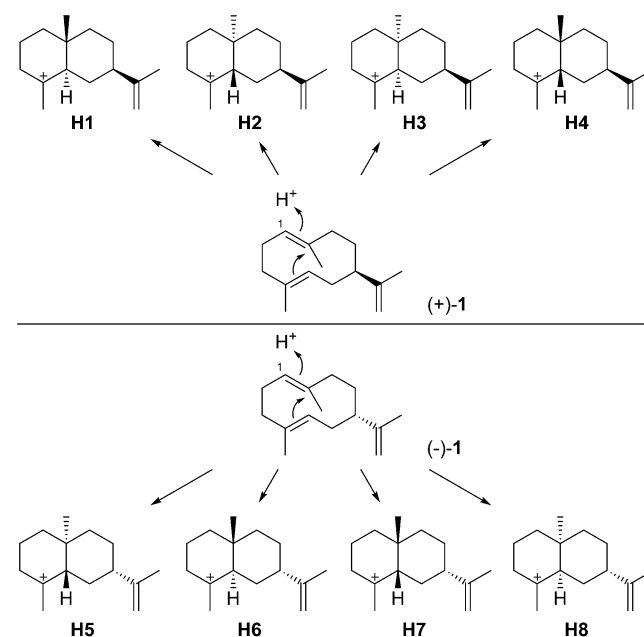
**Scheme 6.** Secondary terpene cyclisations of **1**.

formation of neutral **1** and its reprotonation also an intramolecular or water-mediated proton transfer in cation **B** may directly lead to **H**, **J** or **K**, thus bypassing **1** that would in such cases be better described as a side product rather than an intermediate. However, experimental evidence to distinguish between these alternatives is difficult to obtain, and **1** will preferentially be discussed as an intermediate towards more complex sesquiterpenes in this article. A detailed discussion of the reactions from **1** will follow in the subsequent sections.

### 3. Eudesmanes

#### 3.1. Eudesmanes with a regular skeleton

The protonation-induced cyclisation of **1** can lead to eight stereochemically distinct cationic intermediates (Scheme 7), four of which arise from (+)-**1** (**H1–H4**), while the other four stereoisomers originate from (–)-**1** (**H5–H8**). For each intermediate, simple deprotonations or nucleophilic attack of water are possible. Also, hydride shifts can occur first, which further widens the reachable chemical space of eudesmanes. For many of these possibilities the corresponding structures have been reported.

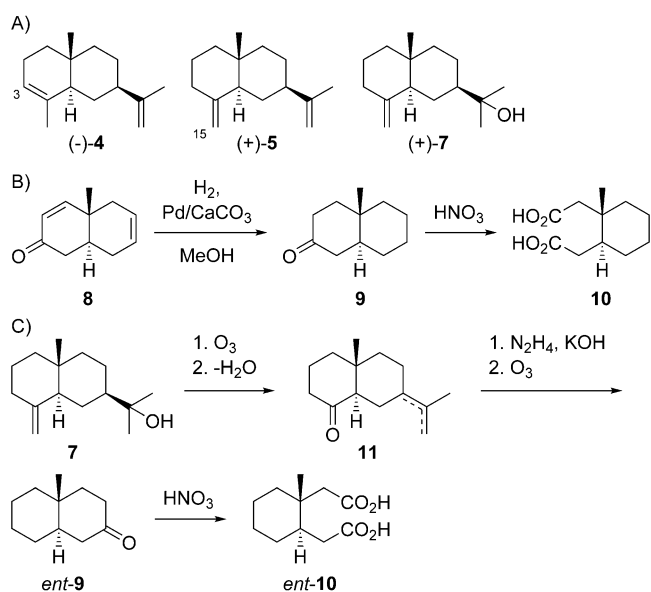


**Scheme 7.** Cyclisations induced by reprotonation of **1** at C1 to **H1–H8**.

#### 3.2. Eudesmanes from cation **H1**

An important intermediate to eudesmanes is **H1**. Deprotonations from C3 and C15 lead to  $\alpha$ -selinene (**4**) and  $\beta$ -selinene (**5**), two compounds that have been isolated more than 100 years ago from celery oil.<sup>[46]</sup> Their structures were elucidated in degradation experiments<sup>[47]</sup> and were correlated to  $\beta$ -eudesmol (**7**, Scheme 8A).<sup>[48–50]</sup> Based on a comparison of physical characteristics of degradation products to those of other *cis*- and *trans*-decalins initially a *cis*-decalin structure was assigned,<sup>[51]</sup>





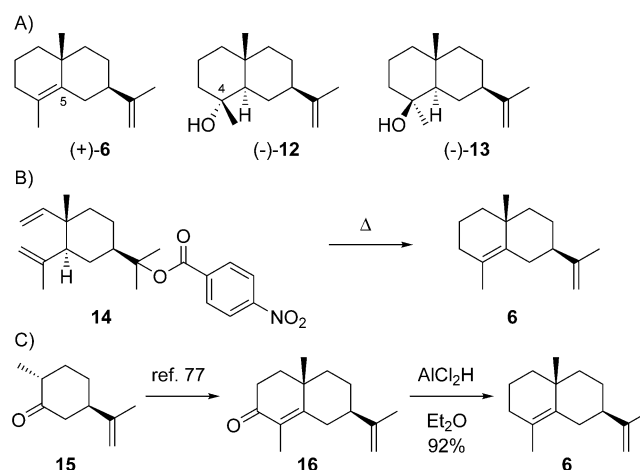
**Scheme 8.** A) Structures of eudesmanes from **H1** and of **7**. B) Chemical correlation of ketone **8** with **10**. C) Chemical correlation of **7** with *ent*-**10**.

but a later conformational re-examination indicated a *trans*-fused ring system.<sup>[52,53]</sup> The absolute configurations of **4** and **5** were determined by chemical correlation through the following arguments. The structure of ketone **8** was established in the classical synthesis of steroids by Woodward.<sup>[54]</sup> Two years later the same group converted **8** into the dicarboxylic acid **10** (Scheme 8B) that was the opposite enantiomer as obtained by degradation of **7** (Schemes 8C)<sup>[55]</sup> that had previously been correlated with **4** and **5** (vide supra).

The optical rotation of **4** was repeatedly found to have a positive value, including the reports from Brazilian rosewood oil ( $[\alpha]_D = +18$ ),<sup>[56]</sup> *Dendropanax trifidus* ( $[\alpha]_D = +68$ )<sup>[57]</sup> and *Cryptotaenia japonica* ( $[\alpha]_D^{15} = +6.3$ ),<sup>[58]</sup> or for **4** obtained by enantioselective synthesis ( $[\alpha]_D = +15.7$ ,  $\text{CHCl}_3$ ).<sup>[59]</sup> Andersen et al. pointed out that minor impurities may result in erroneous data and reported a value of  $[\alpha]_D = -16$  (*c* 0.2, pentane)<sup>[60]</sup> that was confirmed by Maurer and Grieder ( $[\alpha]_D^{20} = -14.5$ ,  $\text{CHCl}_3$ , 1%),<sup>[61]</sup> and in both cases secured by CD spectroscopy. For **5** consistently positive optical rotations with values between  $[\alpha]_D = +31.7$  ( $\text{CHCl}_3$ ) and  $[\alpha]_D = +60$  ( $\text{CHCl}_3$ ) have been given.<sup>[49,57,58,60–66]</sup> Thus, natural  $\alpha$ - and  $\beta$ -selinene from (+)-**1** are characterised as (–)-**4** and (+)-**5**. Complete <sup>1</sup>H- and <sup>13</sup>C-NMR data for **4** and **5** are available.<sup>[66,67]</sup>

Compounds **4** and **5** were identified from various plant sources.<sup>[49,57,58,60–63,66–76]</sup> In some cases **2** was also isolated,<sup>[58,68,69]</sup> sometimes with determined absolute configuration of (–)-**2**,<sup>[61–63]</sup> which supports (+)-**1** as a biosynthetic intermediate, but **1** could also be the true natural product, while **4** and **5** may have been formed spontaneously from **1** during compound isolation (Scheme 4).

An alternative deprotonation of **H1** can lead to selina-4,11-diene (**6**), while the attack of water may result in selin-11-en-4 $\alpha$ -ol (**12**) or neointermedeol (**13**, Scheme 9A). As the stereochemical information at C5 is lost in **6**, this sesquiterpene can



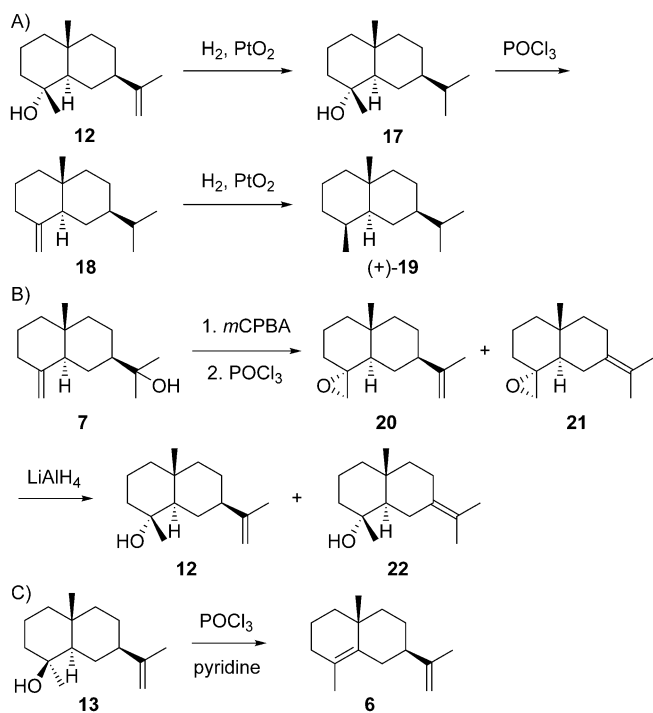
**Scheme 9.** A) Structures of eudesmanes from **H1**. B) Correlation of **14** to **6**. C) Synthesis of **6**.

also arise from **H4**. Conclusions may be possible from co-isolated materials with retained stereochemical information at C5. The absolute configuration of **6** was evident from its formation by pyrolysis of the *p*-nitrobenzoate **14** of (–)-elemol (**3**), leading to (+)-**6** (Scheme 9B).<sup>[65]</sup> This finding is further supported by an enantioselective synthesis of (+)-**6** starting from (+)-*trans*-dihydrocarvone (**15**) through (+)- $\alpha$ -cyperone (**16**),<sup>[77]</sup> followed by reduction of the ketone with  $\text{AlCl}_2\text{H}$  (Scheme 9C).<sup>[78]</sup>

Compound **6** has been isolated from several plants<sup>[62,74,79–83]</sup> with reported positive optical rotations ranging from  $[\alpha]_D^{14} = +32.05$  ( $\text{MeOH}$ )<sup>[79]</sup> to  $[\alpha]_D^{20} = +54.5$  ( $\text{CHCl}_3$ , 1%).<sup>[80]</sup> From *Vernonia glabra* **6** was isolated together with **1**, **2**, **4** and **5** after column chromatography, suggesting that it may have been formed by silicic-acid-catalysed cyclisation of **1**.<sup>[81]</sup> The full<sup>[61]</sup> or partial<sup>[78,79,84]</sup> <sup>1</sup>H-NMR data have frequently been published, but unfortunately no <sup>13</sup>C-NMR data are available from the literature.

The alcohol **12** ( $[\alpha]_D^{20} = -18$ ) was first isolated from *Podocarpus dactyloides* and its structure was correlated to (+)-selinane (**19**), the hydrocarbon corresponding to **4** and **5**, by catalytic hydrogenation to **17**, dehydration with  $\text{POCl}_3$  to **18** and hydrogenation (Scheme 10A), while the  $4\alpha$  orientation of the hydroxy function was deduced from the NMR spectrum, thereby establishing its absolute configuration.<sup>[85]</sup> This structural assignment was confirmed by a synthesis from **7** that was converted into the epoxide and dehydrated with  $\text{POCl}_3$  to yield a mixture of **20** and **21** (Scheme 10B). Epoxide opening with  $\text{LiAlH}_4$  resulted in (–)-**12** and juniper camphor (**22**).<sup>[86]</sup> Furthermore, the racemic compound, along with all other seven stereoisomers, has been synthesised<sup>[87]</sup> and comparative spectroscopic data including <sup>1</sup>H- and <sup>13</sup>C-NMR have been published.<sup>[87,88]</sup> Identical <sup>1</sup>H- and <sup>13</sup>C-NMR data for **12** were reported for the material from *Artemisia barrelieri*<sup>[89]</sup> and *Tanacetum nubigenum*.<sup>[90]</sup> Compound **12** has been isolated from many plant species.<sup>[73,80,82,89–102]</sup>

Neointermedeol (**13**) was first reported from the grass *Boerhaavia intermedia*, with an optical rotation of  $[\alpha]_D^{25} =$

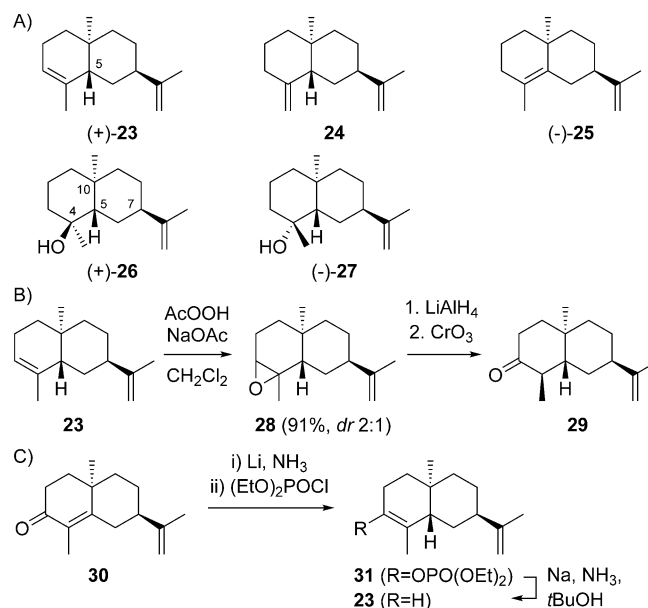


**Scheme 10.** Chemical correlations of 12 with A) 19 and B) 7. C) Correlation of 13 with 6.

+7.5,<sup>[103,104]</sup> while the material isolated later from *Panax ginseng* exhibited a negative optical rotation ( $[\alpha]_D^{22} = -4.8$ ,  $c$  3.45, CHCl<sub>3</sub>).<sup>[84]</sup> To resolve the situation (–)-13 was dehydrated with POCl<sub>3</sub> in pyridine, yielding (+)-6 and thus securing the absolute configuration of 13 (Scheme 10C). The structure of 13 has also been confirmed by synthesis of the racemate.<sup>[87]</sup> Further isolations have been reported from termites including *Subulitermes baileyi*<sup>[105]</sup> and *Amitermes excellens*,<sup>[106]</sup> and from the plants *Geigeria burkei*<sup>[107]</sup> and *Artemisia schmidtiana*.<sup>[108]</sup> Partial <sup>1</sup>H- and full <sup>13</sup>C-NMR data for 13 have been published.<sup>[84,104]</sup>

### 3.3. Eudesmanes from cation H2

Sesquiterpenes arising through H2 occur less frequent in Nature compared to H1 derivatives, but the alcohol 26 (Scheme 11A) is quite widespread. The sesquiterpene 5,10-diepi- $\alpha$ -selinene (23) was first reported from *Dipterocarpus alatus* ( $[\alpha]_D^{20} = +2.1$ ).<sup>[109]</sup> The compound was co-isolated with (7*R*,10*S*)-eudesma-4,11-diene, (–)-25 ( $[\alpha]_D^{20} = -108.6$ ), that could potentially also arise by deprotonation of H3, but if a common terpene cyclisation is assumed, intermediate H2 should be relevant. The absolute configuration of 23 was assigned by epoxidation with peracetic acid to a mixture of stereoisomeric epoxides 28, reduction with LiAlH<sub>4</sub> to yield a mixture of alcohols, and Jones oxidation. From the obtained ketones 29, the enantiomer of a known compound, was isolated as main product (Scheme 11B).<sup>[109]</sup> Further, an enantioselective synthesis of 23 from 30 that is readily accessible from dihydrocarvone 15 was reported, that proceeded by reduction with Li in NH<sub>3</sub> and phosphorylation with (EtO)<sub>2</sub>POCl to 31, followed

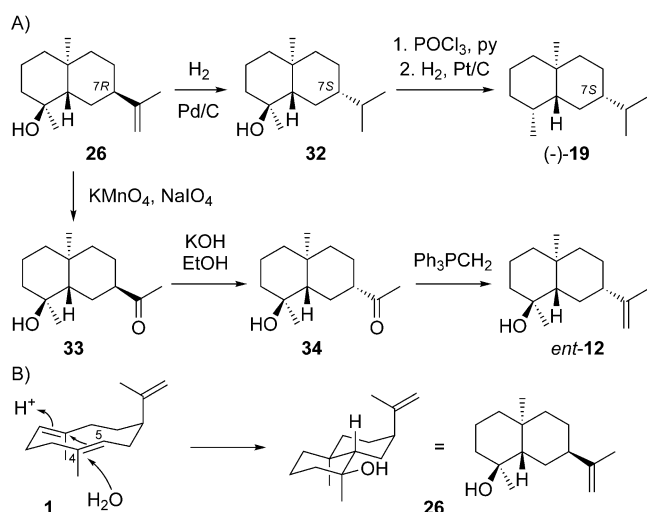


**Scheme 11.** A) Structures of 23–26 and "paradiisol" (27). Chemical correlations of 23 with B) ketone 29 and C) synthetic 30.

by defunctionalisation with Na in NH<sub>3</sub> and *t*BuOH (Scheme 11C).<sup>[110]</sup> Alternatively, 30 can be converted into a mixture of 23, its C5 epimer and 25 by Wolff–Kishner reduction.<sup>[111]</sup> The regioisomer 5,10-diepi- $\beta$ -selinene (24) was first obtained along with 23 by dehydration of a sesquiterpene alcohol with the assigned structure of "paradiisol" (27) from grapefruit (*Citrus paradisi*).<sup>[112]</sup> Subsequent work demonstrated that "paradiisol" was identical with intermedeol (26).<sup>[113]</sup> All three compounds 23–25 were also obtained by hydrolysis of intermedeol  $\beta$ -D-fucopyranoside.<sup>[114]</sup> Compound 23, sometimes accompanied by 24 or 25, has also been reported from several termites.<sup>[106,115,116]</sup> Full <sup>1</sup>H- and <sup>13</sup>C-NMR data of 23 (with missing signals only for quaternary olefinic carbons) and 24 are available from the literature,<sup>[116]</sup> while data for 25 are lacking.

Intermedeol (4*S*,5*S*,7*R*,10*S*)-26 ( $[\alpha]_D^{25} = +10.7$ ) was first reported with 7*S* configuration from *Bothriochloa intermedia*.<sup>[117]</sup> This wrong structural assignment was based on the finding that 26 was converted into (–)-selinene (19) by hydrogenation (Pd/C), dehydration (POCl<sub>3</sub>, pyridine) and hydrogenation (Pt/C, Scheme 12A). The subsequently discovered alcohol 12 (Scheme 9)<sup>[85]</sup> showed different physical characteristics and spectroscopic properties, and thus the structure of *ent*-12 for intermedeol was excluded. Oxidation of 26 with KMnO<sub>4</sub> and NaIO<sub>4</sub> to hydroxyketone 33, followed by epimerisation to 34 and Wittig methylenation gave *ent*-12, supporting a structural revision for intermedeol to 26. The initially observed formation of (–)-19 from 26 was explained by double bond migration and hydrogenation from the sterically less hindered side during Pd catalysis, yielding intermediate 32 with overall epimerisation at C7.<sup>[86]</sup> The structure of 26 was also confirmed by synthesis.<sup>[87,110,111]</sup>

Compound 26 has frequently been isolated from plants.<sup>[66,117–127]</sup> For 26 isolated from *Cymbopogon flexuosus* the



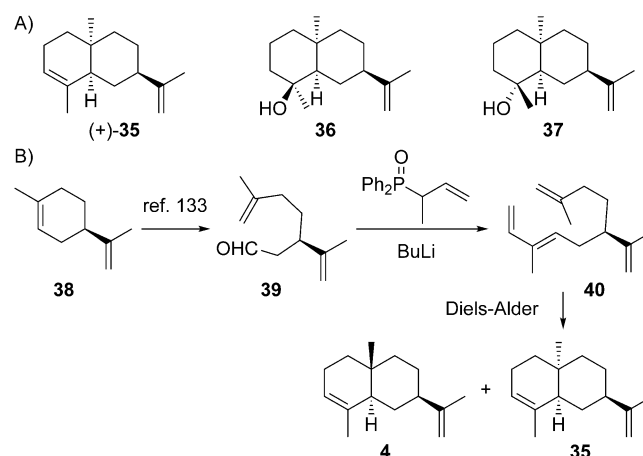
**Scheme 12.** A) Chemical correlations of **26** with (–)-**19** and *ent*-**12**, B) concerted mechanism for the protonation induced cyclisation of **1** to **26**.

opposite absolute configuration was assigned, despite the optical activity of  $[\alpha]_{\text{D}}^{25} = +2$  ( $c$  3.3, MeOH). The compound was named “isointermedeol”<sup>[128]</sup> but this material was likely an impure sample of (+)-**26**.<sup>[129]</sup> Nevertheless, the description of “isointermedeol” caused some confusion, as there is at least one later paper about *Jasonia candicans* with reference to the report of this supposedly new sesquiterpene alcohol.<sup>[130]</sup> For the (+)-intermedeol synthase from *Termitomyces* GC-MS analysis of the products revealed minor amounts of **2**, thereby establishing **1** as a side product and supporting this compound as a biosynthetic intermediate to **26**.<sup>[131]</sup> Another (+)-intermedeol synthase was recently reported from *Streptomyces clavuligerus*.<sup>[132]</sup> Complete <sup>1</sup>H- and <sup>13</sup>C-NMR data of **26** in CDCl<sub>3</sub><sup>[87,88,104,114,119]</sup> or C<sub>6</sub>D<sub>6</sub><sup>[131,132]</sup> have been reported.

Paradisiole (*4R,5S,7R,10S*)-**27** represents the initially assigned structure of a sesquiterpene alcohol from *Citrus paradisi*<sup>[112]</sup> that was later corrected to **26**.<sup>[113]</sup> It may seem surprising that **27** has never been reported as a natural product, while its epimer **26** is widespread, but this is understandable on biosynthetic grounds (Scheme 12B). Starting from the shown conformation of **1**, a concerted protonation induced ring closure and attack of water can lead to **26**, while the formation of **27** by such a process would require a *syn* addition to the C4=C5 double bond of **1** with attack of water from the internal face, which seems sterically impossible. However, compound **27** has been synthesised<sup>[87]</sup> and was obtained as one of the hydrolysis products of intermedeol β-D-fucopyranoside ( $[\alpha]_{\text{D}}^{22} = -17.9$ ,  $c$  0.53, EtOH).<sup>[114]</sup> Full spectroscopic data are available.<sup>[87,88,114]</sup>

### 3.4. Eudesmanes from cation H3

Natural products from **H3** are unknown. Synthetic compounds that could formally arise through **H3** by terpene cyclisation include 10-*epi*-α-selinene (**35**), 7-*epi*-amiteol (**36**) and 5-*epi*-paradisiole (**37**, Scheme 13). Compound **35** was first obtained by Wolff–Kishner reduction of **30**,<sup>[111]</sup> and then from (*R*)-limonene

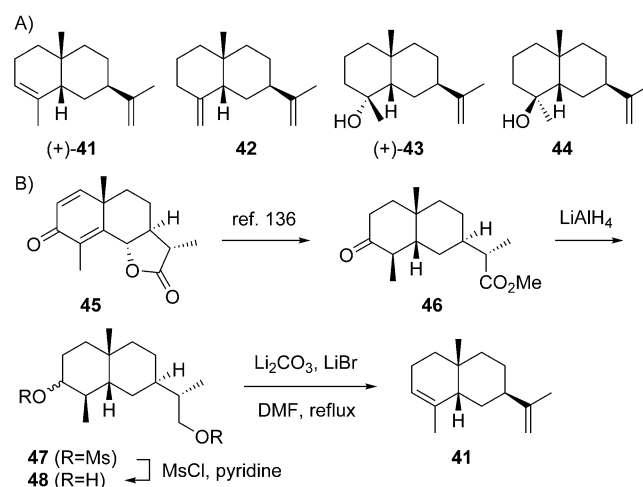


**Scheme 13.** A) Structures of **35**–**37**. B) Enantioselective synthesis of **35**.

(**37**) that can be converted in three steps into the aldehyde **39** (Scheme 13B),<sup>[133]</sup> followed by Wittig–Horner olefination to **40**. An intramolecular Diels–Alder reaction results in the *endo*-adduct **4** and the *exo*-adduct **35** ( $[\alpha]_{\text{D}}^{25} = +102$ , CHCl<sub>3</sub>, 0.7%).<sup>[134]</sup> A similar route was also reported from (*S*)-carvone.<sup>[135]</sup> For **36** and **37** only synthetic routes to the racemates have been established.<sup>[87]</sup> For all three compounds full spectroscopic data have been published.<sup>[87,88,135]</sup>

### 3.5. Eudesmanes from cation H4

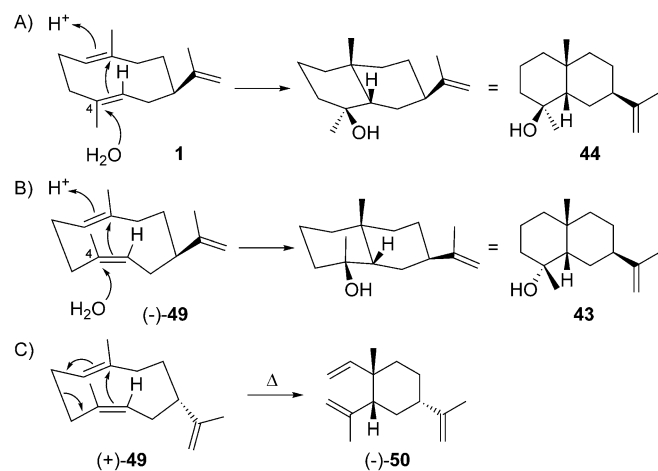
Only a few natural products arising through **H4** are known. Amiteol (+)-**43** ( $[\alpha]_{\text{D}}^{24} = +8$ , CHCl<sub>3</sub>) from the termite *Ami-termes excellens* was the first isolated compound from this class and co-occurred with 5-*epi*-α-selinene (**41**), 5-*epi*-β-selinene (**42**) and **6** in this species (Scheme 14A).<sup>[107]</sup> Although **6** is usually assumed to be formed via **H1**, in *A. excellens* a formation via **H4** is more likely, as this reflects the mechanism for its cometabolites. The absolute configuration of **43** was established by dehydration with SOCl<sub>2</sub>, yielding a mixture of **41**, **42**



**Scheme 14.** A) Structures of **41**–**44**. B) Enantioselective synthesis of **41**.

and (+)-**6** ( $[\alpha]_{\text{D}}^{24} = +30$ ,  $\text{CHCl}_3$ ),<sup>[106]</sup> the same enantiomer as originally reported from *Chamaecyparis formosensis*.<sup>[79]</sup> Furthermore, (+)-**41** was synthesised from  $\alpha$ -santonin (**45**) that was converted into **46** through a known route (Scheme 14B).<sup>[136]</sup> Reduction of **46** to epimeric diols **47**, mesylation to **48** and elimination with  $\text{Li}_2\text{CO}_3$  and  $\text{LiBr}$  in refluxing DMF yielded **41** ( $[\alpha]_{\text{D}}^{25} = +30.1$ ,  $c$  3.50,  $\text{CHCl}_3$ ).<sup>[137]</sup> Syntheses for racemic **43** and 5-*epi*-neointermedeol (**44**) have also been established,<sup>[87]</sup> but despite its tentative GC/MS based identification as constituent of some essential oils compound **44** has not been isolated from natural sources so far. More recently, a terpene synthase for **41** has been identified from the cyanobacterium *Nostoc punctiforme*, but the absolute configuration of the product has not been assigned.<sup>[138]</sup> Full spectroscopic data including IR,  $^1\text{H}$ - and  $^{13}\text{C}$ -NMR are available for **41**,<sup>[137,138]</sup> **43** and **44**.<sup>[87,88]</sup>

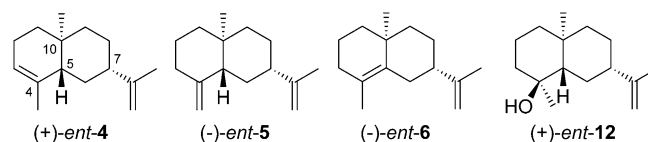
Notably, while the formation of the sesquiterpene hydrocarbons **41**, **42** and **6** should be possible through **H4**, the formation of **43** along this pathway encounters a difficulty that is related to the explanation for the possible formation of **26**, but not of **27**, from **H2** (Scheme 12). Along similar lines (Scheme 15A), the protonation induced cyclisation of **1** starting from a boat–boat conformation can explain the biosynthesis of **44**, while the formation of **43** would require the nucleophilic attack of water from the sterically less accessible *Re* face at C4. However, the formation of **43** is well understandable, if a precursor with a C4=C5 *Z*-configured double bond would be assumed (Scheme 15B). This precursor is known as (–)-helminthogermacrene (**49**) from the fungus *Helminthosporium sativum*<sup>[139]</sup> and later from the termite *Amitermes wheeleri*.<sup>[140]</sup> The enantiomer (+)-**49** was reported from the liverwort *Scapania undulata* and has a very similar EI mass spectrum and GC retention index to **1**, but is less prone to a Cope rearrangement to (–)-*cis*- $\beta$ -elemene (**50**, Scheme 15C).<sup>[141]</sup> Synthetic routes towards racemic **49** have been developed<sup>[139,142]</sup> and the absolute configuration of (+)-**49** was established by chemical correlation to (–)-helmiscapene, a compound discussed in Section 3.8.<sup>[39]</sup>



**Scheme 15.** Protonation induced cyclisations A) of **1** to **44** and B) of **49** to **43**. C) Cope rearrangement of **49** to **50**.

### 3.6. Eudesmanes from cation H5

Compounds derived from (–)-**1** through the enantiomeric series of intermediates **H5**–**H8** have been reported less often compared to those from (+)-**1**, which may be attributed to the fact that still most work has been done on higher plants for which (+)-**1** is the typical enantiomer (Section 2). The cation **H5** gives rise to the known natural products *ent*- $\alpha$ -selinene (*ent*-**4**), *ent*- $\beta$ -selinene (*ent*-**5**), *ent*-selina-4,11-diene (*ent*-**6**) and (4*S*,5*S*,7*S*,10*S*)-eudes-11-en-4-ol (*ent*-**12**, Figure 1).

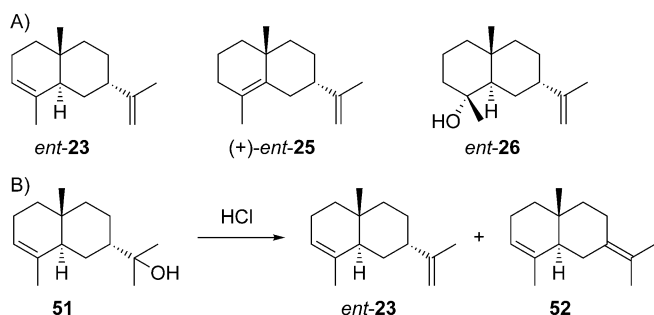


**Figure 1.** Structures of *ent*-**4**–*ent*-**6** and *ent*-**12**.

The first report about naturally occurring enantiomers of selinene sesquiterpenes identified *ent*-**4** as a constituent of the liverwort *Chiloscyphus polyanthus* in 1973. Its absolute configuration was established by CD spectroscopy in comparison to authentic (–)-**4**.<sup>[60]</sup> Compounds *ent*-**4** and *ent*-**6**, likewise established by CD spectroscopy and accompanied by **2**, were subsequently reported from the liverworts *Diplophyllum albicans* and *D. taxifolium*,<sup>[143]</sup> while the liverworts *Riccardia jackii*, *Bazzania spiralis* and *Tylimanthus tenellus* contain different combinations of *ent*-**4**, *ent*-**5** and *ent*-**12**.<sup>[144–147]</sup> Also insects were reported to contain *ent*-**4** and (+)-**2**, exemplified by their occurrence in *Ceroplastes ceriferus*, which is surprising considering the fact that the „normal“ enantiomeric series of compounds is present in the related species *C. rubens*.<sup>[62]</sup> In all these examples the absolute configurations were determined from the optical rotations of the isolated compounds. In *Penicillium roqueforti* also *ent*-**4**, *ent*-**5** and *ent*-**12** may occur; in this case the absolute configurations were assigned based on their biosynthetic relationship to aristolochene (vide infra) that is generated through (–)-**1** in this fungus.<sup>[148]</sup>

### 3.7. Eudesmanes from cation H6

Little is known about eudesmanes arising via cationic intermediate **H6**. The compound 7-*epi*- $\alpha$ -selinene (*ent*-**23**, Scheme 16A) was first reported from *Amyris balsamifera*, a species from which also 7-*epi*- $\alpha$ -eudesmol (**51**, Scheme 16B) was isolated and structurally characterised by NMR spectroscopy. From its positive optical rotation ( $[\alpha]_{\text{D}} = +10$ ,  $c$  1.8,  $\text{CHCl}_3$ ) the authors concluded on the shown absolute configuration for **51**, but a comprehensible explanation for this assignment is missing. Dehydration of **51** yielded a mixture of two products to which the structures of *ent*-**23** and **52** were assigned by NMR spectroscopy, unfortunately without separating the obtained materials and determining their optical rotations. The compounds described as *ent*-**23** and **52** also occurred in the essential oil of *A. balsamifera*.<sup>[149]</sup> One study reported the chro-



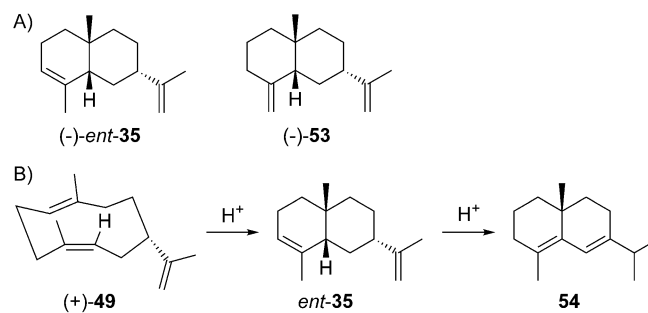
**Scheme 16.** A) Structures of *ent*-23, *ent*-25 and *ent*-26. B) Dehydration of 51.

matographic separation of the compound from *A. balsamifera* and (+)-23 (the latter with a mentioned source „provided by Dr. Wilfried König“) on a chiral stationary GC phase, which represents the only hint in the literature that the structure of *ent*-23 for the essential oil constituent may be correctly assigned.<sup>[150]</sup> Compound *ent*-23 was also reported as major product of a terpene synthase from *Vitis vinifera*.<sup>[150, 151]</sup> Both enantiomers of 23 have been obtained by synthesis from the enantiomers of 15, but optical rotary powers of the products were not measured.<sup>[152]</sup> However, *ent*-23 may have a negative optical rotation, as for 23 from *Dipterocarpus alatus* a low value of  $[\alpha]_D^{20} = +2.1$  was determined.<sup>[109]</sup> This would be consistent with a report by König in which *ent*-23 was published as the (–)-enantiomer, albeit only based on separation by gas chromatography using a chiral stationary phase without isolation.<sup>[153]</sup>

Compound *ent*-25 ( $[\alpha]_D^{16} = +46.5$ ,  $c$  0.85,  $\text{CHCl}_3$ ) has been synthesised using the same strategy as for 6 (Scheme 9C),<sup>[78]</sup> but has not been isolated from any organism. The only report about *ent*-26 from *Monactis macbridei* by Bohlmann and co-workers<sup>[154]</sup> gives a reference to the erroneous “isointermedeol”<sup>[128]</sup> that was corrected shortly after.<sup>[129]</sup> Unfortunately, Bohlmann’s paper does not give an optical rotation for the isolated material so that it is difficult to judge, if the authors of this study were aware of the misassignment of “isointermedeol” at the time of their publication. Overall, this discussion shows that compounds from H6 are not only rare, but if they occur in the literature, the assignments of absolute configurations remain unclear. Since the compounds originate in all cases from higher plants, they may truly be the usual enantiomers, that is, 23, 25 and 26.

### 3.8. Eudesmanes from cations H7 and H8

The literature contains only few reports of compounds that may originate from H7, while no examples from H8 are available.  $\alpha$ -Helmiscapene (*ent*-35, Scheme 17A) was first isolated from *Scapania undulata* and suggested to arise through a “cis-germacrene”,<sup>[155]</sup> a compound that was later described from this species<sup>[141]</sup> after its first identification from *H. sativum* as helminthogermacrene (49).<sup>[139]</sup> In agreement with the positive optical rotation of synthetic 35 (Scheme 13), *ent*-35 was found to be the (–)-enantiomer ( $[\alpha]_D = -100$ ,  $\text{CHCl}_3$ ) and correlated



**Scheme 17.** A) Structures of *ent*-35 and 53. B) Acid-catalysed cyclisation of 49 to *ent*-35 and isomerisation to 54.

to (+)- $\delta$ -selinene (54) by acid-catalysed isomerisation (Scheme 17B). Both *ent*-35 and  $\beta$ -helmiscapene (–)-53 were also found in the liverwort *Radula perrottetii*.<sup>[156]</sup> The acid-catalysed cyclisation of (+)-49 to *ent*-35 suggests that the formation of *ent*-35 from 49 could be non-enzymatic and that germacrene A may indeed not be the precursor of helmiscapenes.<sup>[39]</sup> Full  $^1\text{H}$ - and  $^{13}\text{C}$ -NMR data are available for *ent*-35 and 53.<sup>[39, 156]</sup>

## 4. Rearranged Eudesmanes

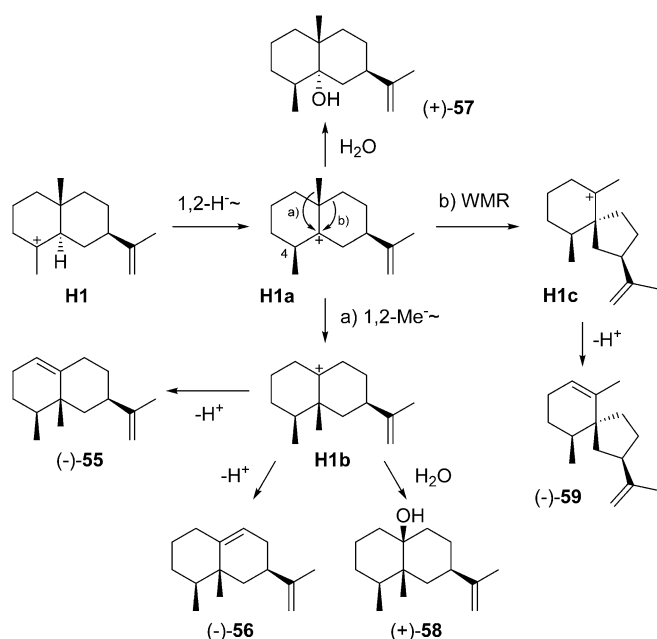
In this section rearranged eudesmanes from H1–H6 will be discussed, while such compounds from H7 and H8 are unknown.

### 4.1. Rearranged eudesmanes from H1

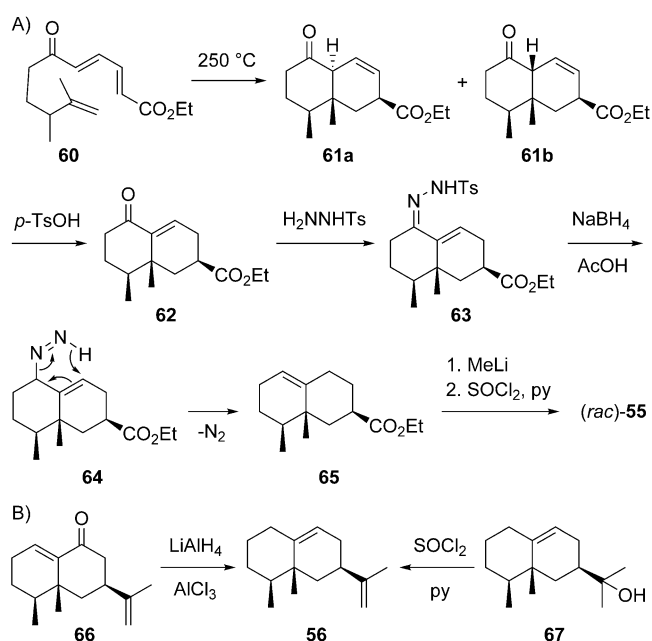
Rearranged eudesmanes can in theory arise from all cations H1–H8 in Scheme 7. An important group of compounds by widespread occurrence in Nature originates from H1. Specifically, this intermediate can undergo a 1,2-hydride migration to H1a that must proceed suprafacially and thus determines the configuration at C4 (Scheme 18; 1,n-hydride or proton migrations as used in this article refer to the distance of n carbons for the migration, not to positional numbers). A subsequent 1,2-methyl group migration leads to H1b (path a) that upon deprotonation yields eremophilene (55) or 4,5-*diepi*-aristolochene (56). Alternatively, H1a can react in a Wagner–Meerwein rearrangement (WMR) with ring contraction to H1c that results in hinesene (59, path b).

Compound 55 was first isolated from *Petasites officinalis* and *P. albus* ( $[\alpha]_D^{20} = -104.2$  and  $[\alpha]_D^{24} = -142.5$ , respectively).<sup>[157–159]</sup> Its structure was initially wrongly assigned,<sup>[160]</sup> but then corrected based on a chemical derivatisation and interpretation of the EI-MS fragmentation behaviour of a thioketal derivative.<sup>[159]</sup> The sesquiterpene 55 was later isolated from several higher plants.<sup>[58, 161–167]</sup> Furthermore, (–)-55 was discovered in the gorgonian *Plexaurella fusifera*<sup>[168]</sup> and along with 2 in the liverwort *Frullania serrata*.<sup>[169]</sup>

An elegant synthesis for (*rac*)-55 has been developed starting from 60 that can give 61a by a Diels–Alder reaction, with partial epimerisation to 61b (Scheme 19A). Both compounds can be converted into 62 by acid-catalysed isomerisation. Reaction with tosylhydrazine leads to 63 that was reduced with



Scheme 18. Biosynthesis of rearranged eudesmanes from H1.

Scheme 19. A) Synthesis of (*rac*)-**55** through a Diels–Alder approach, B) preparation of **56** from the natural products **66** and **67**.

NaBH<sub>4</sub> via **64** to **65**.<sup>[170]</sup> Treatment with MeLi and dehydration with SOCl<sub>2</sub> in pyridine gave **55**.<sup>[171]</sup> Its double bond regioisomer **56** (Scheme 19B) was first obtained from eremophilone (**66**), the first structurally characterised terpene found to violate Ruzicka's isoprene rule,<sup>[172]</sup> by reduction with LiAlH<sub>4</sub> and AlCl<sub>3</sub>,<sup>[173]</sup> and later from eremophil-9-en-11-ol (**67**) by dehydration ([α]<sub>D</sub> = -11.1, c 0.18, CHCl<sub>3</sub>).<sup>[174]</sup> Compound **56** has also been obtained by synthesis from capsidiol,<sup>[175]</sup> but was never isolat-

ed from Nature. Complete <sup>13</sup>C-NMR data are available for **55** and **56**.<sup>[170,175]</sup>

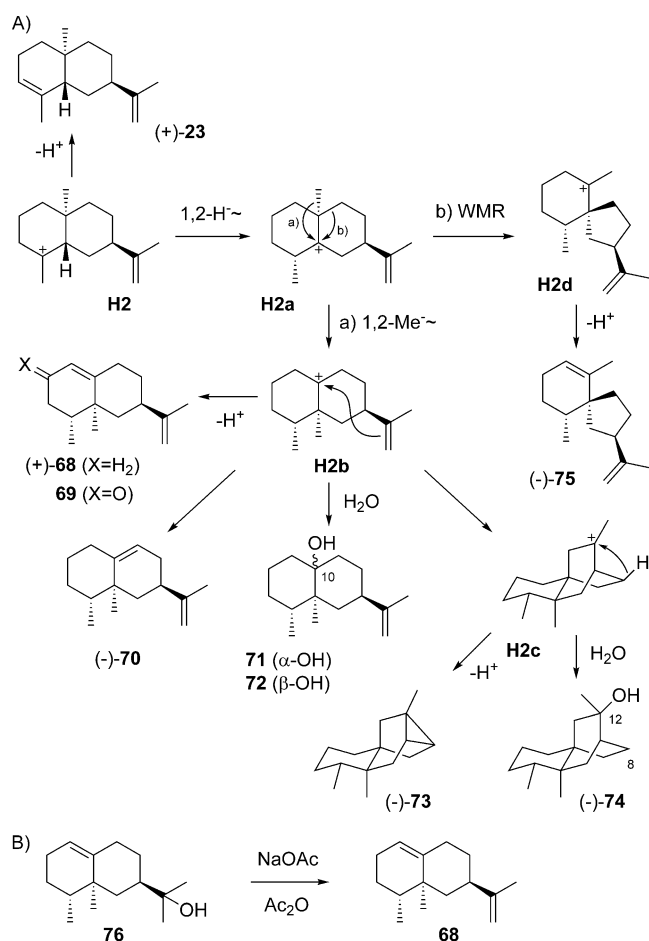
The sesquiterpene alcohol 4αH-eudesma-11-en-4α-ol (**57**), [α]<sub>D</sub> = +32.8 (c 0.7, CHCl<sub>3</sub>), was isolated from *Kleinia pendula* and can arise by attack of water to **H1 a**.<sup>[176]</sup> Similarly, the addition of water to **H1 b** leads to eremophil-11-en-10β-ol (**58**), a compound that is known from *Alpinia intermedia* ([α]<sub>D</sub> = +29.2, c 0.12, CHCl<sub>3</sub>).<sup>[66]</sup> For both alcohols **57** and **58** full <sup>13</sup>C-NMR data were given.<sup>[66]</sup>

Hinesene (**59**) was first isolated from *Rolandra fruticosa* ([α]<sub>D</sub><sup>24</sup> = -44, c 0.1, CHCl<sub>3</sub>).<sup>[177]</sup> The absolute configuration was initially assigned based on the same sign of optical rotation than for hinesol and later confirmed by enantioselective synthesis from santonin.<sup>[178]</sup> The compound is also known from an unspecified liverwort of the genus *Frullania*.<sup>[179]</sup> Full <sup>1</sup>H- and <sup>13</sup>C-NMR data were provided.<sup>[177,178]</sup>

## 4.2. Rearranged eudesmanes from H2

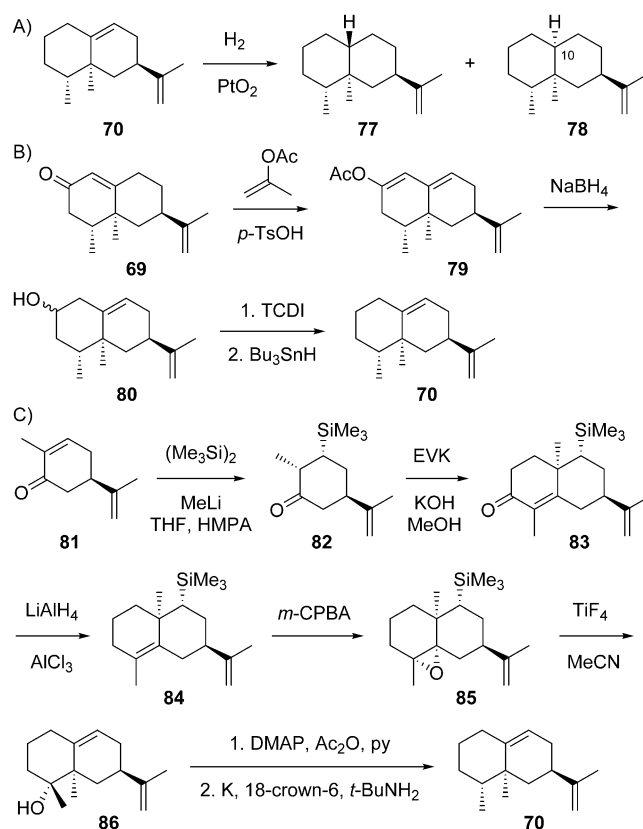
Also rearranged eudesmanes from **H2** constitute an important group of compounds (Scheme 20A), including (+)-valencene (**68**), (-)-aristolochene (**70**), valencene hydrate (**71**) and its C10 epimer **72**, (-)-ishwarane (**73**), (-)-8,12-*seco*-ishwaran-12-ol (**74**) and (-)-agaraspirene (**71**). Compound **73** requires a third cyclisation from **H2 b** to **H2 c** and deprotonation with closure of a cyclopropane ring, while **74** can be explained by attack of water to **H2 c**.

Valencene (**68**) was first isolated from orange oil<sup>[180]</sup> and found to be related to nootkatone (**69**) by oxidative conversion,<sup>[181]</sup> an important value adding transformation for which an artificial enzyme system has been developed.<sup>[182]</sup> Compound **69** is a flavour constituent of citrus fruits and its structure had previously been established.<sup>[183]</sup> The optical rotation of **68** was determined for the material obtained by dehydration of valerianol (**76**, Scheme 20B) with NaOAc in refluxing Ac<sub>2</sub>O ([α]<sub>D</sub> = +73.4, c 5.3, CHCl<sub>3</sub>).<sup>[184]</sup> A synthesis of (*rac*)-**68** similar to the synthesis of (*rac*)-**55** in Scheme 19A has been developed.<sup>[170]</sup> The sesquiterpene **68** is a constituent of the essential oils from numerous plants, but has rarely been isolated. *Bixa orellana* is one of the few sources from which its isolation was mentioned,<sup>[185]</sup> while it was obtained enriched together with **2** in a sesquiterpene hydrocarbon fraction from the liverwort *Porrella acutifolia*.<sup>[186]</sup> The combination of **2** and **68** also occurs in the octocoral *Plexaurella fusifera*,<sup>[188]</sup> while **68** from bacteria is rare, but has been identified from *Streptomyces* sp. FORM5.<sup>[187]</sup> Valencene synthases are known from *Citrus sinensis*,<sup>[188]</sup> *Vitis vinifera*,<sup>[150,151]</sup> and *Callitropsis nootkatensis*,<sup>[189]</sup> in which it occurs together with a valencene oxidase for the biosynthesis of **69**.<sup>[190]</sup> Besides **68**, the terpene synthases from *V. vinifera* were reported to produce (-)-7-*epi*-selinene (*ent*-**23**, Scheme 16)<sup>[150,151]</sup> that must originate from **H6**. It would be easier to understand, if one of the two enzyme products would represent the opposite enantiomer than reported, so that both could arise through a common intermediate. In fact, the configurational assignment for **68** was based on a GC analysis using a chiral stationary phase, but without including a (-)-**68** standard.



**Scheme 20.** A) Biosynthesis of compounds from H2. B) Dehydration of 76 to 68.

Aristolochene (**70**,  $[\alpha]_D^{25} = -76.47$ ) was first isolated from *Aristolochia indica*. Its structure was elucidated by NMR spectroscopy and catalytic hydrogenation, yielding a mixture of (+)-nootkatane (**77**), also obtained by hydrogenation of **68**, and its C10 epimer **78** (Scheme 21 A).<sup>[191]</sup> The structural assignment was later confirmed by a synthesis of **70** from **68**, that was first oxidised to **69**, followed by conversion into the dienol acetate **79** (Scheme 21 B). Deconjugation by reduction with NaBH<sub>4</sub> gave **80** that was defunctionalised with thiocarbonyldiimidazole and Bu<sub>3</sub>SnH to yield **70**.<sup>[192]</sup> Furthermore, an enantioselective synthesis from (*S*)-carvone (**81**) has been developed (Scheme 21 C). After silylation to **82**, a Robinson annelation with ethylvinyl ketone resulted in **83**. Its reduction with excess LiAlH<sub>4</sub> and AlCl<sub>3</sub> to **84** was followed by epoxidation to **85**. Treatment with TiF<sub>4</sub> resulted in epoxide opening with methyl group migration and cleavage of the trimethylsilyl cation to produce **86**, that was defunctionalised in two more steps to **70**.<sup>[193]</sup> Compound **70** was also reported as a side product of valencene synthase from *V. vinifera*<sup>[150]</sup> and as a headspace constituent from *Streptomyces acidiscabies*.<sup>[194]</sup> Both compounds (+)-**68** and (-)-**70** are present in extracts from the liverwort *Dumortiera hirsuta* with absolute configurations established in comparison to authentic standards by GC using a chiral sta-



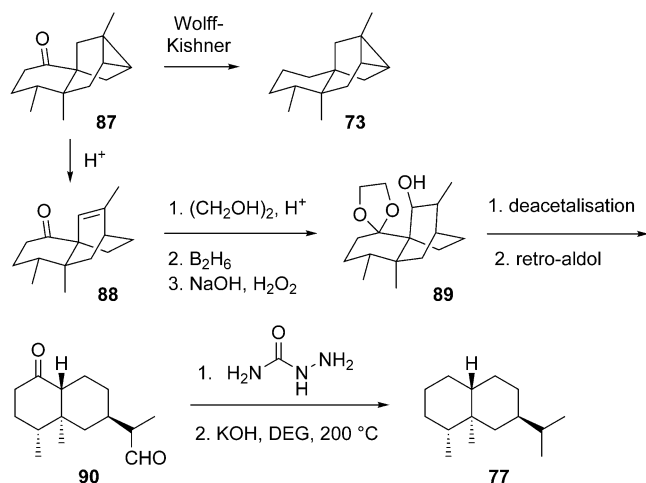
**Scheme 21.** A) Hydrogenation of **70**. B) Synthesis of **70** from **69**, and C) from (*S*)-carvone (**81**).

tionary phase.<sup>[153]</sup> Full <sup>13</sup>C-NMR data for **68**<sup>[170]</sup> and **70**<sup>[193,195,196]</sup> have been reported.

Valencene hydrate (**71**), arising from **H2b** by attack of water, has been isolated from orange juice. For comparison this compound and its C10 epimer **72** were synthesised from **68** by epoxidation and epoxide opening with LiAlH<sub>4</sub>. Unfortunately, no optical rotations were given, but full <sup>13</sup>C-NMR data are available.<sup>[197]</sup>

(-)-Ishwarane (**73**,  $[\alpha]_D = -40.33$ ) was first isolated from *Aristolochia indica* where it co-occurs with biosynthetically linked (+)-ishwarone (**87**) that can be converted into **73** by Wolff-Kishner reduction (Scheme 22).<sup>[191]</sup> Compound **87** undergoes ring opening to (-)-isoishwarone (**88**) by treatment with acid.<sup>[198]</sup> Its further conversion by acetalisation, hydroboration and oxidation leads to **89**, that upon deacetalisation and retroaldol reaction results in **90**. Reduction through the bis-semicarbazone yields (+)-nootkatane (**77**), thus firmly establishing the absolute configuration of **73**.<sup>[199]</sup> Ishwarane was subsequently also found in many other plants,<sup>[185,200–204]</sup> while 8,12-*seco*-ishwaran-12-ol (**74**,  $[\alpha]_D = -165$ , *c* 0.1, CHCl<sub>3</sub>) has only once been reported from *Litsea amara*.<sup>[205]</sup> Its absolute configuration has not been formally established, but was suggested to correspond to that of **73**. Full <sup>13</sup>C-NMR data for **73** and **74** are available.<sup>[206,207]</sup>

(-)-Agarospirene (**75**) was first obtained by pyrolysis of the benzoate ester of agarospirol, a compound isolated from agar-

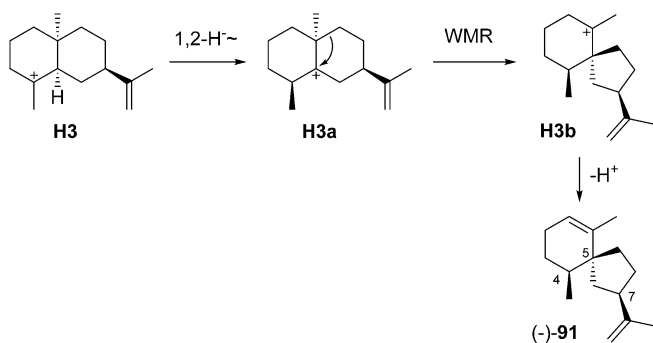


Scheme 22. Chemical correlation of ishwarane (73) with nootkatane (77).

wood.<sup>[207]</sup> Its structure has also been ascribed to a natural product isolated from the liverworts *Scapania robusta* and *Scapania maxima*,<sup>[208,209]</sup> but a later synthesis of 75 ( $[\alpha]_D^{22} = -11$ ,  $c$  0.3) and its stereoisomers demonstrated that the natural product was identical to (–)-hinesene (59).<sup>[178]</sup> Complete <sup>1</sup>H- and <sup>13</sup>C-NMR data for 75 were reported.<sup>[178]</sup>

#### 4.3. Rearranged eudesmanes from H3

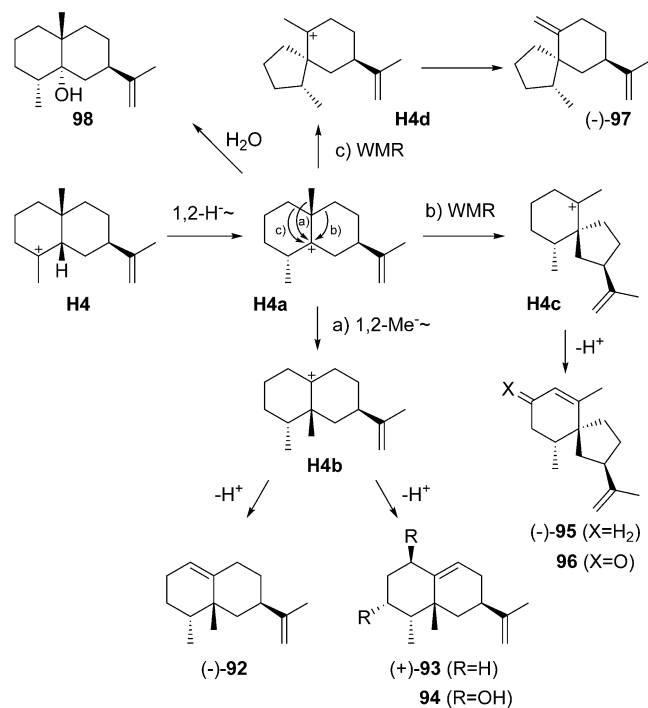
Natural rearranged eudesmanes from H3 are unknown. The only known compound is (4*S*,5*R*,7*R*)-spirovetivadiene (91) that has been obtained by synthesis ( $[\alpha]_D^{22} = -3$ ,  $c$  0.6). Its hypothetical biosynthesis from H3 would require a 1,2-hydride shift to H3a, ring contraction to H3b and deprotonation (Scheme 23). Full <sup>1</sup>H- and <sup>13</sup>C-NMR data are available.<sup>[178]</sup>



Scheme 23. Rearranged eudesmanes from H3: spirovetivadiene (91).

#### 4.4. Rearranged eudesmanes from H4

Known rearranged eudesmanes from intermediate H4 (Scheme 24) are represented by (–)-4-*epi*-eremophilene (92), (+)-5-*epi*-aristolochene (93), (–)-premnaspirodiene (95, also named spirovetivene), (–)-spirolepechinene (96) and 4βH,7αH,10β-eudesm-11-en-4α-ol (98). The unusual sesquiterpene 97 requires a ring contraction to H4d and deprotonation.



Scheme 24. Biosynthesis of rearranged eudesmanes from H4.

Both compounds 92 ( $[\alpha]_D^{25} = -22.7$ ,  $c$  0.17, CHCl<sub>3</sub>) and 93 ( $[\alpha]_D^{25} = +8.13$ ,  $c$  0.16, hexane) were obtained by synthesis from capsidiol (94).<sup>[175,210]</sup> Notably, 93 is also the biosynthetic precursor to 94,<sup>[211]</sup> as was demonstrated by incubation of [1,1-<sup>3</sup>H<sub>2</sub>]FPP with cell-free enzyme preparations from *Nicotiana tabacum*, yielding radioactively labelled 93. Furthermore, <sup>14</sup>C-labelled 93 was incorporated into 94 in feeding experiments with *N. tabacum* and *Capsicum annum*.<sup>[212,213]</sup> Subsequent work resulted in the purification of tobacco 5-*epi*-aristolochene synthase (TEAS),<sup>[214]</sup> cloning of the genes from *N. tabacum* and *C. annum* and expression in *Escherichia coli*,<sup>[215–217]</sup> and determination of the first crystal structure of a plant terpene synthase.<sup>[218]</sup> Based on this structure the active site residue Tyr520 was suggested to be responsible for reprotonation of the intermediate (–)-1. Consistent with this hypothesis, the Y520F enzyme variant gave (–)-1 as a single product.<sup>[219]</sup> Also the 5-*epi*-aristolochene-1,3-dihydroxylase for the biosynthesis of 94 from 93 has been identified.<sup>[220]</sup> For the biotechnological access to 93 the *epi*-aristolochene synthase gene has been heterologously expressed in *E. coli*,<sup>[221]</sup> in *Oryza sativa*,<sup>[222]</sup> and in yeast in which optimisation of the strain and the culture conditions resulted in a high titre production.<sup>[223]</sup> A thermostable variant of EAS has been created.<sup>[224]</sup>

Along similar lines of research, 95 has first been isolated from *Premna latifolia*<sup>[225]</sup> and subsequently from *Lepechinia bulata* ( $[\alpha]_D^{20} = -88$ ,  $c$  0.501, CHCl<sub>3</sub>) in which it co-occurs with 97 ( $[\alpha]_D^{20} = -32$ ,  $c$  0.125, CHCl<sub>3</sub>).<sup>[226]</sup> The premnaspirodiene synthase (also known as vetispirodiene synthase) from *Hyoscyamus muticus* (HPS) has been characterised.<sup>[227,228]</sup> Another sesquiterpene synthase (Tps32) from *Solanum lycopersicum* with 90% sequence identity to HPS was initially described as viridi-

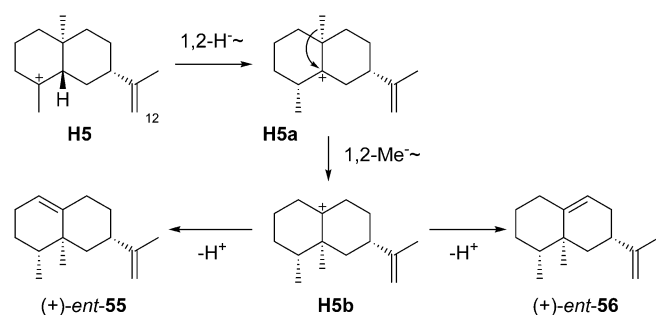


florene synthase,<sup>[229]</sup> but a later study showed that Tps32 is indeed active as prenospiadiene synthase.<sup>[230]</sup> Compound **95** is the parent hydrocarbon of (–)-solavetivone (**96**),<sup>[231,232]</sup> for which a prenaspiadiene oxygenase was reported.<sup>[233]</sup>

A detailed analysis of the product profiles of TEAS and HPS has led to the characterisation of several side products and demonstrated that TEAS produces minor amounts of **95**,<sup>[234]</sup> while HPS generates small quantities of **93** from FPP.<sup>[235]</sup> Domain swapping experiments between TEAS and HPS resulted in enzyme variants making mixtures of **93** and **95** and allowed the identification of domains that conferred specificity for these two products.<sup>[236]</sup> After the crystal structure of TEAS had become available, a systematic and rational approach targeting nine selected residues within and near the active site in all 2<sup>9</sup> = 512 combinations for a functional interconversion between TEAS and HPS was surveyed.<sup>[237,238]</sup> Finally, compound **98** has been isolated from orange juice. <sup>1</sup>H- and <sup>13</sup>C-NMR data for **92**,<sup>[175]</sup> **93**,<sup>[210]</sup> **95**,<sup>[178,226]</sup> **97**,<sup>[226]</sup> and **98**<sup>[197]</sup> have been published.

#### 4.5. Rearranged eudesmanes from H5

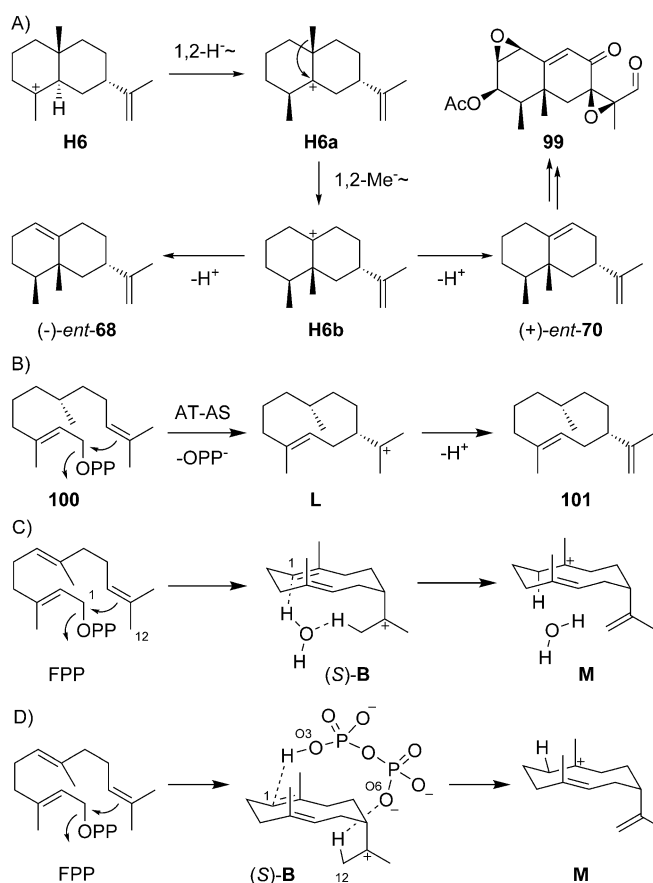
Only a few reports about rearranged eudesmanes from **H5** from Nature are available (Scheme 25). Terpene synthases for *ent*-**55** have been characterised from the myxobacterium *Sorangium cellulosum* ([ $\alpha$ ]<sub>D</sub><sup>25</sup> = +131.7, *c* 1.0, CHCl<sub>3</sub>)<sup>[239]</sup> and the plant pathogenic fungus *Fusarium fujikuroi*.<sup>[240]</sup> The cyclisation mechanism of (+)-eremophilene synthase from *F. fujikuroi* was studied by isotopic labelling experiments that showed selective deprotonation from C12 of FPP in the formation of the intermediate (–)-**1**, allowed to follow the 1,2-hydride shift from **H5** to **H5a**, and demonstrated that the final deprotonation from **H5b** to *ent*-**55** proceeds with loss of the same proton as incorporated in the cyclisation of (–)-**1** to **H5** (Scheme 7).<sup>[240]</sup> A crystal structure of *ent*-**55**<sup>[239]</sup> and full NMR data assignments have been published.<sup>[239,240]</sup> Only a synthetic study towards *ent*-**56** ([ $\alpha$ ]<sub>D</sub><sup>25</sup> = +12.5, *c* 2.5, CHCl<sub>3</sub>) is available.<sup>[241]</sup>



Scheme 25. Biosynthesis of rearranged eudesmanes from **H5**.

#### 4.6. Rearranged eudesmanes from H6

Rearranged molecules from **H6** (Scheme 26A) are (–)-valencene (*ent*-**68**) and (+)-aristolochene (*ent*-**70**) that has been isolated from *Aspergillus terreus* ([ $\alpha$ ]<sub>D</sub> = +79.4, *c* 0.0176, hexane),<sup>[192,196]</sup> and *Penicillium roqueforti*, in which it occurs to-



Scheme 26. A) Biosynthesis of rearranged eudesmanes from **H6**. B) Cyclisation of (*R*)-5,6-dihydro-FPP (**100**) to **101** by AT-AS. C) Proposed water-mediated proton transfer from (*S*)-**B** to **M** in the biosynthesis of *ent*-**70**.

gether with **2**.<sup>[148,242,243]</sup> The absolute configuration has been established by synthesis of (–)-**70** from (+)-valencene (**68**).<sup>[192]</sup> (+)-Aristolochene synthase was first isolated from *P. roqueforti* (PR-AS)<sup>[244]</sup> and is also present in *A. terreus* (AT-AS).<sup>[245]</sup> Subsequent gene cloning and expression gave efficient access to the recombinant enzymes.<sup>[246,247]</sup> A biphasic flow reactor system for the biocatalytic production of *ent*-**70** has been developed.<sup>[248]</sup>

Notably, PR-AS produces a mixture of *ent*-**70** as the main and *ent*-**68** and (–)-**1** as side products, while AT-AS yields *ent*-**70** as a single product.<sup>[249,250]</sup> Isotopic labelling experiments demonstrated that the cyclisation of FPP to *ent*-**70** proceeds with inversion of configuration at C1 and the specific loss of a proton from C12.<sup>[245]</sup> The E252Q variant of PR-AS yielded (–)-germacrene A (**1**) as the only product.<sup>[250]</sup> Further support of (–)-**1** as an intermediate was obtained by the observed cyclisation of (*R*)-5,6-dihydro-FPP (**100**) to the germacrene A analogue **101** by AT-AS (Scheme 26B).<sup>[251]</sup> Similar experiments have been carried out with fluorinated FPP analogues.<sup>[252,253]</sup> On the other hand, instead of a true pathway intermediate, (–)-**1** could only be a shunt product. Allemann and co-workers have argued for this view, as (–)-**1** was not accepted as a substrate by PR-AS,<sup>[249]</sup> and a computational study showed feasibility of a water-mediated direct proton transfer from (*S*)-**B** to **M** that could further cyclise to **H6** (Scheme 26C).<sup>[254]</sup> However, the

same workers later excluded this possibility experimentally, because the incorporation of deuterium from D<sub>2</sub>O at C1 of *ent*-70 proceeded with *Re* face attack.<sup>[255]</sup> Based on the crystal structure of PR-AS the active site residue Tyr92 was suggested to serve as a general acid in the reprotonation of (–)-1,<sup>[256]</sup> but also this hypothesis was disfavoured by site-directed mutagenesis.<sup>[250]</sup> A more detailed picture was subsequently obtained by the crystal structure of AT-AS, providing evidence that the diphosphate anion is ideally positioned to act as a general acid and base relevant for i) the deprotonation of (S)-B, with the proton taken up by O6, and ii) the reprotonation of the resulting (–)-1 with donation of a different proton from O3 (this process may also be concerted with 1 as a highly transient species, Scheme 26D).<sup>[257]</sup> The results of a site-directed mutagenesis suggest that the thus formed eudesmane cation **H6** is stabilised by W334 of PR-AS or W308 of AT-AS.<sup>[258]</sup> Cationic azanalogues of **H6** have been shown to efficiently inhibit catalysis by PR-AS.<sup>[259,260]</sup>

The sesquiterpene hydrocarbon *ent*-70 is the biosynthetic precursor to PR toxin (**99**),<sup>[261]</sup> a potent mycotoxin that targets transcription and protein biosynthesis with a lethal dose of LD<sub>50</sub> = 5 mg kg<sup>-1</sup> in mice,<sup>[262–264]</sup> and a series of other oxidation products that are likely pathway intermediates.<sup>[265–269]</sup> Surprisingly, despite the potential of mycotoxin biosynthesis *P. roqueforti* is traditionally used for the production of blue cheese, which is explainable by the rapid degradation of **99** under cheese fermentation conditions.<sup>[270]</sup> Biosynthetic hypotheses linking these oxidised metabolites have been investigated by feeding of labelled precursors<sup>[148,269]</sup> and discussed on the grounds of the biosynthetic gene cluster,<sup>[271–273]</sup> but apart from the aristolochene synthase and the poorly characterised eremofortin C oxidase<sup>[274]</sup> for the installation of the aldehyde function in **99** little is known about the enzymes involved in fungal toxin biosynthesis.

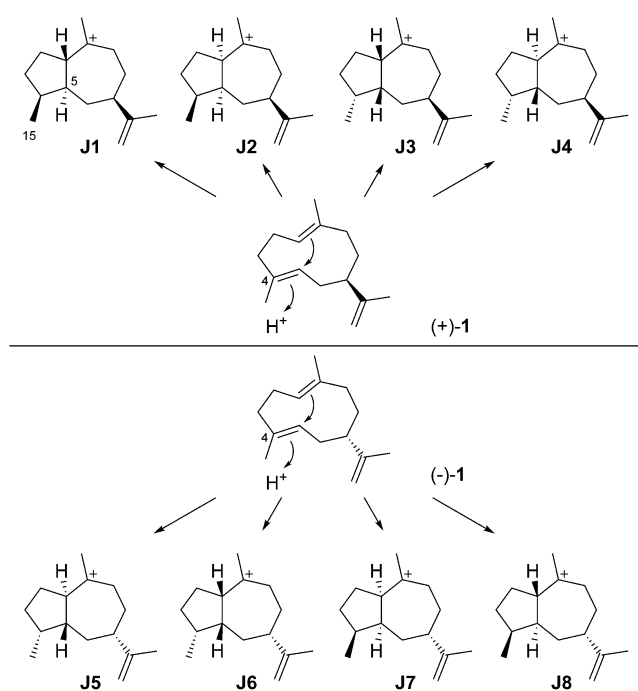
## 5. Guaianes

### 5.1. Guaianes formed by C4 protonation of germacrene A

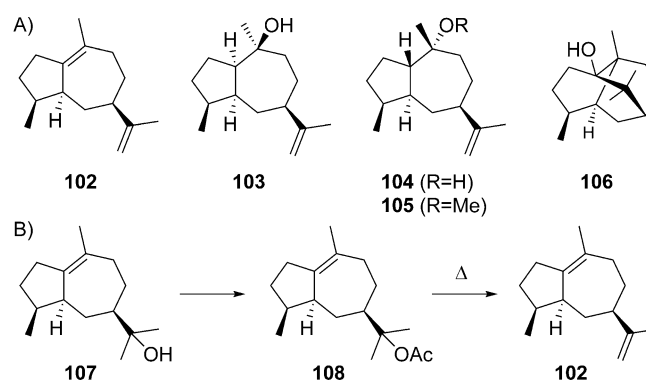
Eight cationic intermediates can be formed from the enantiomers of **1** by protonation at C4 and ring closure (Scheme 27). These cations exhibit four stereogenic centres, leading to a maximum number of 2<sup>4</sup> = 16 possible stereoisomers, but two of the stereogenic centres are not set independently, since the C4/C5 double bond in **1** is *E*-configured and the ring closure proceeds by *anti* addition, that is, Me15 and H5 must be arranged *trans*. Thus, only eight stereoisomers are relevant to this pathway, namely **J1–J4** from (+)-**1**, and their enantiomers **J5–J8** from (–)-**1**.

### 5.2. Guaianes formed from cations **J1** and **J2**

Guaianes from cations **J1** and **J2** include δ-guaiene (**102**) and pogostol (**103**, Scheme 28A). δ-Guaiene is also named α-bulnesene and can in principle be generated by the deprotonation of **J1** or **J2**, while **103** derives from **J2** by *Si* face attack of water. Compound **102** was first isolated from the patchouli oil



Scheme 27. Cyclisations induced by reprotonation of **1** at C4 to **J1–J8**.



Scheme 28. A) Guaianes derived from **J1** and **J2**, initially reported structures of pogostol (**104**) and pogostol methyl ether (**105**), and patchoulol (**106**). B) Synthesis of **102** from bulnesol (**107**).

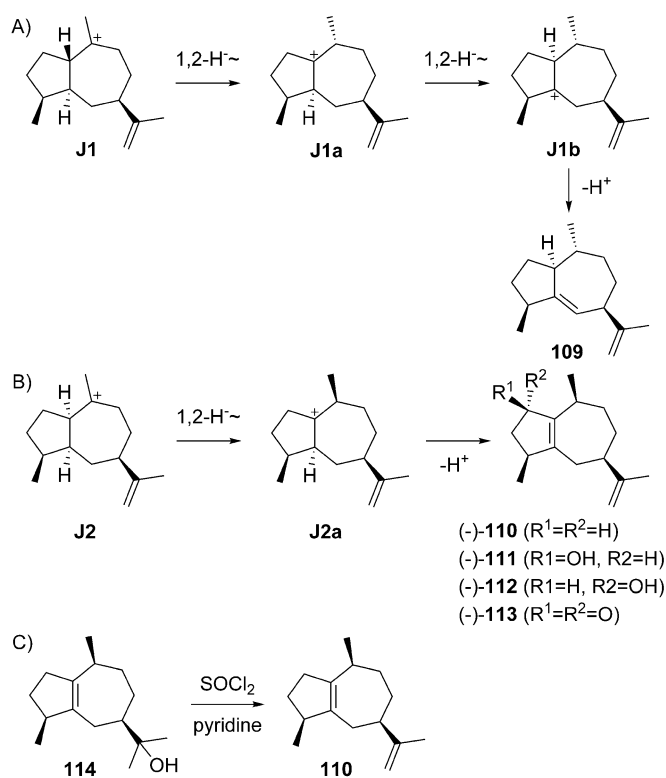
of *Pogostemon cablin* and given its premier name δ-guaiene in 1950. Initially, only the planar structure with insecure positioning of double bonds was determined, with a reported optical rotation close to zero of  $[\alpha]_D = +0.32$ .<sup>[275]</sup> Later, bulnesol (**107**) was chemically converted into **102** by pyrolysis of its acetate **108** (Scheme 28B), leading to a material with an  $[\alpha]_D = 0$ ,<sup>[276]</sup> that was thus inconclusive for assigning the absolute configuration of **102** from the fully established structure of **106**.<sup>[277,278]</sup> Because **102** is accompanied by patchouli alcohol (**106**) in *P. cablin*, it was suggested that both compounds should have coinciding absolute configurations, but at this time for **106** still a wrong structure was assumed (*vide infra*).<sup>[276]</sup> A subsequent stereoselective synthesis from α-cyperone (**16**, Scheme 9) and comparison of the optical rotatory dispersion (o.r.d.) curves of

synthetic and natural **102** finally established its structure.<sup>[279,280]</sup> Compound **102** is known from several other plants<sup>[281–284]</sup> including *Piper fimbriatum*,<sup>[285]</sup> in which it occurs together with **2**. In addition, **102** can be produced by cultured cells from *Aquilaria crassna* and *Aquilaria sinensis*,<sup>[286,287]</sup> resulting in the discovery of the  $\delta$ -guaiene synthase from *A. crassna*.<sup>[288]</sup> Compound **102** is also one of the main products of the  $\alpha$ -guaiene synthase from *V. vinifera*<sup>[289]</sup> and a side product of the patchoulol synthase from *P. cablin*.<sup>[22,290]</sup> The complete <sup>1</sup>H- and <sup>13</sup>C-NMR data of **102** are available.<sup>[286]</sup>

Pogostol (**103**) was first isolated from *P. cablin* ( $[\alpha]_D = -20.2$ ,  $c$  8.7).<sup>[291]</sup> Since then, **103** was reported from various other plant sources<sup>[292–296]</sup> and is known from the fungus *Geniculosporium*.<sup>[297]</sup> A relative configuration was first assigned for pogostol O-methyl ether (**105**) from *Artabotrys stenopetalus*,<sup>[298]</sup> followed by the assignment of the relative configuration of **104** for pogostol by Weyerstahl and co-workers.<sup>[293]</sup> A subsequent synthesis of the reported structures **104** and **105** for pogostol and its methyl ether demonstrated that both assignments were erroneous.<sup>[299]</sup> Amand et al. then gave a correction as **103**.<sup>[295]</sup> Although pogostol is long known and fairly widespread in Nature, the absolute configuration still remains to be determined. For unclear reasons the structure of *ent*-**103** has been assigned to the CAS number of pogostol (21698-41-9), while in fact **103** may be more likely, because this corresponds to the main product **106** of the patchoulol synthase from *P. cablin* that also makes **103** as a side product.<sup>[22]</sup> <sup>1</sup>H- and <sup>13</sup>C-NMR data of **103** are reported in the literature.<sup>[292–295,297]</sup>

The sesquiterpene 1,4-*diepi*- $\gamma$ -gurjunene (**109**, Scheme 29A) was isolated from the sponge *Cymbastela hooperi* ( $[\alpha]_D = +34.6$ ,  $c$  0.11,  $\text{CHCl}_3$ ).<sup>[300]</sup> The formation of this compound can be understood from **J1** by two sequential 1,2-hydride shifts via **J1a** to **J1b** and deprotonation. Since the absolute configuration of **109** has not been determined, it may also be derived from intermediate **J5**. Full <sup>1</sup>H- and <sup>13</sup>C-NMR data have been provided for **109**.<sup>[300]</sup>

$\alpha$ -Guaiene (**110**, Scheme 29B) may instead arise from **J2** by 1,2-hydride migration to **J2a** and deprotonation. It is the universal precursor leading under simple aerial oxidation conditions to many fragrant volatiles of industrial importance such as (*R*)- and (*S*)-rotundols (**111** and **112**) and rotundone (**113**) that exhibit a pleasant peppery or woody aroma.<sup>[301–303]</sup> Compound **110** ( $[\alpha]_D^{19} = -64.5$ ,  $c$  3.584, dioxane) was initially obtained by dehydration of guaiol (**114**, Scheme 29C).<sup>[304]</sup> With the absolute configuration of **114** being specified,<sup>[305]</sup> the full structure of compound **110** was also affirmed. Natural sources of **110** include several plant species<sup>[63,284,285,306–310]</sup> and cell cultures from *Aquilaria crassna* and *A. sinensis*.<sup>[286,287]</sup> A recombinant  $\alpha$ -guaiene synthase has been reported from *V. vinifera*,<sup>[289]</sup> and **110** is also a side product of  $\delta$ -guaiene synthase from *A. crassna*<sup>[288]</sup> and patchoulol synthase from *P. cablin*.<sup>[22,290]</sup> The biosynthesis of **110** is also possible from **K1** (Scheme 32, Section 5.5) by 1,2-hydride shift and deprotonation, but the co-occurrence with **102** in several species,<sup>[284–287,307,308]</sup> whose formation can best be understood from **J1** or **J2**, together with the observation of both compounds in the product profiles of several terpene synthases<sup>[22,288–290]</sup> speaks in favour of a common



**Scheme 29.** Biosynthesis of guaienes from A) **J1** and B) **J2**. C) Chemical correlation of **110** with guaiol (**114**).

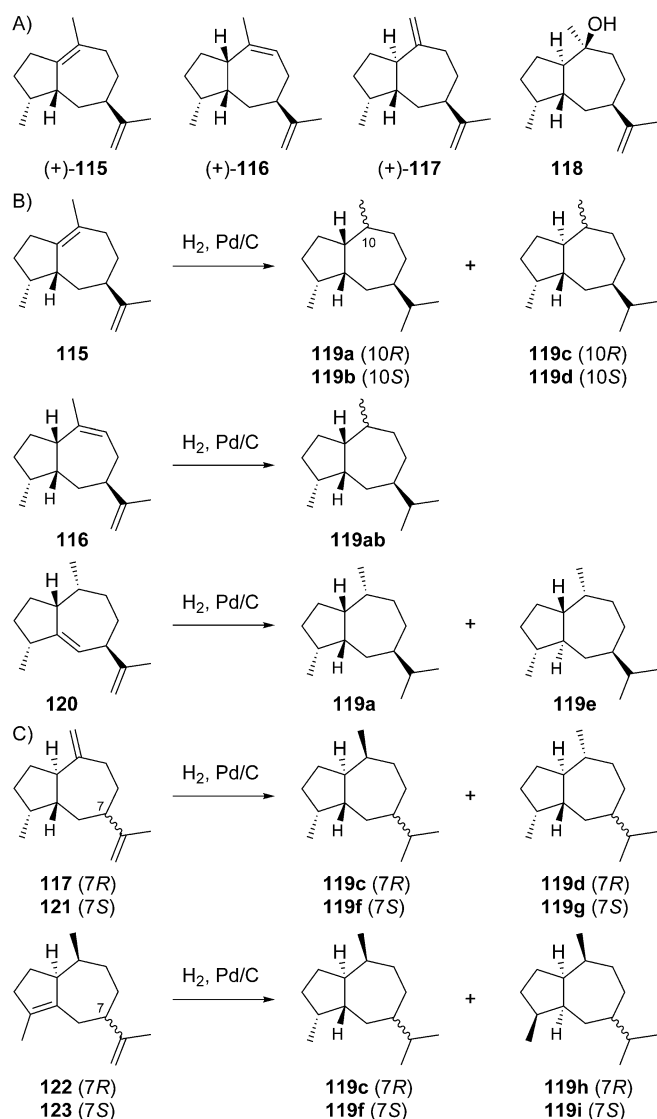
biosynthesis through **J2**. Full <sup>1</sup>H- and <sup>13</sup>C-NMR data of **110** are provided.<sup>[284,308]</sup>

### 5.3. Guaienes formed from cations **J3** and **J4**

Guaienes from **J3** and **J4** include guaia-1(10),11-diene (**115**) that is accessible through both cations by deprotonation, and guaia-9,11-diene (**116**) obtainable by loss of a proton from **J3** (Scheme 30A). Deprotonation of **J4** can lead to guaia-10(14),11-diene (**117**), a compound for which we revise the structure here based on the reason given below, while the attack of water to **J4** can give 4,5-*diepi*-pogostol (**118**). For **118** this discussion is hypothetical, because this compound was only obtained in racemic form by synthesis and is not known as natural product.<sup>[299]</sup>

The hydrocarbons (+)-**115** and (+)-**116** were both isolated only from the fruits of *Peucedanum tauricum*.<sup>[311]</sup> Their co-occurrence in one organism suggests that they may have the same cationic precursor **J3**. The absolute configurations of **115** and **116** were specified by comparison of their hydrogenation products to those obtained from (+)- $\gamma$ -gurjunene (**120**, Scheme 30B),<sup>[312]</sup> leading to one common product (**119a**) from all three materials, as judged by GC analysis using two different chiral stationary phases.

Guaia-10(14),11-diene (**117**) is only known from *Abies koreana*.<sup>[121]</sup> Its absolute configuration was elaborated using the same hydrogenation strategy as for **115** and **116** with chemical correlation to aciphyllene (**122**, Scheme 30C). At the stage of

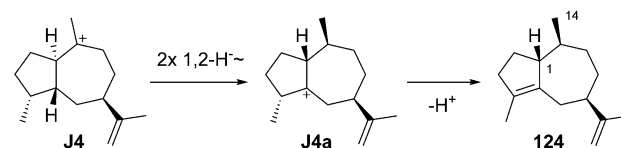


**Scheme 30.** A) Structures of **115**–**118**. Correlations through hydrogenation products B) of **115** and **116** to **120** and C) of revised **117** to aciphyllene (**122**, see text).

this work the structure of **123** with *7S* stereochemistry was assigned for aciphyllene,<sup>[284]</sup> which would have led to the hydrogenation products **119f** and **119i**, and therefore the structure of **121** was concluded for the natural product from *A. koreana* expected to give the hydrogenation products **119f** and **119g**. However, shortly after the structure of aciphyllene underwent a revision to (*7R*)-**122**.<sup>[313]</sup> In conclusion, the truly obtained hydrogenation products from aciphyllene were **119c** and **119h**, with the consequence that the natural product from *A. koreana* must be revised herewith to **117**, expected to give **119c** and **119d**.

The synthetic compound 1-*epi*-aciphyllene (**124**) has been prepared from guaial (**114**),<sup>[314]</sup> but has not been discovered from Nature so far. Indeed, its biosynthesis is not easily understood, as its formation through the **K** series (Scheme 32, Section 5.5) of cations cannot lead to a *cis*-orientation of H1 and Me14. If **124** exists at all as a natural product, two sequential 1,2-hydride migrations from **J4** to **J4a** and deprotonation

could explain its formation (Scheme 31). Full <sup>1</sup>H- and <sup>13</sup>C-NMR data for **124** were reported,<sup>[314]</sup> but unfortunately no optical rotation that would be useful for comparison in case of its future isolation.



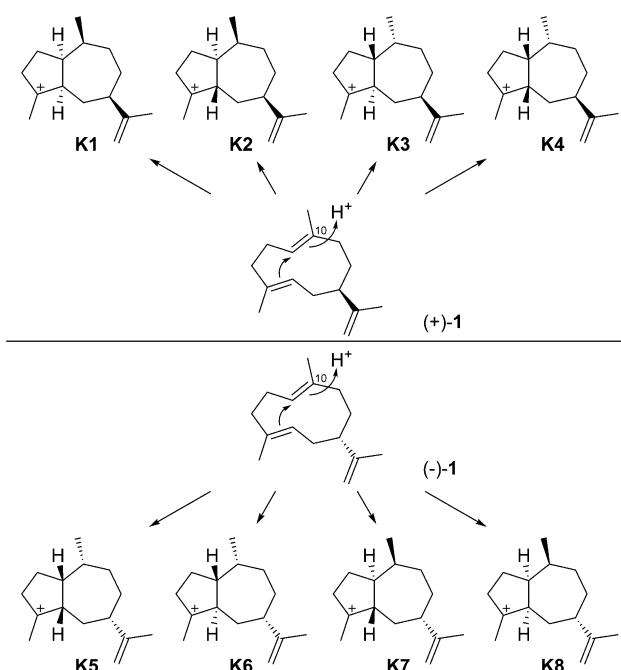
**Scheme 31.** Hypothetical biosynthesis of 1-*epi*-aciphyllene (**124**).

#### 5.4. Guaianes formed from cations **J5**–**J8**

Despite the fact that for **103** the absolute configuration has not been determined and this compound could in principle arise through **J6**, no guaianes from **J5**–**J8** are known. The absolute configuration of 1,4-*diepi*- $\gamma$ -gurjunene (**109**) from *C. hooperi* would be most interesting to know, as sponges may produce the optical antipodes of plant compounds.

#### 5.5. Guaianes formed by C10 protonation of germacrene **A**

Considering the discussion above, there are also only four logical cationic intermediates (**K1**–**K4**) after the cyclisation from (+)-**1** initiated by C10 protonation (Scheme 32). Likewise, (–)-**1** can produce four additional candidates (**K5**–**K8**).

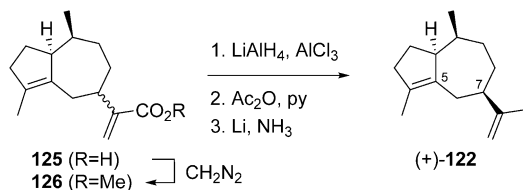


**Scheme 32.** Cyclisations induced by reprotonation of **1** at C10 to **K1**–**K8**.

#### 5.6. Guaianes formed from cations **K1** and **K2**

A deprotonation from C5 of **K1** or **K2** provides aciphyllene (**122**), also named guaia-4,11-diene. Compound **122** was first

isolated from *Lindera glauca* in 1983 ( $[\alpha]_D^{20} = +153.0$ ).<sup>[284]</sup> Its structure was erroneously elucidated by Kubota et al. as that of 7-*epi*-aciphyllylene (**123**) by chemical correlation with aciphyllic acid (**125**, Scheme 33),<sup>[284,315]</sup> a compound that had been reported with 7*S* configuration.<sup>[316]</sup> The structure was later corrected to **122** by synthesis from (+)-dihydrocarvone (**15**).<sup>[313]</sup> Whether this means that also **125** should be revised to have 7*R* configuration or the material had undergone epimerisation at C7 during the transformations into **122** remains unclear at this stage. However, since Kubota and co-workers<sup>[315]</sup> as well as Liu and Yu<sup>[317]</sup> have reported different NMR data for “aciphyllic acid”, in both cases with 7*S* configuration, at least one of these structures must be wrong. Thus it may be likely that the Japanese workers have indeed started their correlation of “aciphyllic acid” to **122** from a material with 7*R* configuration. (+)-Aciphyllylene (**122**) was later also found in *Dumortiera hirsuta*,<sup>[153]</sup> and with undetermined absolute configuration from the essential oil of *Xylopi rubescens*.<sup>[310]</sup> It is also known as a side product of the recombinant patchoulol synthase from *Pogostemon cablin*,<sup>[290]</sup> a multi-product terpene synthase for which all products retain the (7*R*) stereochemistry introduced in the intermediate (+)-1 and thus further supporting the structural reassignment for **122**. Moreover, total syntheses from (*R*)-limonene by Srikrishna et al.<sup>[318]</sup> and from guaial (**114**) by Huang et al.<sup>[314]</sup> were conducted. The <sup>1</sup>H- and <sup>13</sup>C-NMR data of **122** have been published.<sup>[153,284]</sup>

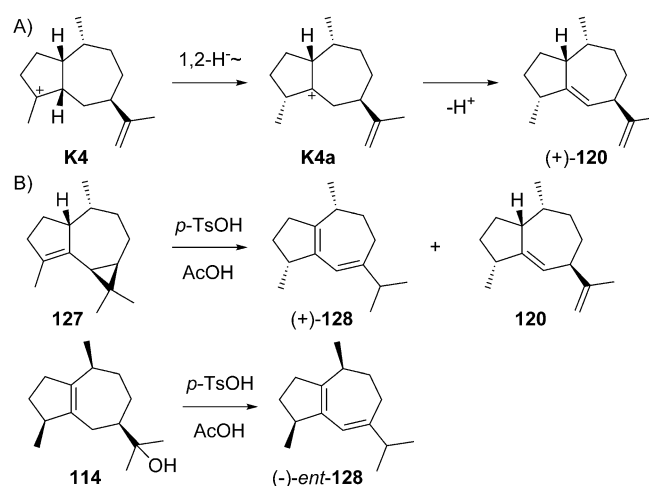


Scheme 33. Chemical correlation of “aciphyllic acid” to **122** (corrected structure).

### 5.7. Guaianes formed from cations K3 and K4

One of the most important sesquiterpenes derived from the K series is (+)- $\gamma$ -gurjunene (**120**). Its formation can be understood from **K4** by 1,2-hydride shift to **K4a** and deprotonation (Scheme 34A). This component was first discovered from the gurjun balsams of several species of *Dipterocarpus* ( $[\alpha]_D = +147$ ,  $\text{CHCl}_3$ ).<sup>[314,319]</sup> Its absolute configuration was illuminated by correlation with  $\alpha$ -gurjunene (**127**) and guaial (**114**, Scheme 34B).<sup>[312]</sup> While treatment of **127** with acid gave the isomerisation products (+)-**128** and **120** identical to natural (+)- $\gamma$ -gurjunene, the isomerisation of **114** produced (–)-*ent*-**128**. Compound **120** was also isolated from *Persea gamblei*.<sup>[320]</sup> Complete <sup>1</sup>H- and <sup>13</sup>C-NMR data have been published.<sup>[300,319,321]</sup>

Compound (–)-*ent*-**123** (Figure 2) is only known as a synthetic material ( $[\alpha]_D^{24} = -13.2$ ,  $c$  0.35,  $\text{CHCl}_3$ ) and could, as a hypothetical natural product, arise from **K3** or **K4** by deprotonation. It is wrongly presented in the synthesis paper that corrects the structure of (+)-aciphyllylene (**122**) as the assigned structure of this natural product (**123**, Scheme 30), while it rep-



Scheme 34. A) Biosynthesis of **120**. B) Correlation of **120** with **127** and **114**.

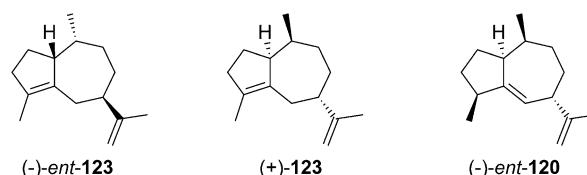


Figure 2. Structures of synthetic compounds *ent*-**123**, **123** and *ent*-**120**.

resents in fact its enantiomer. Full <sup>1</sup>H- and <sup>13</sup>C-NMR data are available.<sup>[313]</sup>

### 5.8. Guaianes formed from cations K5–K8

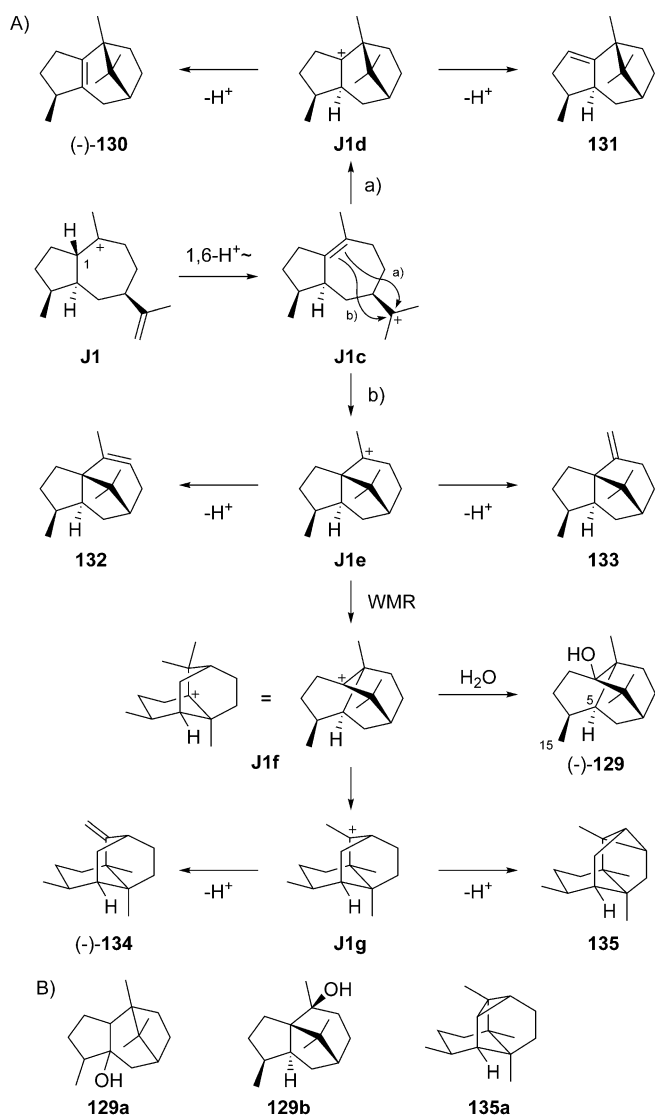
Natural products from the cations **K5–K8** are unknown. Synthetic compounds (Figure 2) include (+)-7-*epi*-aciphyllylene (**123**) obtained from (*R*)-limonene ( $[\alpha]_D^{27} = +13.5$ ,  $c$  1.3,  $\text{CHCl}_3$ ).<sup>[318]</sup> and (–)- $\gamma$ -gurjunene (*ent*-**120**) made accessible through an enantioselective Morita-Baylis–Hillman reaction using an enantiopure phosphine catalyst ( $[\alpha]_D^{20} = -121.1$ ,  $c$  0.1  $\text{CHCl}_3$ ).<sup>[322]</sup> For both compounds full NMR data were provided.<sup>[318,322]</sup>

## 6. Cyclised and Rearranged Guaianes

Further cyclisations eventually with skeletal rearrangements are important for two groups of compounds originating from **J1** and **J3**, while no examples from the other cations of the **J** series or from cations of the **K** series are known.

### 6.1. Compounds from J1

Compounds from **J1** include patchouli alcohol (**129**), the patchoulenes **130–133** and seychellenes **134** and **135** (Scheme 35A). The common biosynthesis of these compounds can be understood from **J1** by a long range proton shift from C1 into the isopropenyl group to **J1c**, followed by cyclisation to **J1d** (path a) and deprotonation to  $\beta$ -patchoulene (**130**) and



**Scheme 35.** A) Biosynthesis of cyclised and rearranged guaianes from J1. B) Initially assigned structures for (-)-patchouli alcohol (129a and 129b) and cycloseychellene (135a).

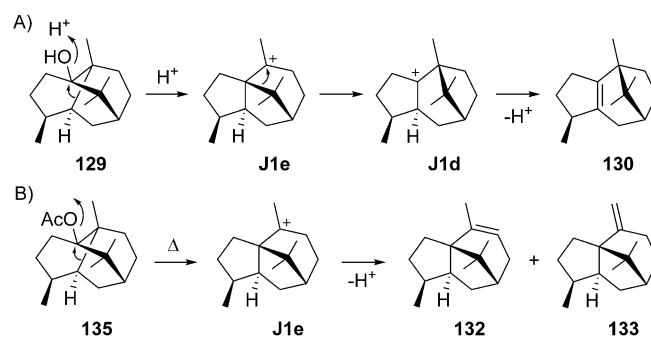
$\delta$ -patchoulene (131). An alternative cyclisation from J1c to J1d (path b) and deprotonation yields  $\alpha$ -patchoulene (132) and  $\gamma$ -patchoulene (133). A Wagner–Meerwein rearrangement of J1e to J1f gives access to patchouli alcohol (129) by attack of water, while a methyl group migration to J1g and deprotonation results in seychellene (134) or cycloseychellene (135). This pathway is in agreement with feeding experiments using radioactively labelled (4R)-[2-<sup>14</sup>C,4-<sup>3</sup>H]mevalonic acid,<sup>[323,324]</sup> and with deuterium incorporation from (2-<sup>2</sup>H)FPP at C5 of 129 and several side products from patchoulol synthase,<sup>[22,290]</sup> while a reported additional deuteration at C15 is difficult to understand.

Patchouli alcohol or patchoulol (-)-129 was first isolated as the main constituent from patchouli oil (*P. cablin*) in 1869.<sup>[325]</sup> The oil is one of the most important industrial fragrances that is widely used in perfumery and cosmetics products. Its planar structure was described more than 80 years later as that of

129a (Scheme 35 B).<sup>[326]</sup> A structural revision based on chemical transformations and a synthesis from (+)-camphor through 132 resulted in the assignment of structure 129b.<sup>[327–329]</sup> However, a subsequent X-ray analysis of the chromic acid diester surprisingly led to the structure of 129,<sup>[330]</sup> suggesting that during the synthesis of this compound from 132 a similar skeletal rearrangement as in the biosynthesis must have taken place. A later synthesis from (*R*)-carvone (*ent*-38) resulted in (-)-129 ( $[\alpha]_D^{25} = -121.3$ ,  $c$  2.3, CHCl<sub>3</sub>).<sup>[331]</sup> Compound (-)-129 was also isolated from plants of the genera *Valeriana*<sup>[332–334]</sup> and *Nardostachys*<sup>[335,336]</sup> The complete <sup>13</sup>C NMR data of 129 are available.<sup>[290,333,337]</sup>

The patchoulenes 130–133 and seychellenes 134 and 135 have been reported to co-occur with 129 in several species,<sup>[307,332,334–336,338,339]</sup> and also many of these compounds are observed as products of the patchoulol synthase,<sup>[22,290]</sup> supporting their common biosynthesis through shared intermediates (Scheme 35 A) and corresponding absolute configurations. Formally, the absolute configuration of 130 ( $[\alpha]_D^{30} = -42.6$ ,  $c$  10.51, CHCl<sub>3</sub>) was specified by chemical correlation with patchouli alcohol through acid treatment, at a time when 129b was believed to be the correct structure of this sesquiterpene alcohol. Pyrolysis of patchoulyl acetate (135) yielded a mixture of 132 and 133, and dehydration with POCl<sub>3</sub> resulted in a mixture of mainly 132 with 130 and 133.<sup>[328]</sup> A reinterpretation of the results from these experiments included a Wagner–Meerwein rearrangement (Scheme 36).<sup>[340]</sup> Compound 131 was first obtained by the acid-catalysed transformation of 129<sup>[341]</sup> and later isolated from patchouli oil.<sup>[342]</sup> The complete <sup>1</sup>H and <sup>13</sup>C NMR data of 130 are available,<sup>[308]</sup> while those of 131–133 are lacking.

Seychellene (134, Scheme 35 A),  $[\alpha]_D = -72$  ( $c$  0.4, CHCl<sub>3</sub>),<sup>[343]</sup> was first found in patchouli oil (“hydrocarbon G”),<sup>[307]</sup> followed by structure elucidation through chemical degradation.<sup>[340,343]</sup> A total synthesis of (-)-134 from (*R*)-carvone (*ent*-81) confirmed its absolute configuration.<sup>[344]</sup> Cycloseychellene (135) was reported to possess the structure of 135a (Scheme 35 B) when it was first isolated from *P. cablin* in 1973.<sup>[339]</sup> In 1981, Welch et al. synthesised ( $\pm$ )-135a and found that the spectral and chromatographic properties of the synthetic hydrocarbon differed significantly from those of the natural product.<sup>[345]</sup> A re-examination of the NMR spectra of cycloseychellene indicating that its



**Scheme 36.** A) Acid promoted conversion of 129 into 130. B) Pyrolysis of patchoulyl acetate (135) to patchoulenes 132 and 133.

structure should be corrected to that of **135**.<sup>[346]</sup> The <sup>1</sup>H- and <sup>13</sup>C-NMR data of **134** are available from the literature.<sup>[308,344]</sup>

## 6.2. Compounds from J3

The biosynthesis of rotundene (**136**), isotrundene (**137**) and cyperene (**138**) can be understood from **J3** (Scheme 37A). Its cyclisation to **J3a** (path a) and deprotonation yields **136** and **137**, while a 1,2-hydride shift to **J3b** (path b) followed by a 1,5-proton shift to **J3c**, cyclisation to **J3d** and deprotonation result in **138**. This common biosynthetic pathway nicely explains the co-occurrence of **136**–**138** in *Cyperus rotundus*.<sup>[347]</sup> Compound **136** ( $[\alpha]_D^{20} = -16.3$ ) was first reported from *C. rotundus* and *C. scariosus*,<sup>[348]</sup> and later also from *C. alopecuroides*,<sup>[349]</sup> but at this stage only with the planar structure. (–)-Isotrundene (**137**) was isolated from *C. rotundus* whose relative configuration was determined by NOESY.<sup>[347]</sup> This allowed to demonstrate that **136** has the same skeleton by conversion into rotundol (**139**) through oxymercuration and dehydration with POCl<sub>3</sub> (Scheme 37B). The absolute configuration of **136**, and thus also of **137**, was determined by ozonolysis to **140**, decar-

boxylation to a mixture of epimers **141ab**, Wittig methylenation to **142ab** and catalytic hydrogenation to **119ab** (Scheme 37C). One of these hydrocarbons was identical to **119a** obtained by hydrogenation of **120** (Scheme 30C). Complete <sup>1</sup>H- and <sup>13</sup>C-NMR data for **137** have been reported,<sup>[347]</sup> but are lacking for **136**.

The sesquiterpene **138** ( $[\alpha]_D^{20} = -20.0$ , neat), was first isolated from *Cyperus rotundus*.<sup>[350,351]</sup> Its absolute configuration was resolved by the chemical correlation through its hydrogenation product that was identical to a material derived from **129** by dehydration with POCl<sub>3</sub> and hydrogenation.<sup>[352,353]</sup> The (–)-enantiomer of **138** was later isolated from several other plants.<sup>[177,349,354–367]</sup> Full <sup>1</sup>H- and <sup>13</sup>C-NMR data in CDCl<sub>3</sub> and C<sub>6</sub>D<sub>6</sub> have been reported.<sup>[367,368]</sup>

## 7. Conclusions

Germacrene A shows a unique and interesting chemistry mainly characterised by its reactivity towards acid-catalysed cyclisations and its thermal lability in a Cope rearrangement to β-elemene. Similar observations have been made for other germacrenes,<sup>[369]</sup> suggesting that the high ring strain associated with the 10-membered ring in these systems may be a strong driving force for the observed reactions leading to much less strained compounds with 6-membered rings. The reactivity built up by the ring strain is also used in enzymatic reactions towards sesquiterpenes for which germacrene A serves as an important intermediate. In enzyme reactions not only the formation of 6–6 bicyclic compounds, but also of 5–7 bicyclic derivatives can be achieved, and for both cases follow-up chemistry by skeletal rearrangements can further increase the structural variability. Subsequent steps include oxidative and other modifications after terpene cyclisation, leading to numerous derivatives for each compound presented in this review, which further underlines the central importance of germacrene A in sesquiterpene biosynthesis.

## Acknowledgements

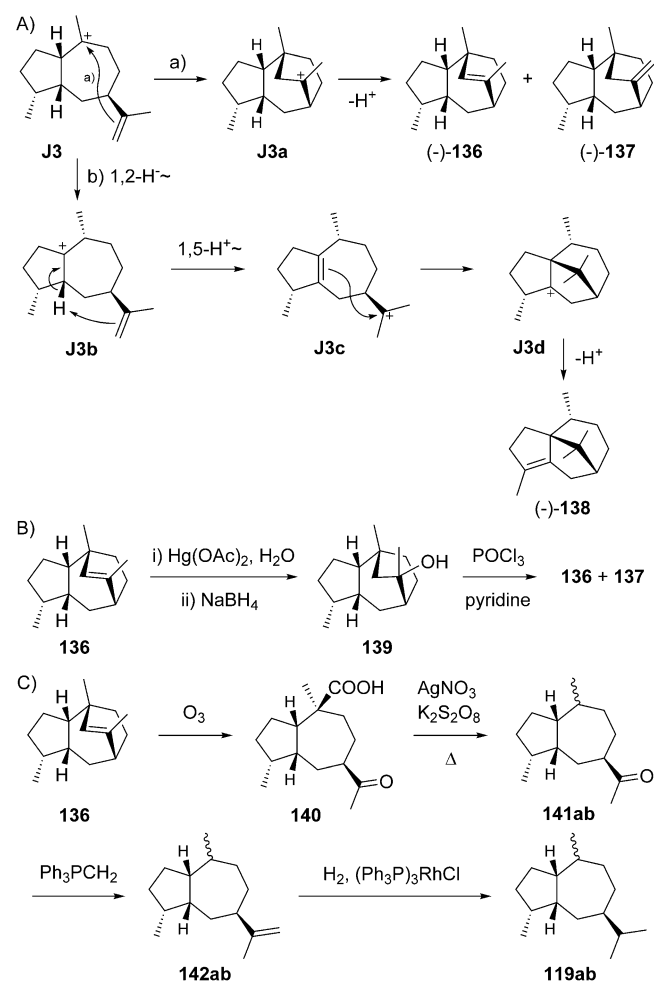
Open access funding enabled and organized by Projekt DEAL.

## Conflict of interest

The authors declare no conflict of interest.

**Keywords:** biosynthesis · configuration determination · enzyme catalysis · germacrene A · sesquiterpenes

- [1] J. S. Dickschat, *Nat. Prod. Rep.* **2016**, *33*, 87–110.
- [2] M. B. Quin, C. M. Flynn, C. Schmidt-Dannert, *Nat. Prod. Rep.* **2014**, *31*, 1449–1473.
- [3] D. W. Christianson, *Chem. Rev.* **2017**, *117*, 11570–11648.
- [4] F. Chen, D. Tholl, J. Bohlmann, E. Pichersky, *Plant J.* **2011**, *66*, 212–229.
- [5] X. Chen, T. G. Köllner, Q. Jia, A. Norris, B. Santhanam, P. Rabe, J. S. Dickschat, G. Shaulsky, J. Gershenzon, F. Chen, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 12132–12137.



**Scheme 37.** A) Biosynthesis of **136**–**138** from **J3**. Chemical correlation of **B**) **136** to **137**, and C) **136** to **119a**, the hydrogenation product of (+)-γ-gurjunene.

- [6] P. Rabe, J. Rinkel, B. Nubbemeyer, T. G. Köllner, F. Chen, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2016**, *55*, 15420–15423; *Angew. Chem.* **2016**, *128*, 15646–15649.
- [7] A. J. Weinheimer, W. W. Youngblood, P. H. Washecheck, T. K. B. Karns, L. S. Ciereszko, *Tetrahedron Lett.* **1970**, *11*, 497–500.
- [8] V. Sýkora, V. Herout, F. Sorm, *Coll. Czech. Chem. Commun.* **1956**, *21*, 267–268.
- [9] A. D. Wagh, S. K. Paknikar, S. C. Bhattacharyya, *Tetrahedron* **1964**, *20*, 2647–2654.
- [10] R. W. Dunlop, R. J. Wells, *Aust. J. Chem.* **1979**, *32*, 1345–1351.
- [11] W. S. Bowers, C. Nishino, M. E. Montgomery, L. R. Nault, M. W. Nielson, *Science* **1977**, *196*, 680.
- [12] C. Nishino, W. S. Bowers, M. E. Montgomery, L. R. Nault, M. W. Nielson, *J. Chem. Ecol.* **1977**, *3*, 349–357.
- [13] J. A. Faraldos, S. Wu, J. Chappell, R. M. Coates, *Tetrahedron* **2007**, *63*, 7733–7742.
- [14] W. A. König, A. Rieck, I. Hardt, B. Gehrke, K.-H. Kubeczka, H. Muhle, *J. High Resolut. Chromatogr.* **1994**, *17*, 315–320.
- [15] J.-W. de Kraker, M. C. R. Franssen, A. de Groot, W. A. König, H. J. Bouwmeester, *Plant Physiol.* **1998**, *117*, 1381–1392.
- [16] H. J. Bouwmeester, J. Kodde, F. W. A. Verstappen, I. G. Altug, J.-W. de Kraker, T. E. Wallaart, *Plant Physiol.* **2002**, *129*, 134–144.
- [17] S. Irmisch, S. T. Krause, G. Kunert, J. Gershenzon, J. Degenhardt, T. G. Köllner, *BMC Plant Biol.* **2012**, *12*, 84.
- [18] J.-W. de Kraker, M. C. R. Franssen, A. de Groot, T. Shibata, H. J. Bouwmeester, *Phytochemistry* **2001**, *58*, 481–487.
- [19] M. Majidi, Q. Liu, G. Karimzadeh, M. A. Malboobi, J. Beekwilder, K. Cankar, R. de Vos, S. Todorovic, A. Simonovic, H. J. Bouwmeester, *Phytochemistry* **2011**, *72*, 1739–1750.
- [20] A. M. Ramirez, N. Saillard, T. Yang, M. C. R. Franssen, H. J. Bouwmeester, M. A. Jongsma, *PLoS ONE* **2013**, *8*, e65030.
- [21] B. Menin, C. Comino, E. Portis, A. Moglia, K. Cankar, H. J. Bouwmeester, S. Lanteri, J. Beekwilder, *Plant Sci.* **2012**, *190*, 1–8.
- [22] F. Deguerry, L. Pastore, S. Wu, A. Clark, J. Chappell, M. Schalk, *Arch. Biochem. Biophys.* **2006**, *454*, 123–136.
- [23] M.-Y. Kim, Y.-J. Chang, M.-H. Bang, N.-I. Baek, J. Jin, C.-H. Lee, S.-U. Kim, *J. Plant Biol.* **2005**, *48*, 178–186.
- [24] M. H. Bennett, J. W. Mansfield, M. J. Lewis, M. H. Beale, *Phytochemistry* **2002**, *60*, 255–261.
- [25] C. M. Berteau, A. Voster, F. W. A. Verstappen, M. Maffei, J. Beekwilder, H. J. Bouwmeester, *Arch. Biochem. Biophys.* **2006**, *448*, 3–12.
- [26] D. M. Martin, S. Aubourg, M. B. Schouwey, L. Daviet, M. Schalk, O. Toub, S. T. Lund, J. Bohlmann, *BMC Plant Biol.* **2010**, *10*, 226.
- [27] J. C. Göpfert, G. MacNevin, D.-K. Ro, O. Spring, *BMC Plant Biol.* **2009**, *9*, 86.
- [28] J. Göpfert, A.-K. Bülow, O. Spring, *Nat. Prod. Commun.* **2010**, *5*, 709–715.
- [29] T. Shimada, T. Endo, A. Rodriguez, H. Fujii, M. Nakano, A. Sugiyama, T. Shimizu, L. Pena, M. Omura, *Sci. Hortic.* **2012**, *145*, 102–108.
- [30] T. Benabdelkader, Y. Guitton, B. Pasquier, J. L. Magnard, F. Jullien, A. Kameli, L. Legendre, *Physiol. Plant.* **2015**, *153*, 43–57.
- [31] L. Pazouki, H. R. Memari, A. Kännaste, R. Bichele, Ü. Niinemets, *Front. Plant Sci.* **2015**, *6*, 111.
- [32] T.-D. Nguyen, J. A. Faraldos, M. Vardakou, M. Salmon, P. E. O'Maille, D.-K. Ro, *Biochem. Biophys. Res. Commun.* **2016**, *479*, 622–627.
- [33] J. Rinkel, J. S. Dickschat, *Org. Lett.* **2019**, *21*, 2426–2429.
- [34] P. Rabe, T. Schmitz, J. S. Dickschat, *Beilstein J. Org. Chem.* **2016**, *12*, 1839–1850.
- [35] L. Ding, H. Goerls, K. Dornblut, W. Lin, A. Maier, H.-H. Fiebig, C. Hertweck, *J. Nat. Prod.* **2015**, *78*, 2963–2967.
- [36] M. L. Wise, M. Urbansky, G. L. Helms, R. M. Coates, R. Croteau, *J. Am. Chem. Soc.* **2002**, *124*, 8546–8547.
- [37] J. Rinkel, P. Rabe, L. zur Horst, J. S. Dickschat, *Beilstein J. Org. Chem.* **2016**, *12*, 2317–2324.
- [38] J. Piel, *Nat. Prod. Rep.* **2009**, *26*, 338–362.
- [39] A. M. Adio, C. Paul, H. Tesso, P. Kloth, W. A. König, *Tetrahedron: Asymmetry* **2004**, *15*, 1631–1635.
- [40] P. Rabe, L. Barra, J. Rinkel, R. Riclea, C. A. Citron, T. A. Klapschinski, A. Janusko, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2015**, *54*, 13448–13451; *Angew. Chem.* **2015**, *127*, 13649–13653.
- [41] L. Lauterbach, J. Rinkel, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2018**, *57*, 8280–8283; *Angew. Chem.* **2018**, *130*, 8412–8415.
- [42] R. B. Woodward, R. Hoffmann, *J. Am. Chem. Soc.* **1965**, *87*, 395–397.
- [43] K. N. Houk, Y.-T. Lin, F. K. Brown, *J. Am. Chem. Soc.* **1986**, *108*, 554–556.
- [44] D. K. Lewis, B. Brandt, L. Crockford, D. A. Glenar, G. Rauscher, J. Rodriguez, J. E. Baldwin, *J. Am. Chem. Soc.* **1993**, *115*, 11728–11734.
- [45] W. von E. Doering, W. R. Roth, *Tetrahedron* **1962**, *18*, 67–74.
- [46] F. W. Semmler, F. Risse, *Chem. Ber.* **1912**, *45*, 3301–3307.
- [47] L. Ruzicka, M. Stoll, *Helv. Chim. Acta* **1922**, *5*, 923–936.
- [48] L. Ruzicka, A. H. Wind, D. R. Koolhaas, *Helv. Chim. Acta* **1931**, *14*, 1132–1151.
- [49] B. S. Tyagi, B. B. Ghatge, S. C. Bhattacharyya, *Tetrahedron* **1963**, *19*, 1189–1193.
- [50] T. C. Jain, R. J. Striha, *Phytochemistry* **1976**, *15*, 847–848.
- [51] L. Ruzicka, D. R. Koolhaas, A. H. Wind, *Helv. Chim. Acta* **1931**, *14*, 1171–1178.
- [52] D. H. R. Barton, *J. Chem. Soc.* **1953**, 1027–1040.
- [53] D. H. R. Barton, *Chem. Ind.* **1953**, 664.
- [54] R. B. Woodward, F. Sondheimer, D. Taub, K. Heusler, W. M. MacLamore, *J. Am. Chem. Soc.* **1952**, *74*, 4223–4251.
- [55] B. Riniker, J. Kalvoda, D. Arigoni, A. Fürst, O. Jeger, A. M. Gold, R. B. Woodward, *J. Am. Chem. Soc.* **1954**, *76*, 313–314.
- [56] Y. R. Naves, P. Ardizio, *Bull. Soc. Chim. Fr.* **1956**, 292–297.
- [57] K. Yoshihara, Y. Hirose, *Bull. Chem. Soc. Jpn.* **1978**, *51*, 3395–3396.
- [58] T. Okude, S. Hayashi, *Bull. Chem. Soc. Jpn.* **1970**, *43*, 2984–2985.
- [59] G. L. Chetty, G. S. Krishna Rao, S. Dev, D. K. Banerjee, *Tetrahedron* **1966**, *22*, 2311–2318.
- [60] N. H. Andersen, B. Shunk, C. R. Costin, *Experientia* **1973**, *29*, 645–647.
- [61] B. Maurer, A. Grieder, *Helv. Chim. Acta* **1977**, *60*, 2177–2190.
- [62] Y. Naya, F. Miyamoto, T. Takemoto, *Experientia* **1978**, *34*, 984–986.
- [63] L. A. Smedman, E. Zavarin, R. Teranishi, *Phytochemistry* **1969**, *8*, 1457–1470.
- [64] A. S. Bawdekar, G. R. Kelkar, *Tetrahedron* **1965**, *21*, 1521–1528.
- [65] C. Ganter, B. Keller-Wojtkiewicz, *Helv. Chim. Acta* **1971**, *54*, 183–206.
- [66] H. Itokawa, H. Morita, T. Kobayashi, K. Watanabe, Y. Itaka, *Chem. Pharm. Bull.* **1987**, *35*, 2860–2868.
- [67] H. J. Williams, I. Sattler, G. Moyna, A. I. Scott, A. A. Bell, S. B. Vinson, *Phytochemistry* **1995**, *40*, 1633–1636.
- [68] Y. Ohta, Y. Hirose, *Agr. Biol. Chem.* **1966**, *30*, 1196–1201.
- [69] M. G. Moshonas, P. E. Shaw, M. K. Veldhuis, *J. Agric. Food Chem.* **1972**, *20*, 751–752.
- [70] F. Bohlmann, L. Fiedler, *Chem. Ber.* **1978**, *111*, 408–410.
- [71] O. N. Devgan, M. M. Bokadia, *Curr. Sci.* **1967**, *36*, 205–206.
- [72] P. Raharivelomanana, J.-P. Bianchini, A. R. P. Ramanoelina, J. R. E. Ra-soarahona, R. Faure, A. Cambon, *Phytochemistry* **1998**, *47*, 1085–1088.
- [73] R. M. Smith, S. Siwatibau, *Phytochemistry* **1975**, *14*, 2013–2015.
- [74] J. H. G. Lago, C. B. Brochini, N. F. Roque, *Phytochemistry* **2002**, *60*, 333–338.
- [75] S. S. Martin, J. H. Langenheim, E. Zavarin, *Phytochemistry* **1972**, *11*, 3049–3051.
- [76] H. Hendriks, T. M. Malingre, S. Batterman, R. Bos, *Phytochemistry* **1975**, *14*, 814–815.
- [77] Zhaoming Xiong, J. Yang, Y. Li, *Tetrahedron: Asymmetry* **1996**, *7*, 2607–2612.
- [78] Z. Zhang, Z. Xiong, G. Zheng, Y. Li, *Tetrahedron: Asymmetry* **2001**, *12*, 2137–2141.
- [79] T. Toda, Y. S. Cheng, T. Nozoe, *Chem. Pharm. Bull.* **1967**, *15*, 903–905.
- [80] S. Melching, W. A. König, *Phytochemistry* **1999**, *51*, 517–523.
- [81] F. Bohlmann, N. Ates, J. Jakupovic, *Phytochemistry* **1983**, *22*, 1159–1162.
- [82] B. M. Lawrence, K. M. Weaver, *Planta Med.* **1974**, *25*, 385–388.
- [83] J. H. G. Lago, C. B. Brochini, N. F. Roque, *Phytochemistry* **2002**, *60*, 333–338.
- [84] H. Iwabuchi, M. Yoshikura, W. Kamisako, *Chem. Pharm. Bull.* **1989**, *37*, 509–510.
- [85] R. E. Corbett, R. A. J. Smith, *Tetrahedron Lett.* **1967**, *8*, 1009–1012.
- [86] G. L. Chetty, V. B. Zalkow, L. H. Zalkow, *Tetrahedron Lett.* **1968**, *9*, 3223–3225.
- [87] R. P. W. Kesselmann, J. B. P. A. Wijnberg, A. J. Minnaard, R. E. Walinga, A. de Groot, *J. Org. Chem.* **1991**, *56*, 7237–7244.



- [88] R. P. W. Kesselmans, J. B. P. A. Wijnberg, A. de Groot, T. A. van Beek, *J. Essent. Oil Res.* **1992**, *4*, 201–217.
- [89] J. A. Marco, J. F. Sanz, A. Yuste, M. Carda, J. Jakupovic, *Phytochemistry* **1991**, *30*, 3661–3668.
- [90] C. S. Chanotiya, S. S. Sammal, C. S. Mathela, *Ind. J. Chem. B* **2005**, *44B*, 1922–1926.
- [91] A. F. Barrero, J. F. Sanchez, J. E. Oltra, J. Altarejos, N. Ferrol, A. Barragan, *Phytochemistry* **1991**, *30*, 1551–1554.
- [92] C.-Y. Duh, S.-K. Wang, Y.-L. Weng, M. Y. Chiang, C.-F. Dai, *J. Nat. Prod.* **1999**, *62*, 1518–1521.
- [93] J.-H. Sheu, K.-C. Chang, P.-J. Sung, C.-Y. Duh, Y.-C. Shen, *J. Chin. Chem. Soc.* **1999**, *46*, 253–257.
- [94] J. Jakupovic, C. Zdero, M. Grenz, F. Tschritzis, L. Lehmann, S. M. Hashemi-Nejad, F. Bohlmann, *Phytochemistry* **1989**, *28*, 1119–1131.
- [95] P. Weyerstahl, H. Marschall-Weyerstahl, C. Christiansen, *Flavour Fragr. J.* **1989**, *4*, 93–98.
- [96] P. Weyerstahl, H. Marschall, *Flavour Fragr. J.* **1992**, *7*, 73–77.
- [97] C. Botega, F. M. Pagliosa, V. da S. Bolzani, M. Yoshida, O. R. Gottlieb, *Phytochemistry* **1993**, *32*, 1331–1333.
- [98] J.-M. Oger, P. Richomme, H. Guinaudeau, J. P. Bouchara, A. Fournet, *J. Essent. Oil Res.* **1994**, *6*, 493–497.
- [99] A. I. Suárez, Z. Blanco, F. delle Monache, R. S. Compagnone, F. Arvelo, *Nat. Prod. Res.* **2004**, *18*, 421–426.
- [100] P. E. Oliveira da Cruz, E. V. Costa, V. R. de Souza Moraes, P. C. de Lima Nogueira, M. E. Vendramin, A. Barison, A. G. Ferreira, A. P. do Nascimento Prata, *Biochem. Syst. Ecol.* **2011**, *39*, 872–875.
- [101] J.-L. Gu, Z.-J. Li, H.-X. Zhang, Z.-Z. Du, *Chem. Biodiversity* **2014**, *11*, 1398–1405.
- [102] X.-L. Wu, X.-J. Xiong, W.-Q. Lu, H. Huang, Y.-H. Shen, Z.-J. Wu, W.-S. Chen, *Molecules* **2016**, *21*, 1031.
- [103] V. B. Zalkow, A. M. Shaligram, L. H. Zalkow, *Chem. Ind.* **1964**, 194–195.
- [104] L. H. Zalkow, J. T. Baxter, R. J. McClure, M. M. Gordon, *J. Nat. Prod.* **1980**, *43*, 598–608.
- [105] G. D. Prestwich, M. S. Collins, *Biochem. Syst. Ecol.* **1981**, *9*, 83–88.
- [106] Y. Naya, G. D. Prestwich, S. G. Spanton, *Tetrahedron Lett.* **1982**, *23*, 3047–3050.
- [107] F. Bohlmann, C. Zdero, M. Ahmed, *Phytochemistry* **1982**, *21*, 1679–1691.
- [108] J. Jakupovic, T. V. Chau-Thi, U. Warning, F. Bohlmann, H. Greger, *Phytochemistry* **1986**, *25*, 1663–1667.
- [109] E. Klein, W. Rojahn, *Tetrahedron Lett.* **1970**, *11*, 279–282.
- [110] J. W. Huffman, C. A. Miller, A. R. Pinder, *J. Org. Chem.* **1976**, *41*, 3705–3710.
- [111] L. H. Zalkow, M. Smith, G. L. Chetty, A. W. Chaligram, P. Ingwalson, *J. Org. Chem.* **1976**, *41*, 3710–3714.
- [112] H. Sulser, J. R. Scherer, K. L. Stevens, *J. Org. Chem.* **1971**, *36*, 2422–2426.
- [113] J. W. Huffman, L. H. Zalkow, *Tetrahedron Lett.* **1973**, *14*, 751–754.
- [114] A. San Feliciano, M. Medarde, B. del Rey, J. M. M. del Corral, A. F. Barrero, *Phytochemistry* **1990**, *29*, 3207–3211.
- [115] R. Baker, H. R. Coles, M. Edwards, D. A. Evans, P. E. Howse, S. Walmsley, *J. Chem. Ecol.* **1981**, *7*, 135–145.
- [116] E. Everaerts, Y. Roisin, J.-L. le Quere, O. Bonnard, J. M. Pasteels, *J. Chem. Ecol.* **1993**, *19*, 2865–2879.
- [117] L. H. Zalkow, V. B. Zalkow, D. R. Brannon, *Chem. Ind.* **1963**, 38.
- [118] P. Weyerstahl, V. K. Kaul, M. Weihrauch, H. Marschall-Weyerstahl, *Planta Med.* **1987**, *53*, 66–72.
- [119] P. M. Abreu, R. G. Noronha, *Flavour Fragr. J.* **1997**, *12*, 79–83.
- [120] S. Baran, S. H. von Reuss, W. A. König, D. Kalemba, *Flavour Fragr. J.* **2007**, *22*, 78–83.
- [121] C. S. Mathela, A. Tiwari, R. C. Padalia, C. S. Chanotiya, *Ind. J. Chem. B* **2008**, *47*, 1249–1253.
- [122] H.-J. Park, S.-H. Kwon, K.-O. Yoo, I.-C. Sohn, K.-T. Lee, H.-K. Lee, *Planta Med.* **2000**, *66*, 783–784.
- [123] P. S. Beauchamp, V. Dev, D. R. Docter, R. Ehsani, G. Vita, A. B. Melkani, C. S. Mathela, A. T. Bottini, *J. Essent. Oil Res.* **1996**, *8*, 117–121.
- [124] C. L. Cantrell, J. A. Klun, C. T. Bryson, M. Kobaisy, S. O. Duke, *J. Agric. Food Chem.* **2005**, *53*, 5948–5953.
- [125] Y. Saito, M. Taniguchi, T. Komiya, A. Ohsaki, Y. Okamoto, X. Gong, C. Kuroda, M. Tori, *Tetrahedron* **2013**, *69*, 8505–8510.
- [126] Y.-L. Li, Y.-X. Gao, H.-Z. Jin, L. Shan, W.-L. Chang, X.-W. Yang, H.-W. Zeng, N. Wang, A. Steinmetz, W.-D. Zhang, *Phytochemistry* **2015**, *117*, 135–143.
- [127] L.-L. Liu, T. K. Q. Ha, W. Ha, W. K. Oh, J.-L. Yang, Y.-P. Shi, *J. Nat. Prod.* **2017**, *80*, 298–307.
- [128] R. K. Thappa, K. L. Dhar, C. K. Atal, *Phytochemistry* **1979**, *18*, 671–672.
- [129] J. W. Huffman, A. R. Pinder, *Phytochemistry* **1980**, *19*, 2468–2469.
- [130] A. A. Ahmed, A. A. Mahmoud, *Tetrahedron* **1998**, *54*, 8141–8152.
- [131] I. Burkhardt, N. Kreuzenbeck, C. Beemelmans, J. S. Dickschat, *Org. Biomol. Chem.* **2019**, *17*, 3348–3355.
- [132] P. Rabe, J. Rinkel, T. A. Klapschinski, L. Barra, J. S. Dickschat, *Org. Biomol. Chem.* **2016**, *14*, 158–164.
- [133] R. R. Heath, R. E. Doolittle, P. E. Sonnet, J. H. Tumlinson, *J. Org. Chem.* **1980**, *45*, 2910–2912.
- [134] D. Caine, B. Stanhope, S. Fiddler, *J. Org. Chem.* **1988**, *53*, 4124–4127.
- [135] D. Caine, B. Stanhope, *Tetrahedron* **1987**, *43*, 5545–5555.
- [136] M. Ando, T. Asao, N. Hiratsuka, K. Takase, T. Nozoe, *Bull. Chem. Soc. Jpn.* **1980**, *53*, 1425–1434.
- [137] M. Ando, K. Arai, K. Kikuchi, K. Isogai, *J. Nat. Prod.* **1994**, *57*, 1189–1199.
- [138] S. A. Agger, F. Lopez-Gallego, T. R. Hoye, C. Schmidt-Dannert, *J. Bacteriol.* **2008**, *190*, 6084–6096.
- [139] R. E. K. Winter, F. Dron, D. Arigoni, *J. Org. Chem.* **1980**, *45*, 4786–4789.
- [140] R. H. Scheffrahn, J. J. Sims, R. K. Lee, M. K. Rust, *J. Nat. Prod.* **1986**, *49*, 699–701.
- [141] A. M. Adio, C. Paul, P. Kloth, W. A. König, *Phytochemistry* **2004**, *65*, 199–206.
- [142] J. E. McMurry, P. Kocovsky, *Tetrahedron Lett.* **1985**, *26*, 2171–2172.
- [143] Y. Ohta, N. H. Andersen, C.-B. Liu, *Tetrahedron* **1977**, *33*, 617–628.
- [144] Y. Asakawa, M. Toyota, R. Takeda, C. Suire, T. Takemoto, *Phytochemistry* **1981**, *20*, 725–728.
- [145] A. Matsuo, O. Ihsii, M. Suzuki, M. Nakayama, S. Hayashi, *Z. Naturforsch. Sect. B* **1982**, *37*, 1636–1639.
- [146] K. Kondo, M. Toyota, Y. Asakawa, *Phytochemistry* **1990**, *29*, 2197–2199.
- [147] M. Toyota, I. Omatsu, J. Braggins, Y. Asakawa, *Chem. Pharm. Bull.* **2004**, *52*, 481–484.
- [148] N. L. Brock, J. S. Dickschat, *ChemBioChem* **2013**, *14*, 1189–1193.
- [149] T. A. van Beek, R. Kleis, M. A. Posthumus, A. van Veldhuizen, *Phytochemistry* **1989**, *28*, 1909–1911.
- [150] J. Lückner, P. Bowen, J. Bohlmann, *Phytochemistry* **2004**, *65*, 2649–2659.
- [151] D. M. Martin, O. Toub, A. Chiang, B. C. Lo, S. Ohse, S. T. Lund, J. Bohlmann, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 7245–7250.
- [152] F. Shimoma, H. Kondo, S. Yuuya, T. Suzuki, H. Hagiwara, M. Ando, *J. Nat. Prod.* **1998**, *61*, 22–28.
- [153] Y. Saritas, N. Bülow, C. Fricke, W. A. König, H. Muhle, *Phytochemistry* **1998**, *48*, 1019–1023.
- [154] F. Bohlmann, C. Zdero, R. M. King, H. Robinson, *Liebigs Ann. Chem.* **1984**, 503–511.
- [155] N. H. Andersen, P. Bissonette, C.-B. Liu, B. Shunk, Y. Ohta, C.-L. W. Tseng, A. Moore, S. Huneck, *Phytochemistry* **1977**, *16*, 1731–1751.
- [156] H. Tesso, W. A. König, Y. Asakawa, *Phytochemistry* **2005**, *66*, 941–949.
- [157] L. Novotný, V. Herout, F. Sorm, *Tetrahedron Lett.* **1961**, *2*, 697–701.
- [158] J. Hochmannová, L. Novotny, V. Herout, *Collect. Czech. Chem. C* **1962**, *27*, 1870–1876.
- [159] J. Krepinsky, O. Motl, L. Dolejs, L. Novotny, V. Herout, R. B. Bates, *Tetrahedron Lett.* **1968**, *9*, 3315–3318.
- [160] J. Hochmannová, V. Herout, *Collect. Czech. Chem. C* **1964**, *29*, 2369–2376.
- [161] L. Novotný, J. Toman, F. Stary, A. D. Marquez, V. Herout, F. Sorm, *Phytochemistry* **1966**, *5*, 1281–1287.
- [162] F. Bohlmann, W.-R. Abraham, *Phytochemistry* **1978**, *17*, 1629–1635.
- [163] F. Bohlmann, K.-H. Knoll, *Liebigs Ann. Chem.* **1979**, 470–472.
- [164] S. Dupré, M. Grenz, J. Jakupovic, F. Bohlmann, H. M. Niemeyer, *Phytochemistry* **1991**, *30*, 1211–1220.
- [165] M. Tori, H. Nakamizo, K. Mihara, M. Sato, Y. Okamoto, K. Nakashima, M. Tanaka, Y. Saito, M. Sono, X. Gong, Y. Shen, R. Hanai, C. Kuroda, *Phytochemistry* **2008**, *69*, 1158–1165.
- [166] M. Tori, A. Watanabe, S. Matsuo, Y. Okamoto, K. Tachikawa, S. Takaoka, X. Gong, C. Kuroda, R. Hanai, *Tetrahedron* **2008**, *64*, 4486–4495.

- [167] H. Nagano, A. Torihata, M. Matsushima, R. Hanai, Y. Saito, M. Baba, Y. Tanio, Y. Okamoto, Y. Takashima, M. Ichihara, X. Gong, C. Kuroda, M. Tori, *Helv. Chim. Acta* **2009**, *92*, 2071–2081.
- [168] J. L. Frenz-Ross, J. J. Enticknap, R. G. Kerr, *Mar. Biotechnol.* **2008**, *10*, 572–578.
- [169] Y. Asakawa, X. Lin, K. Kondo, Y. Fukuyama, *Phytochemistry* **1991**, *30*, 4019–4024.
- [170] F. Näf, R. Decorzant, W. Thommen, *Helv. Chim. Acta* **1982**, *65*, 2212–2223.
- [171] R. M. Coates, J. E. Shaw, *J. Org. Chem.* **1970**, *35*, 2597–2601.
- [172] A. R. Penfold, J. L. Simonsen, *J. Chem. Soc.* **1939**, 87–89.
- [173] L. H. Zalkow, G. L. Chetty, *J. Org. Chem.* **1975**, *40*, 1833–1834.
- [174] H. Itokawa, H. Morita, K. Watanabe, S. Mihashi, Y. Itaka, *Chem. Pharm. Bull.* **1985**, *33*, 1148–1153.
- [175] Y. Zhao, D. J. Schenk, S. Takahashi, J. Chappell, R. M. Coates, *J. Org. Chem.* **2004**, *69*, 7428–7435.
- [176] A. H. Elmi, M. H. Farah, E. Fattorusso, S. Magno, L. Mayol, *Phytochemistry* **1987**, *26*, 3069–3071.
- [177] J. Jakupovic, M. Grenz, F. Bohlmann, D. C. Wasshausen, R. M. King, *Phytochemistry* **1989**, *28*, 1937–1941.
- [178] G. Blay, L. Cardona, A. M. Collado, B. Garcia, V. Morcillo, J. R. Pedro, *J. Org. Chem.* **2004**, *69*, 7294–7302.
- [179] M. Tori, M. Aoki, K. Nakashima, Y. Asakawa, *Phytochemistry* **1995**, *39*, 99–103.
- [180] G. L. K. Hunter, W. B. Brogden, *J. Food Sci.* **1965**, *30*, 1–4.
- [181] G. L. K. Hunter, W. B. Brogden, *J. Food Sci.* **1965**, *30*, 876–878.
- [182] S. Schulz, M. Girhard, S. K. Gaßmeyer, V. D. Jäger, D. Schwarze, A. Vogel, V. B. Urlacher, *ChemCatChem* **2015**, *7*, 601–604.
- [183] W. D. MacLeod, *Tetrahedron Lett.* **1965**, *6*, 4779–4783.
- [184] H. Hikino, N. Suzuki, T. Takemoto, *Chem. Pharm. Bull.* **1968**, *16*, 832–838.
- [185] B. M. Lawrence, J. W. Hogg, *Phytochemistry* **1973**, *12*, 2995.
- [186] T. Hashimoto, H. Irita, M. Tanaka, S. Takaoka, Y. Asakawa, *Phytochemistry* **2000**, *53*, 593–604.
- [187] U. Groenhagen, M. Maczka, J. S. Dickschat, S. Schulz, *Beilstein J. Org. Chem.* **2014**, *10*, 1421–1432.
- [188] L. Sharon-Asa, M. Shalit, A. Frydman, E. Bar, D. Holland, E. Or, U. Lavi, E. Lewinsohn, Y. Eyal, *Plant J.* **2003**, *36*, 664–674.
- [189] J. Beekwilder, A. van Houwelingen, K. Cankar, A. D. J. van Dijk, R. M. de Jong, G. Stoop, H. Bouwmeester, J. Achkar, T. Sonke, D. Bosch, *Plant Biotechnol. J.* **2014**, *12*, 174–182.
- [190] K. Cankar, A. van Houwelingen, M. Goedbloed, R. Renirie, R. M. de Jong, H. Bouwmeester, D. Bosch, T. Sonke, J. Beekwilder, *FEBS Lett.* **2014**, *588*, 1001–1007.
- [191] T. R. Govindachari, P. A. Mohamed, P. C. Parthasarathy, *Tetrahedron* **1970**, *26*, 615–619.
- [192] D. E. Cane, E. J. Salaski, P. C. Prabhakaran, *Tetrahedron Lett.* **1990**, *31*, 1943–1944.
- [193] G. Blay, L. Cardona, A. M. Collado, B. Garcia, J. R. Pedro, *J. Org. Chem.* **2006**, *71*, 4929–4936.
- [194] P. Rabe, C. A. Citron, J. S. Dickschat, *ChemBioChem* **2013**, *14*, 2345–2354.
- [195] G. I. Birnbaum, A. Stoessl, S. H. Grover, J. B. Stothers, *Can. J. Chem.* **1974**, *52*, 993–1005.
- [196] D. E. Cane, B. J. Rawlings, C.-C. Yang, *J. Antibiot.* **1987**, *40*, 1331–1334.
- [197] R. Näf, A. Velluz, A. P. Meyer, *J. Essent. Oil Res.* **1996**, *8*, 587–595.
- [198] A. K. Ganguly, K. W. Gopinath, T. R. Govindachari, K. Nagarajan, B. R. Pai, P. C. Parthasarathy, *Tetrahedron Lett.* **1969**, *10*, 133–136.
- [199] T. R. Govindachari, K. Nagarajan, P. C. Parthasarathy, *Chem. Commun.* **1969**, 823a–824.
- [200] C. A. Carollo, A. R. Hellmann, J. Maximo de Siqueira, *Biochem. Syst. Ecol.* **2005**, *33*, 647–649.
- [201] L. C. Teng, J. F. DeBardleben, *Experientia* **1971**, *27*, 14–15.
- [202] P. Weyerstahl, U. Splittgerber, H. Marschall, *Flavour Fragr. J.* **1995**, *10*, 365–370.
- [203] J. Gupta, M. Ali, K. K. Pillai, A. Velasco-Negueruela, M. J. Perez-Alonso, F. Gomez Contreras, *J. Essent. Oil Res.* **1997**, *9*, 667–672.
- [204] R. Ratnayake, B. M. Ratnayake Bandara, S. Wijesundara, J. K. MacLeod, P. Simmonds, V. Karunaratne, *Nat. Prod. Res.* **2008**, *22*, 1393–1402.
- [205] S. A. Achmad, E. L. Ghisalberti, E. H. Hakim, L. Makmur, M. Manurung, *Phytochemistry* **1992**, *31*, 2153–2154.
- [206] R. M. Cory, J. B. Sothers, *Org. Magn. Reson.* **1978**, *11*, 252–257.
- [207] K. R. Varma, M. L. Maheshwari, S. C. Bhattacharyya, *Tetrahedron* **1965**, *21*, 115–138.
- [208] C. L. Wu, *J. Hattori Bot. Lab.* **1984**, *56*, 221–226.
- [209] C. L. Wu, T. J. Lee, *Proc. Natl. Sci. Coun. ROC B* **1983**, 428–431.
- [210] I. M. Whitehead, D. F. Ewing, D. R. Threlfall, D. E. Cane, P. C. Prabhakaran, *Phytochemistry* **1990**, *29*, 479–482.
- [211] I. M. Whitehead, D. F. Ewing, D. R. Threlfall, *Phytochemistry* **1988**, *27*, 1365–1370.
- [212] I. M. Whitehead, D. R. Threlfall, D. F. Ewing, *Phytochemistry* **1989**, *28*, 775–779.
- [213] Y. Yoshizawa, T. Yamaura, S. Kawaii, T. Hoshino, J. Mizutani, *Biosci. Biotechnol. Biochem.* **1994**, *58*, 305–308.
- [214] U. Vögeli, J. W. Freeman, J. Chappell, *Plant Physiol.* **1990**, *93*, 182–187.
- [215] P. J. Facchini, J. Chappell, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 11088–11092.
- [216] K. Back, S. Yin, J. Chappell, *Arch. Biochem. Biophys.* **1994**, *315*, 527–532.
- [217] K. Back, S. He, K. U. Kim, D. H. Shin, *Plant Cell Physiol.* **1998**, *39*, 899–904.
- [218] C. M. Starks, K. Back, J. Chappell, J. P. Noel, *Science* **1997**, *277*, 1815–1820.
- [219] K. A. Rising, C. M. Starks, J. P. Noel, J. Chappell, *J. Am. Chem. Soc.* **2000**, *122*, 1861–1866.
- [220] L. Ralston, S. T. Kwon, M. Schoenbeck, J. Ralston, D. J. Schenk, R. M. Coates, J. Chappell, *Arch. Biochem. Biophys.* **2001**, *393*, 222–235.
- [221] V. J. J. Martin, Y. Yoshikuni, J. D. Keasling, *Biotechnol. Bioeng.* **2001**, *75*, 497–503.
- [222] H.-J. Lee, D.-E. Lee, S.-B. Ha, S.-W. Jang, I.-J. Lee, S.-B. Ryu, K. Back, *Plant Sci.* **2004**, *166*, 881–887.
- [223] S. Takahashi, Y. Yeo, B. T. Greenhagen, T. McMullin, L. Song, J. Maurina-Brunker, R. Rosson, J. P. Noel, J. Chappell, *Biotechnol. Bioeng.* **2007**, *97*, 170–181.
- [224] J. E. Diaz, C.-S. Lin, K. Kunishiro, B. K. Feld, S. K. Avrantinis, J. Bronson, J. Greaves, J. G. Saven, G. A. Weiss, *Protein Sci.* **2011**, *20*, 1597–1606.
- [225] C. Rao, G. V. S. Raju, P. G. Krishna, *Indian J. Chem. B* **1982**, *21*, 267–268.
- [226] M. D. Eggers, V. Sinnwell, E. Stahl-Biskup, *Phytochemistry* **1999**, *51*, 987–990.
- [227] K. Back, J. Chappell, *J. Biol. Chem.* **1995**, *270*, 7375–7381.
- [228] J. R. Mathis, K. Back, C. Starks, J. Noel, C. D. Poulter, J. Chappell, *Biochemistry* **1997**, *36*, 8340–8348.
- [229] V. Falara, T. A. Akhtar, T. T. H. Nguyen, E. A. Spyropoulou, P. M. Bleeker, I. Schauvinhold, Y. Matsuba, M. E. Bonini, A. L. Schilmiller, R. L. Last, R. C. Schuurink, E. Pichersky, *Plant Physiol.* **2011**, *157*, 770–789.
- [230] 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.
- [231] D. T. Coxon, K. R. Price, B. Howard, S. F. Osman, E. B. Kalan, R. M. Zacharius, *Tetrahedron Lett.* **1974**, *15*, 2921–2924.
- [232] J. R. Hwu, J. M. Wetzel, *J. Org. Chem.* **1992**, *57*, 922–928.
- [233] S. Takahashi, Y.-S. Yeo, Y. Zhao, P. E. O'Maille, B. T. Greenhagen, J. P. Noel, R. M. Coates, J. Chappell, *J. Biol. Chem.* **2007**, *282*, 31744–31754.
- [234] P. E. O'Maille, J. Chappell, J. P. Noel, *Arch. Biochem. Biophys.* **2006**, *448*, 73–82.
- [235] H. J. Koo, C. R. Vickery, Y. Xu, G. V. Louie, P. E. O'Maille, M. Bowman, C. M. Nartey, M. D. Burkart, J. P. Noel, *J. Antibiot.* **2016**, *69*, 524–533.
- [236] K. Back, J. Chappell, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 6841–6845.
- [237] B. T. Greenhagen, P. E. O'Maille, J. P. Noel, J. Chappell, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 9826–9831.
- [238] P. E. O'Maille, A. Malone, N. Dellas, B. A. Hess, L. Smentek, I. Sheehan, B. T. Greenhagen, J. Chappell, G. Manning, J. P. Noel, *Nat. Chem. Biol.* **2008**, *4*, 617–623.
- [239] A. Schiffrin, T. T. B. Ly, N. Günnewich, J. Zapp, V. Thiel, S. Schulz, F. Hanemann, Y. Khatri, R. Bernhardt, *ChemBioChem* **2015**, *16*, 337–344.
- [240] I. Burkhardt, T. Siemon, M. Henrodt, L. Studt, S. Rösler, B. Tudzynski, M. Christmann, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2016**, *55*, 8748–8751; *Angew. Chem.* **2016**, *128*, 8890–8893.
- [241] K. Hatzellis, G. Pagona, A. Spyros, C. Demetzos, H. E. Katerinopoulos, *J. Nat. Prod.* **2004**, *67*, 1996–2001.
- [242] K. Karlshøj, T. O. Larsen, *J. Agric. Food Chem.* **2005**, *53*, 708–715.

- [243] J. C. R. Demyttenaere, A. Adams, K. van Belleghem, N. de Kimpe, W. A. König, A. V. Tkachev, *Phytochemistry* **2002**, *59*, 597–602.
- [244] T. M. Hohn, R. D. Plattner, *Arch. Biochem. Biophys.* **1989**, *272*, 137–143.
- [245] D. E. Cane, P. C. Prabhakaran, E. J. Salaski, P. H. M. Harrison, H. Noguchi, B. J. Rawlings, *J. Am. Chem. Soc.* **1989**, *111*, 8914–8916.
- [246] D. E. Cane, Z. Wu, R. H. Proctor, T. M. Hohn, *Arch. Biochem. Biophys.* **1993**, *304*, 415–419.
- [247] D. E. Cane, I. Kang, *Arch. Biochem. Biophys.* **2000**, *376*, 354–364.
- [248] X. Tang, R. K. Allemann, T. Wirth, *Eur. J. Org. Chem.* **2017**, 414–418.
- [249] M. J. Calvert, P. R. Ashton, R. K. Allemann, *J. Am. Chem. Soc.* **2002**, *124*, 11636–11641.
- [250] B. Felicetti, D. E. Cane, *J. Am. Chem. Soc.* **2004**, *126*, 7212–7221.
- [251] D. E. Cane, Y. S. Tsantrizos, *J. Am. Chem. Soc.* **1996**, *118*, 10037–10040.
- [252] D. J. Miller, F. Yu, R. K. Allemann, *ChemBioChem* **2007**, *8*, 1819–1825.
- [253] D. J. Miller, F. Yu, D. W. Knight, R. K. Allemann, *Org. Biomol. Chem.* **2009**, *7*, 962–975.
- [254] R. K. Allemann, N. J. Young, S. Ma, D. G. Truhlar, J. Gao, *J. Am. Chem. Soc.* **2007**, *129*, 13008–13013.
- [255] D. J. Miller, J. Gao, D. G. Truhlar, N. J. Young, V. Gonzalez, R. K. Allemann, *Org. Biomol. Chem.* **2008**, *6*, 2346–2354.
- [256] J. M. Caruthers, I. Kang, J. Rynkiewicz, D. E. Cane, D. W. Christianson, *J. Biol. Chem.* **2000**, *275*, 25533–25539.
- [257] E. Y. Shishova, L. di Costanzo, D. E. Cane, D. W. Christianson, *Biochemistry* **2007**, *46*, 1941–1951.
- [258] A. Deligeorgopoulou, S. E. Taylor, S. Forcat, R. K. Allemann, *Chem. Commun.* **2003**, 2162–2163.
- [259] J. A. Faraldos, B. Kariuki, R. K. Allemann, *J. Org. Chem.* **2010**, *75*, 1119–1125.
- [260] J. A. Faraldos, R. K. Allemann, *Org. Lett.* **2011**, *13*, 1202–1205.
- [261] R.-D. Wei, H. K. Schnoes, P. A. Hart, F. M. Strong, *Tetrahedron* **1975**, *31*, 109–114.
- [262] Y. Moule, M. Jemmali, N. Rousseau, *Chem.-Biol. Interact.* **1976**, *14*, 207–216.
- [263] Y. Moulé, M. Jemmali, N. Darracq, *FEBS Lett.* **1978**, *88*, 341–344.
- [264] F.-C. Chen, C.-F. Chen, R.-D. Wei, *Toxicon* **1982**, *20*, 433–441.
- [265] S. Moreau, A. Gaudemer, A. Lablache-Combier, J. Biguet, *Tetrahedron Lett.* **1976**, *17*, 833.
- [266] S. Moreau, M. Cancan, *J. Org. Chem.* **1977**, *42*, 2632–2634.
- [267] S. Moreau, J. Biguet, A. Lablache-Combier, F. Baert, M. Foulon, C. Delfosse, *Tetrahedron* **1980**, *36*, 2989–2997.
- [268] D. Sørensen, A. Raditsis, L. A. Trimble, B. A. Blackwell, M. W. Sumarah, J. D. Miller, *J. Nat. Prod.* **2007**, *70*, 121–123.
- [269] R. Riclea, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2015**, *54*, 12167–12170; *Angew. Chem.* **2015**, *127*, 12335–12338.
- [270] S.-C. Chang, K.-L. Lu, S.-F. Yeh, *Appl. Environ. Microbiol.* **1993**, *59*, 981–986.
- [271] H. H. Jeleń, *J. Agric. Food Chem.* **2002**, *50*, 6569–6574.
- [272] P. I. Hidalgo, R. V. Ullán, S. M. Albillos, O. Montero, M. Á. Fernández-Bodega, C. García-Estrada, M. Fernández-Aguado, J.-F. Martín, *Fungal Genet. Biol.* **2014**, *62*, 11–24.
- [273] P. I. Hidalgo, E. Poirier, R. V. Ullán, J. Piqueras, L. Meslet-Cladiere, E. Coton, M. Coton, *Appl. Microbiol. Biotechnol.* **2017**, *101*, 2043–2056.
- [274] S.-C. Chang, Y.-H. Wei, M.-L. Liu, R.-D. Wei, *Appl. Environ. Microbiol.* **1985**, *49*, 1455–1460.
- [275] F. Šorm, L. Dolejš, O. Knessl, J. Pliva, *Coll. Czech. Chem. Commun.* **1950**, *15*, 82–95.
- [276] R. B. Bates, R. C. Slagel, *J. Am. Chem. Soc.* **1962**, *84*, 1307–1308.
- [277] L. Dolejš, A. Mironov, F. Šorm, *Coll. Czech. Chem. Commun.* **1961**, *26*, 1015–1020.
- [278] H. Minato, *Tetrahedron Lett.* **1961**, *2*, 280–284.
- [279] E. Piers, K. F. Cheng, *J. Chem. Soc. Chem. Commun.* **1969**, *0*, 562–563.
- [280] E. Piers, K. F. Cheng, *Can. J. Chem.* **1970**, *48*, 2234–2245.
- [281] J. P. Minyard, J. H. Tumlinson, A. C. Thompson, P. A. Hedin, *J. Agric. Food Chem.* **1966**, *14*, 332–336.
- [282] M. Ishihara, T. Tsuneya, M. Shiga, K. Uneyama, *Phytochemistry* **1991**, *30*, 563–566.
- [283] S. Hayashi, N. Hayashi, K. Yano, M. Okano, T. Matsuura, *Bull. Chem. Soc. Jpn.* **1968**, *41*, 234–236.
- [284] H. Nii, K. Furukawa, M. Iwakiri, T. Kubota, *Nippon Nogeik. Kaishi* **1983**, *57*, 733–741.
- [285] M. Mundina, R. Vila, F. Tomi, M. P. Gupta, T. Adzet, J. Casanova, S. Canigual, *Phytochemistry* **1998**, *47*, 1277–1282.
- [286] M. Ito, K.-I. Okimoto, T. Yagura, G. Honda, F. Kiuchi, Y. Shimada, *J. Essent. Oil Res.* **2005**, *17*, 175–180.
- [287] Y. Kumeta, M. Ito, *Plant Biotechnol.* **2009**, *26*, 307–315.
- [288] Y. Kumeta, M. Ito, *Plant Physiol.* **2010**, *154*, 1998–2007.
- [289] D. P. Drew, T. B. Andersen, C. Sweetman, B. L. Moller, C. Ford, H. T. Simonsen, *J. Exp. Bot.* **2016**, *67*, 799–808.
- [290] J. A. Faraldos, S. Wu, J. Chappell, R. M. Coates, *J. Am. Chem. Soc.* **2010**, *132*, 2998–3008.
- [291] H. Hikino, K. Ito, T. Takemoto, *Chem. Pharm. Bull.* **1968**, *16*, 1608–1610.
- [292] H. Itokawa, H. Morita, K. Watanabe, Y. Iitaka, *Chem. Lett.* **1984**, *13*, 451–452.
- [293] P. Weyerstahl, H. Marschall, U. Splittgerber, D. Wolf, *Flavour Fragr. J.* **2000**, *15*, 153–173.
- [294] D. B. Stierle, E. Goldstein, K. Parker, T. Bugni, C. Baarson, J. Gress, D. Blake, *J. Nat. Prod.* **2003**, *66*, 1097–1100.
- [295] S. Amand, A. Langenfeld, A. Blond, J. Dupont, B. Nay, S. Prado, *J. Nat. Prod.* **2012**, *75*, 798–801.
- [296] F.-W. Dong, Z.-K. Wu, L. Yang, C.-T. Zi, D. Yang, R.-J. Ma, Z.-H. Liu, H.-R. Luo, J. Zhou, J.-M. Hu, *Phytochemistry* **2015**, *118*, 51–60.
- [297] L. Barra, B. Schulz, J. S. Dickschat, *ChemBioChem* **2014**, *15*, 2379–2383.
- [298] T. C. Fleischer, R. D. Waigh, P. G. Waterman, *J. Nat. Prod.* **1997**, *60*, 1054–1056.
- [299] K. I. Booker-Milburn, H. Jenkins, J. P. H. Charmant, P. Mohr, *Org. Lett.* **2003**, *5*, 3309–3312.
- [300] G. M. König, A. D. Wright, *J. Org. Chem.* **1997**, *62*, 3837–3840.
- [301] A.-C. Huang, S. Burrett, M. A. Sefton, D. K. Taylor, *J. Agric. Food Chem.* **2014**, *62*, 10809–10815.
- [302] A.-C. Huang, M. A. Sefton, C. J. Sumby, E. R. T. Tiekink, D. K. Taylor, *J. Nat. Prod.* **2015**, *78*, 131–145.
- [303] A.-C. Huang, M. A. Sefton, D. K. Taylor, *J. Agric. Food Chem.* **2015**, *63*, 1932–1938.
- [304] K. Takeda, H. Minato, S. Nosaka, *Tetrahedron* **1961**, *13*, 308–318.
- [305] H. Minato, *Tetrahedron* **1962**, *18*, 365–371.
- [306] V. H. Kapadia, V. G. Naik, M. S. Wadia, S. Dev, *Tetrahedron Lett.* **1967**, *8*, 4661–4667.
- [307] N. Tsubaki, K. Nishimura, Y. Hirose, *Bull. Chem. Soc. Jpn.* **1967**, *40*, 597–600.
- [308] O. Rakotonirainy, E. M. Gaydou, R. Faure, I. Bombarda, *J. Essent. Oil Res.* **1997**, *9*, 321–327.
- [309] K. Snajberk, E. Zavarin, *Phytochemistry* **1975**, *14*, 2025–2028.
- [310] A. A. Craveiro, C. H. S. Andrade, F. J. A. Matos, J. W. Alencar, *J. Nat. Prod.* **1979**, *42*, 669–671.
- [311] H. Tesso, W. A. König, K.-H. Kubeczka, M. Bartnik, K. Glowinski, *Phytochemistry* **2005**, *66*, 707–713.
- [312] C. Ehret, G. Ourisson, *Tetrahedron* **1969**, *25*, 1785–1799.
- [313] G. Blay, B. Garcia, E. Molina, J. R. Pedro, *Tetrahedron* **2007**, *63*, 9621–9626.
- [314] A.-C. Huang, C. J. Sumby, E. R. T. Tiekink, D. K. Taylor, *J. Nat. Prod.* **2014**, *77*, 2522–2536.
- [315] H. Nii, K. Furukawa, M. Iwakiri, T. Kubota, *Nippon Nogeik. Kaishi* **1983**, *57*, 725–732.
- [316] F. Bohlmann, C. Zdero, *Chem. Ber.* **1975**, *108*, 1902–1910.
- [317] Y. Liu, D. Yu, *Acta Pharm. Sin.* **1985**, *20*, 514–518.
- [318] A. Srikrishna, V. H. Pardeshi, *Tetrahedron* **2010**, *66*, 8160–8168.
- [319] A. Messer, K. McCormick, Sunjaya, H. H. Hagedorn, F. Tumbel, J. Meinwald, *J. Chem. Ecol.* **1990**, *16*, 3333–3352.
- [320] S. C. Joshi, R. C. Padalia, D. S. Bisht, C. S. Mathela, *Chem. Biodiversity* **2009**, *6*, 1364–1373.
- [321] R. Faure, E. M. Gaydou, O. Rakotonirainy, *J. Chem. Soc. Perkin Trans. 2* **1987**, 341–344.
- [322] R. Mato, R. Manzano, E. Reyes, L. Carrillo, U. Uria, J. L. Vicario, *J. Am. Chem. Soc.* **2019**, *141*, 9495–9499.
- [323] A. Akhila, P. K. Sharma, R. S. Thakur, *Phytochemistry* **1987**, *26*, 2705–2707.
- [324] A. Akhila, P. K. Sharma, R. S. Thakur, *Phytochemistry* **1988**, *27*, 2105–2108.
- [325] H. Gal, *Liebigs Ann. Chem.* **1869**, *150*, 374–376.
- [326] W. Treibs, *Liebigs Ann. Chem.* **1949**, *564*, 141–151.
- [327] G. Büchi, R. E. Erickson, *J. Am. Chem. Soc.* **1956**, *78*, 1262–1263.

- [328] G. Büchi, R. E. Erickson, N. Wakabayashi, *J. Am. Chem. Soc.* **1961**, *83*, 927–938.
- [329] G. Büchi, W. D. MacLeod, *J. Am. Chem. Soc.* **1962**, *84*, 3205–3206.
- [330] M. Dobler, J. D. Dunitz, B. Gubler, H. P. Weber, G. Büchi, J. Padilla, *Proc. Chem. Soc.* **1963**, 383.
- [331] A. Srikrishna, G. Satyanarayana, *Tetrahedron: Asymmetry* **2005**, *16*, 3992–3997.
- [332] C. S. Narayanan, K. S. Kulkarni, A. S. Vaidya, S. Kanthamani, G. L. Kumari, B. V. Bapat, S. K. Paknikar, S. N. Kulkarni, G. R. Kelkar, S. C. Bhattacharyya, *Tetrahedron* **1964**, *20*, 963–968.
- [333] K. Nishiya, T. Tsujiyama, T. Kimura, K. Takeya, H. Itokawa, Y. Itaka, *Phytochemistry* **1995**, *39*, 713–714.
- [334] C. S. Mathela, C. S. Chanotiya, S. S. Sammal, A. K. Pant, S. Pandey, *Chem. Biodiversity* **2005**, *2*, 1174–1182.
- [335] G. Rücker, U. Kretzschmar, *Planta Med.* **1972**, *21*, 1–4.
- [336] G. Rücker, J. Tautges, M. L. Maheswari, D. B. Saxena, *Phytochemistry* **1976**, *15*, 224.
- [337] D. H. R. Barton, J.-C. Beloeil, A. Billion, J. Boivin, J.-Y. Lallemand, S. Mergui, *Helv. Chim. Acta* **1987**, *70*, 273–280.
- [338] G. Wolff, G. Ourisson, *Tetrahedron* **1969**, *25*, 4903–4914.
- [339] S. J. Terhune, J. W. Hogg, B. M. Lawrence, *Tetrahedron* **1973**, *14*, 4705–4706.
- [340] G. Büchi, W. D. MacLeod, J. Padilla, *J. Am. Chem. Soc.* **1964**, *86*, 4438–4444.
- [341] B. D. Mookherjee, R. W. Trenkle, W. O. Ledig, *J. Agric. Food Chem.* **1974**, *22*, 771–773.
- [342] S. K. Paknikar, J. Veeravalli, *Indian J. Chem. B* **1980**, *19*, 432.
- [343] G. Wolff, G. Ourisson, *Tetrahedron Lett.* **1968**, *9*, 3849–3852.
- [344] A. Srikrishna, G. Ravi, *Tetrahedron* **2008**, *64*, 2565–2571.
- [345] S. C. Welch, J. M. Gruber, C. Y. Chou, M. R. Willcott, *J. Org. Chem.* **1981**, *46*, 4816–4817.
- [346] M. R. Willcott, P. A. Morrison, J. M. Assercq, S. C. Welch, *J. Org. Chem.* **1981**, *46*, 4819–4820.
- [347] M. M. Sonwa, W. A. König, *Phytochemistry* **2001**, *58*, 799–810.
- [348] S. K. Paknikar, O. Motl, K. K. Chakravarti, *Tetrahedron Lett.* **1977**, *18*, 2121–2124.
- [349] M. Mekem Sonwa, W. A. König, K.-H. Kubeczka, O. Motl, *Phytochemistry* **1997**, *45*, 1435–1439.
- [350] B. Trivedi, O. Motl, V. Herout, F. Sorm, *Coll. Czech. Chem. Commun.* **1964**, *29*, 1675–1688.
- [351] B. Trivedi, O. Motl, J. Smolikova, F. Sorm, *Tetrahedron Lett.* **1964**, *5*, 1197–1201.
- [352] H. Hikino, K. Aota, T. Takemoto, *Chem. Pharm. Bull.* **1965**, *13*, 628–630.
- [353] H. Hikino, K. Aota, T. Takemoto, *Chem. Pharm. Bull.* **1966**, *14*, 890–896.
- [354] S. B. Nerali, P. S. Kalsi, *Tetrahedron Lett.* **1965**, *6*, 4053–4056.
- [355] M. Okazaki, Y. Ishizaki, Y. Tanahashi, H. Hirota, T. Takahashi, *Yakugaku Zasshi* **1974**, *94*, 881–884.
- [356] P. Joseph-Nathan, J. D. Hernandez, L. U. Roman, E. Garcia, V. Mendoza, *Phytochemistry* **1982**, *21*, 669–672.
- [357] N. G. Bisset, M. A. Diaz-Parra, C. Ehret, G. Ourisson, *Phytochemistry* **1967**, *6*, 1396–1405.
- [358] K. Yano, T. Nishijima, *Phytochemistry* **1974**, *13*, 1207–1208.
- [359] J.-I. Iwamura, K. Komaki, K. Komai, N. Hirao, *Nippon Nogeik. Kaishi* **1978**, *52*, 379–383.
- [360] F. Bohlmann, J. Ziesche, R. M. King, H. Robinson, *Phytochemistry* **1981**, *20*, 1335–1338.
- [361] F. Bohlmann, P. Singh, J. Jakupovic, R. M. King, H. Robinson, *Phytochemistry* **1982**, *21*, 1343–1347.
- [362] P. Joseph-Nathan, J. D. Hernandez, L. U. Roman, E. Garcia, V. Mendoza, S. Mendoza, *Phytochemistry* **1982**, *21*, 1129–1132.
- [363] E. Garcia, G. V. Mendoza, J. A. Guzman, *J. Nat. Prod.* **1988**, *51*, 150–151.
- [364] J. Jakupovic, V. Castro, F. Bohlmann, *Phytochemistry* **1987**, *26*, 451–455.
- [365] R. E. Perdue, F. Tschritzis, J. Jakupovic, *Phytochemistry* **1993**, *34*, 1075–1077.
- [366] A. Aldhafer, M. Langat, B. Ndunda, D. Chirchir, J. O. Midiwo, A. Njue, S. Schwikkard, M. Carew, D. Mulholland, *Phytochemistry* **2017**, *144*, 1–8.
- [367] J. Havlik, M. Budesinsky, P. Kloucek, L. Kokoska, I. Valterova, S. Vasickova, V. Zeleny, *Phytochemistry* **2009**, *70*, 414–418.
- [368] P. Joseph-Nathan, E. Martinez, R. L. Santillan, J. R. Wesener, H. Günther, *Org. Magn. Res.* **1984**, *22*, 308–311.
- [369] A. M. Adio, *Tetrahedron* **2009**, *65*, 1533–1552.

---

Manuscript received: May 1, 2020

Revised manuscript received: May 20, 2020

Accepted manuscript online: May 22, 2020

Version of record online: September 30, 2020

## Appendix B

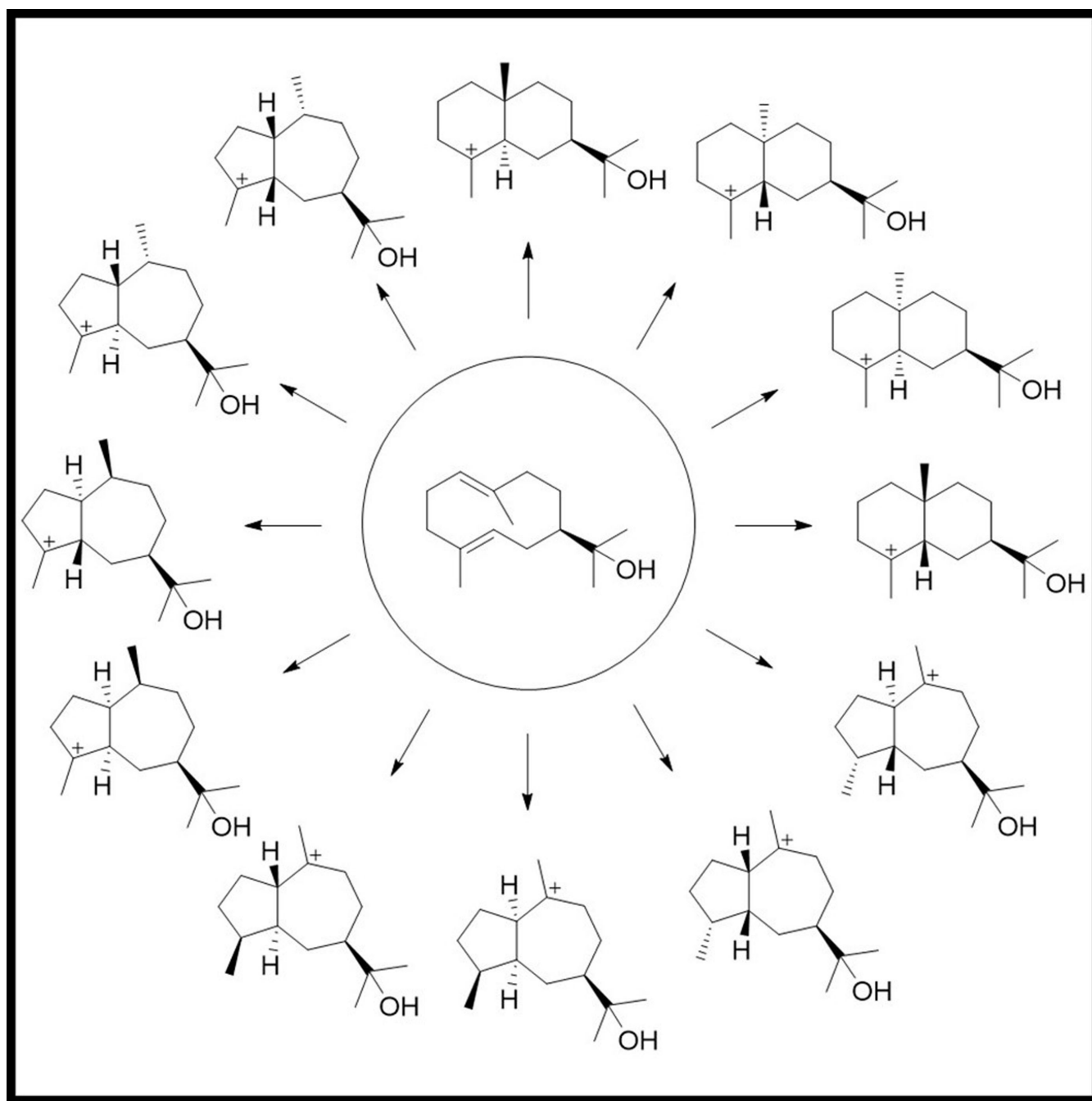
### Hedycaryol – Central Intermediates in Sesquiterpene Biosynthesis, Part II

*Chem. Eur. J.* **2022**, *28*, e202200405

DOI: 10.1002/chem.202200405

# Hedycaryol – Central Intermediates in Sesquiterpene Biosynthesis, Part II

Houchao Xu<sup>[a]</sup> and Jeroen S. Dickschat<sup>\*[a]</sup>



**Abstract:** The known sesquiterpenes that arise biosynthetically from hedycaryol are summarised. Reasonings for the assignments of their absolute configurations are discussed. The analysis provided here suggests that reprotonations at the C1=C10 double bond of hedycaryol are directed toward

C1 and generally lead to 6–6 bicyclic compounds, while reprotonations at the C4=C5 double bond occur at C4 and result in 5–7 bicyclic compounds. Read more in the Review by H. Xu and J. S. Dickschat (DOI: 10.1002/chem.202200405).

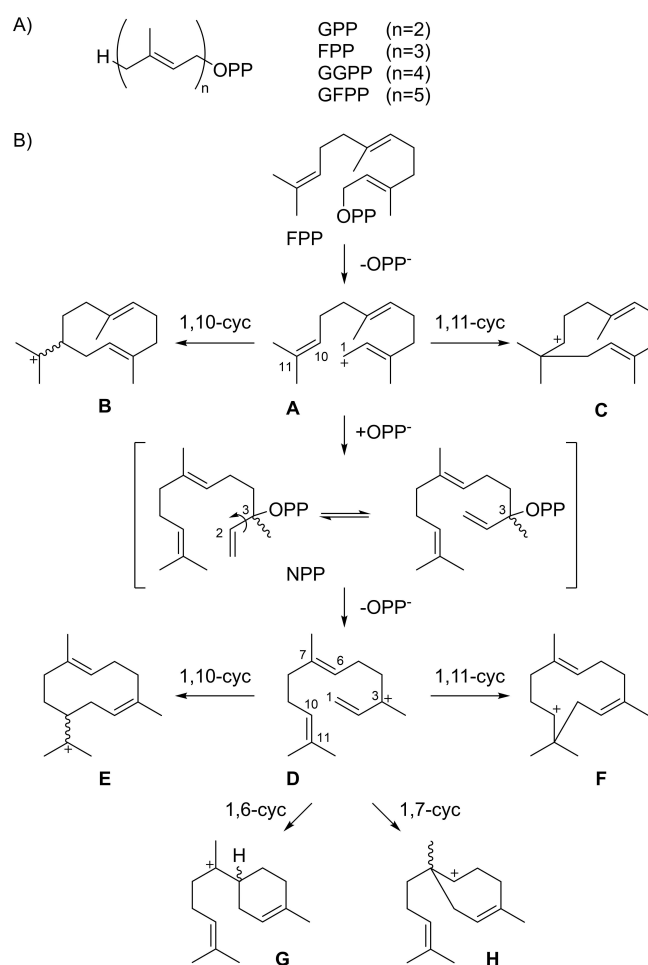
## 1. Introduction

Terpenoids represent the largest class of natural products, exhibit an extraordinary structural diversity and complexity, and are often associated with remarkable biological and pharmaceutical activities.<sup>[1]</sup> Their carbon skeletons are assembled through the action of terpene synthases from only a few acyclic precursors, oligoprenyl diphosphates, that contain multiples of five carbon units with an alkene function and a methyl branch and follow the general formula  $H-(C_5H_8)_n-OPP$  (Scheme 1A). During the past decades, many type I terpene synthases have been characterised from plants,<sup>[2–4]</sup> bacteria,<sup>[4,5]</sup> fungi<sup>[4,6]</sup> and protists<sup>[7]</sup> that act on their substrates through diphosphate abstraction, followed by a cationic cascade reaction to yield usually (poly)cyclic terpene hydrocarbons or alcohols. Subclasses of these enzymes include monoterpene synthases for the conversion of geranyl diphosphate (GPP,  $C_{10}$ ,  $n=2$ ) and sesquiterpene synthases that act on farnesyl diphosphate (FPP,  $C_{15}$ ,  $n=3$ ). For diterpene and sesterterpene synthases<sup>[4,8]</sup> the substrates geranylgeranyl diphosphate (GGPP,  $C_{20}$ ,  $n=4$ ) and geranyl farnesyl diphosphate (GFPP,  $C_{25}$ ,  $n=5$ ) with their multiple reactive double bonds allow for highly complex cyclisation cascades, leading to a fascinating structural complexity from a simple acyclic molecule in just one enzymatic step. Site-directed mutagenesis experiments gave detailed insights into terpene synthase catalysis and made enzymes with new functions available,<sup>[9]</sup> and also the conversion of non-natural substrate analogues is possible,<sup>[10]</sup> making terpene synthases particularly interesting for the enzymatic synthesis of molecules with highly complex architectures. Finally, heterologous expression approaches in engineered yeast<sup>[11]</sup> or *Escherichia coli* strains<sup>[12]</sup> add to the successful methodical repertoire of modern terpene synthase applications.

Type I terpene synthases ionise oligoprenyl diphosphates through the abstraction of diphosphate to yield a highly reactive allyl cation that can subsequently undergo a cascade reaction composed of several elementary steps including cyclisation reactions by intramolecular attack of an alkene function to a cationic centre, Wagner-Meerwein rearrange-

ments, hydride or proton shifts, and a final deprotonation or capture with water. In some cases the deprotonation to an electrically neutral compound is followed by a reprotonation event to initiate a second cyclisation cascade. Herein, for the deprotonation-reprotonation sequence combined experimental and theoretical studies have revealed the importance of main chain carbonyl oxygens and an active site water for the bacterial selinadiene synthase.<sup>[13,14]</sup>

For the conversion of FPP by sesquiterpene synthases different initial cyclisation events are possible (Scheme 1B).<sup>[15,16]</sup> After ionisation of FPP to the farnesyl cation (A), a 1,10-cyclisation can lead to the (*E,E*)-germacradienyl cation (B) or a 1,11-cyclisation may result in the (*E,E*)-humulyl cation (C).



**Scheme 1.** Terpene biosynthesis. A) Structures of oligoprenyl diphosphates. B) Cyclisation modes of FPP towards sesquiterpenes.

[a] H. Xu, Prof. Dr. J. S. Dickschat  
 Kekulé-Institute of Organic Chemistry and Biochemistry  
 University of Bonn  
 Gerhard-Domagk-Straße 1, 53121 Bonn (Germany)  
 E-mail: dickschat@uni-bonn.de

© 2022 The Authors. Chemistry - A European Journal published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Alternatively, the abstracted diphosphate can re-attack at C3 to give nerolidyl diphosphate (NPP) that can undergo a conformational change through rotation around its C2-C3 single bond. Its reionisation to **D** opens four more cyclisation options through 1,10-cyclisation to the (*Z,E*)-germacradienyl cation (**E**), 1,11-cyclisation to the (*Z,E*)-humulyl cation (**F**), 1,6-cyclisation to the bisabolyll cation (**G**) and 1,7-cyclisation to **H**. For all chiral intermediates both enantiomers can be reached through these processes.

Intermediate **B** can be deprotonated to yield germacrene **A** that is a widespread intermediate towards many eudesmane and guaiane sesquiterpene hydrocarbons that can be formed through its reprotonation-induced transannular reactions. The accumulated knowledge about this class of sesquiterpenes was recently summarised by us in a review article in this journal.<sup>[17]</sup> We have also performed a computational study to explore the chemical space through downstream hydride shifts for the different stereoisomers of the guaianes, showing that (supra-facial) 1,2-hydride shifts are always possible, while 1,3-hydride migrations can only be realised for certain geometries of the guaiane skeletons.<sup>[18]</sup> As an alternative to the deprotonation to germacrene **A**, cation **B** can also be captured by water to yield the sesquiterpene alcohol hedycaryol, which is a likewise important intermediate toward many sesquiterpene alcohols. Here we provide a comprehensive overview of the chemistry of hedycaryol and the compounds derived from it through terpene cyclase mediated downstream cyclisations.

Jeroen S. Dickschat studied Chemistry at TU Braunschweig and completed his PhD in 2004. He then moved for postdoctoral stays to Saarland University and the University of Cambridge. In 2008, he became a group leader at TU Braunschweig. In 2014, he was appointed Professor of Organic Chemistry and Biochemistry at the University of Bonn. His research interests span the synthesis and biosynthesis of natural products.



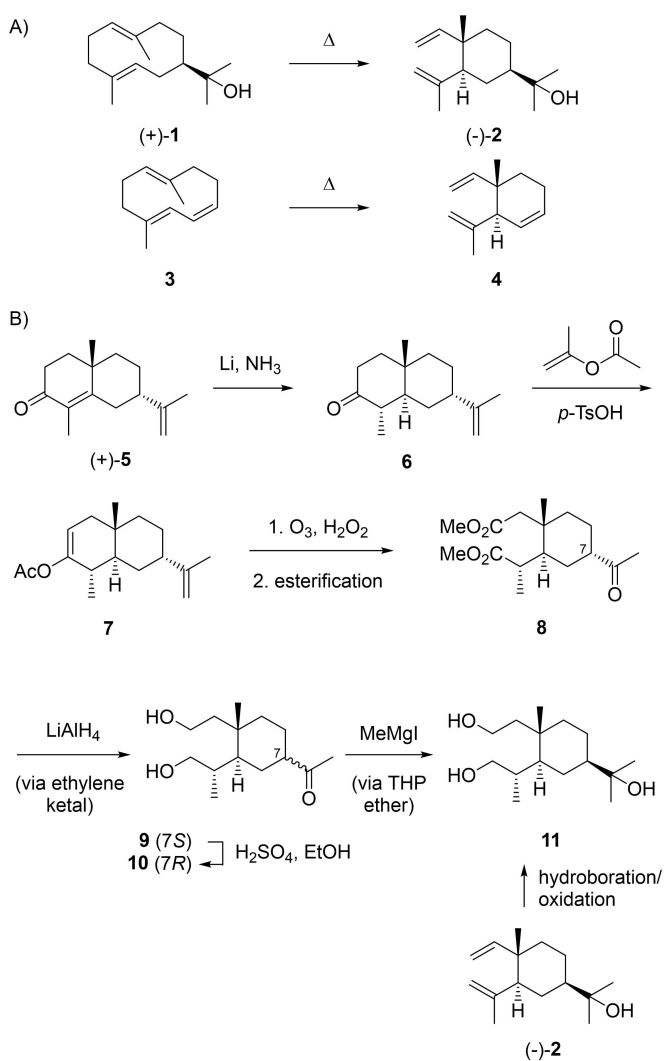
Houchao Xu graduated from China Pharmaceutical University with a B.Sc. degree in 2015. He then obtained his M.Sc. degree from Kunming Institute of Botany, Chinese Academy of Sciences. In September 2019, he started his doctoral study in the group of Prof. Dickschat at the University of Bonn. His research focuses on the chemical synthesis and biosynthesis of terpenes and polyketides.



## 2. Hedycaryol

### 2.1. Structure elucidation and occurrence in Nature

Without detailed knowledge about its structure, in 1916 Semmler and Liao discovered the first monocyclic sesquiterpene alcohol elemol (**2**, Scheme 2A) that was isolated from a fraction of the essential oil of the Philippine tree *Canarium luzonicum* (elemi) obtained by fractional distillation.<sup>[19]</sup> After establishment of its constitution by Sorm and coworkers,<sup>[20]</sup> the compound was also found to be the main constituent (60%) of the essential oil from *Hedycarya angustifolia*, a small tree native to Australia.<sup>[21]</sup> The missing optical activity of the chiral compound geijerene (**4**), the main constituent in the steam distillates from *Geijera parviflora*, was explained by Jones and Sutherland through their discovery that pregeijerene (**3**) is the true plant natural product that undergoes a Cope rearrangement during compound isolation.<sup>[22]</sup> Subsequently, the same workers also described **2** as the product of a thermal Cope



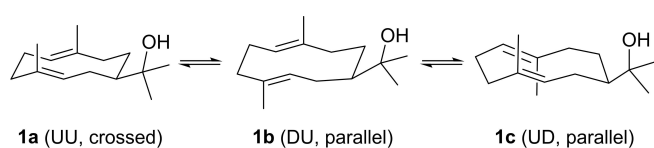
**Scheme 2.** (–)-Elemol (**2**), the Cope rearrangement product of (+)-hedycaryol (**1**). A) Cope rearrangements of **1** and pregeijerene (**3**). B) Absolute configuration of (–)-**2** by chemical correlation to (+)-10-*epi*- $\alpha$ -cyperone (**5**).



rearrangement of hedycaryol (**1**).<sup>[23]</sup> The absolute configuration of **2** has been established independently by chemical correlations to tetrahydroaussurea lactone<sup>[24]</sup> and (+)-10-*epi*- $\alpha$ -cyperone (**5**) in a procedure involving epimerisation of the side chain attached to C7 (Scheme 2B).<sup>[25]</sup> Reduction of **5** with Li in ammonia gave *trans*-fused **6** that was converted with isopropenyl acetate and *p*-TsOH into enol ester **7**, followed by ozonolysis and esterification to **8**. Reduction with LiAlH<sub>4</sub> via ketalisation with ethylene glycol gave **9** that was easily epimerised under acidic conditions to **10**. Its reaction with MeMgI via protection of the alcohol functions as tetrahydropyranyl (THP) ethers yielded **11**, the same triol that was also obtained through hydroboration and oxidation of **2**.<sup>[25]</sup>

Elemol (**2**) was later reisolated from various plants including *Juniperus sabina* and *J. scopulorum*,<sup>[26,27]</sup> *Chamaecyparis obtusa*,<sup>[28]</sup> *Citrus sinensis* and *C. nobilis*,<sup>[29–31]</sup> *Saussurea lappa*,<sup>[32]</sup> *Cinnamomum camphora*,<sup>[33]</sup> *Fokienia hodginsii*,<sup>[34]</sup> *Calycanthus floridus*,<sup>[35]</sup> *Bunium cylindricum*,<sup>[36]</sup> *Ginkgo biloba*,<sup>[37]</sup> *Amyris balsamifera*,<sup>[38]</sup> *Canarium zeylanicum*,<sup>[39]</sup> *Bothriocloa intermedia*,<sup>[40]</sup> *Commiphora abyssinica*,<sup>[41]</sup> *Santolina oblongifolia*,<sup>[42]</sup> *Cymbopogon proximus*,<sup>[43]</sup> *Eremophila flaccida*,<sup>[44]</sup> *Piper ribesoides*,<sup>[45]</sup> *Monocyclanthus vignei*,<sup>[46]</sup> *Neocallitropsis pancheri*,<sup>[47]</sup> *Cryptomeria japonica*,<sup>[48]</sup> and *Eucalyptus maculata*,<sup>[49]</sup> which demonstrates the widespread occurrence of **1** in nature. After its first report from *H. angustifolia*,<sup>[23]</sup> compound **1** was subsequently also isolated from the undistilled oils of the plants *Phebalium ozothamnoides*,<sup>[50]</sup> *Rubus rosifolius*,<sup>[51]</sup> *Thujopsis dolabrata*,<sup>[52]</sup> *Thymus praecox*,<sup>[53]</sup> *Cryptomeria japonica* and *C. fortunei*,<sup>[54]</sup> and *Chamaecyparis obtusa*.<sup>[55]</sup> For the optical rotation of **2** low negative values between  $[\alpha]_D = -2$  and  $-9.7$  are given in the literature,<sup>[24,26,27,30,32,43,46,48]</sup> while for **1** positive values between  $[\alpha]_D = +24.5$  and  $+32.7$  were reported.<sup>[23,50–52]</sup> The enantiomer (–)-**1** is only known from the bacterial hedycaryol synthase (HcS) from *Kitasatospora setae* ( $[\alpha]_D^{25} = -21.3$ ) whose Cope rearrangement gives (+)-**2** ( $[\alpha]_D^{25} = +10.0$ ).<sup>[56]</sup> This finding reflects the observation that also in other cases bacteria and fungi produce the enantiomers of plant terpenes.<sup>[57–59]</sup>

Because of its strained 10-membered ring **1** exists as a mixture of three conformers **1a** with both Me groups attached to the ring up (UU) and crossed double bonds, and **1b** and **1c** with parallel double bonds and each one Me group up and one down (DU, UD) (Scheme 3).<sup>[60,61]</sup> Their fairly slow interconversion causes line broadening in the NMR spectra, and therefore the NMR data assignment was a long standing problem that was only recently solved through a <sup>13</sup>C- and stereoselective <sup>2</sup>H-labelling approach.<sup>[56]</sup> Complete NMR data for **2** have also been published.<sup>[47]</sup> The structure and absolute configuration of (+)-**1**

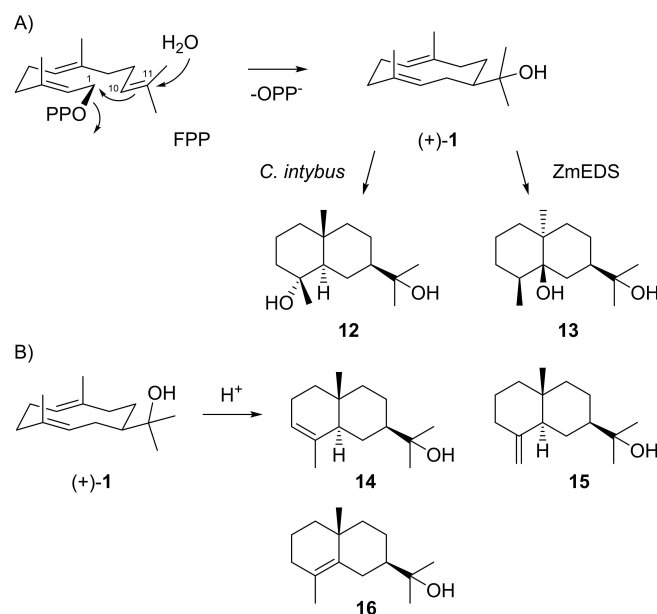


**Scheme 3.** Conformers of **1**. U=Me group at 10-membered ring up, D=Me group down. „Crossed“ and „parallel“ refers to relative orientations of double bonds.

have been further secured by an enantioselective synthesis from (–)-guaiaol.<sup>[62]</sup>

## 2.2. Biosynthesis, enzymatic and non-enzymatic cyclisation

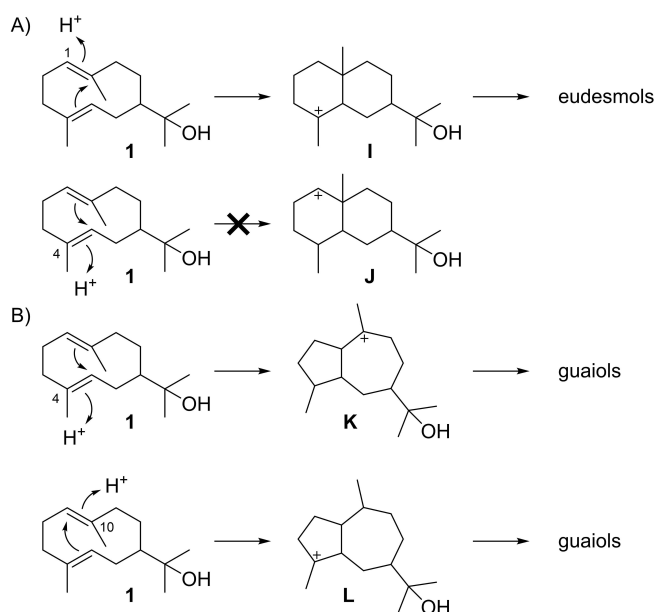
The biosynthesis of **1** by type I terpene synthases proceeds through the abstraction of diphosphate from FPP to initiate a 1,10-cyclisation and attack of water to C11 (Scheme 4A). Selective hedycaryol synthases for **1** are known from the plants *Populus trichocarpa* (PtTPS7),<sup>[63]</sup> *Camellia brevistyla* (CbTPS1),<sup>[64]</sup> and *Liquidambar formosana* (LfTPS01),<sup>[65]</sup> in all cases with undetermined absolute configuration, and for (–)-**1** from *Kitasatospora setae*,<sup>[56]</sup> whose product was initially erroneously assigned as (2Z,6E)-hedycaryol; for this bacterial enzyme also a crystal structure is available.<sup>[66]</sup> In addition, the diterpene synthase VenA from *Streptomyces venezuelae* that converts GGPP into venezuelaene A has a reported side activity with FPP as hedycaryol synthase.<sup>[67]</sup> For the diterpene synthase spiroviolene synthase from *Streptomyces violens*<sup>[68]</sup> ancestral sequence reconstruction resulted in a functional switch to a hedycaryol synthase.<sup>[69]</sup> As will be discussed in detail in this review article, **1** is an important biosynthetic intermediate, as exemplified by its reported biotransformation into cryptomeridiol (**12**) by a mortared root suspension of chicory (*Cichorium intybus*).<sup>[70]</sup> Hedycaryol (**1**) is also a proposed intermediate in the biosynthesis of eudesmane-2 $\alpha$ ,11-diol (**13**), the product of the sesquiterpene synthase ZmEDS from *Zea mays*.<sup>[71]</sup> Herein, the downstream enzymatic cyclisations of **1** are initiated by reprotonation, however, care has to be taken to distinguish enzymatic from non-enzymatic transformations, as it is well known that **1** can also undergo an efficient non-enzymatic acid



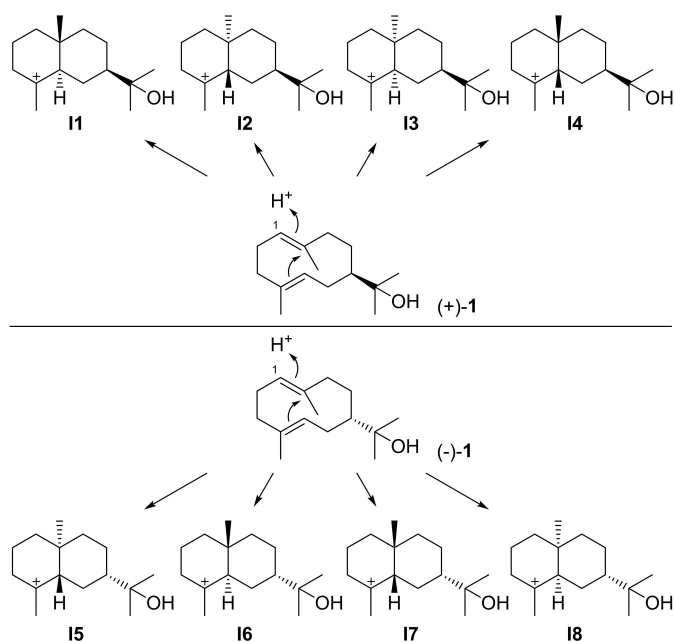
**Scheme 4.** A) Biosynthesis of **1** from FPP and its conversion into **12** and **13**. B) Acid-catalysed reaction to eudesmol **14–16**.

catalysed transannular reaction to yield a mixture mainly composed of  $\alpha$ -,  $\beta$ - and  $\gamma$ -eudesmol (**14** - **16**, Scheme 4B).<sup>[23,72,73]</sup>

Terpene synthases can further convert **1** into eudesmols or guaiaols through the protonation induced reactions shown in Scheme 5. Reprotonation of **1** at C1 can lead to **I**, the precursor to eudesmols, while the alternative reprotonation at C4 results in the secondary cation **J** that is disfavoured. For guaiaols either a protonation at C4 to **K** or at C10 to **L** are possible. The subsequent sections will give a detailed discussion of known compounds arising from **1** via these reactions.



Scheme 5. Possible terpene cyclisation modes for **1**.



Scheme 6. Cyclisation reactions of **1** induced by reprotonation at C1 towards intermediates **I1**-**I8**.

### 3. Eudesmols

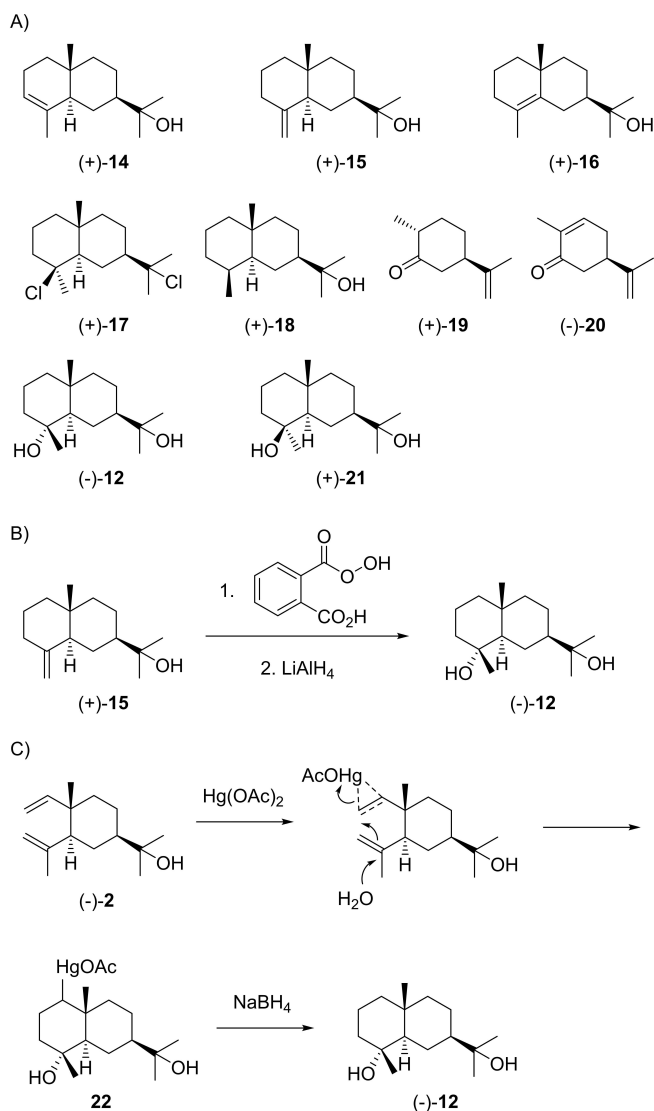
#### 3.1. Cyclisation modes from hedycaryol to eudesmols

Eudesmols can arise from (+)-**1** through protonation at C1 that can induce the cyclisation to the four stereochemically distinct intermediates **I1**-**I4** (Scheme 6). The corresponding protonation induced cyclisations from (–)-**1** gives rise to their enantiomers **I5**-**I8**. All these intermediates can potentially react by three alternative deprotonations, addition of water or intramolecular attack of the hydroxy function at the cation. Further compounds can be formed, if first a 1,2-hydride shifts occurs that may be followed by skeletal rearrangements.

#### 3.2. Eudesmols from cation **I1**

Cation **I1** can undergo deprotonations to yield  $\alpha$ -eudesmol (**14**),  $\beta$ -eudesmol (**15**) or  $\gamma$ -eudesmol (**16**, Scheme 7A). Ruzicka and coworkers demonstrated that the initially obtained “eudesmol” was a mixture of **14** and **15** of varying composition, which explained the observed variations in melting points and optical rotations.<sup>[74]</sup> Their separation from *Eucalyptus macarthurii* was first reported by McQuillin and Parrack in 1956. While the separation of **14** and **15** through chromatography on alumina or repeated recrystallisation could not fully be achieved, crystallisation of the 3,5-dinitrobenzoate esters and their saponification gave access to the pure compounds, establishing positive optical rotations for **14** ( $[\alpha]_D = +28.6$ ) and **15** ( $[\alpha]_D = +63.8$ ).<sup>[75]</sup> The same study also reported on the  $\gamma$ -isomer **16** ( $[\alpha]_D = +62.5$ ) that was obtained from (+)-selinene dihydrochloride (**17**) by elimination and hydrolysis.<sup>[75]</sup> The absolute configuration of **15** was established by Woodward and coworkers through correlation with the steroids.<sup>[76]</sup> All three eudesmols **14**–**16** yield the same hydrogenation product (+)-**18**, confirming their consistent absolute configurations.<sup>[75]</sup> Further proof for this assignment was obtained by synthesis of eudesmols **14**–**16** from (+)-dihydrocarvone (**19**).<sup>[77,78]</sup>

The alcohols **14**–**16** were frequently obtained as a mixture from various plants including different *Eucalyptus* species,<sup>[79,80]</sup> *Thuja occidentalis*,<sup>[81]</sup> and *Phebalium ozothamnoides*,<sup>[50]</sup> while the pure compounds were isolated from *Callitropsis araucarioides*,<sup>[82]</sup> *Cordia trichotoma*,<sup>[83]</sup> and *Cryptomeria japonica*.<sup>[48]</sup> Finally, **14** was also isolated from the liverwort *Porella perrottetiana*, but in this case the material showed a negative optical rotation ( $[\alpha]_D = -6.9$ ).<sup>[84]</sup> The suggested revision of the optical rotation of **14** with the structure as shown in Scheme 7A from a positive to a negative value, based on a synthetic transformation of (+)-**15** into (–)-**14**<sup>[84]</sup> conflicts all previous consistent chemical correlations. Also a later study reported a negative optical rotation for **14** obtained by total synthesis from (–)-carvone (**20**).<sup>[85]</sup> Despite the unclear situation, the structure of **14** is currently assigned with a negative optical rotation to CAS number 473–16-5. Final conclusions require further investigations (cf. also discussion in Section 3.6. about *ent*-**14** derived from **I5**). *Pterocarpus santalinus* is a reported source of pure (+)-**15**, but its comparably low optical rotation ( $[\alpha]_D^{30} = +36.0$ ) may point to a contamination



**Scheme 7.** A) Eudesmols derived from I1 and related compounds. B) Chemical correlation of (+)-15 with (-)-12 and C) of (-)-2 with (-)-12.

with (+)-14.<sup>[86]</sup> All three compounds 14–16 have been isolated from *Neocallitropsis pancheri* with full assignment of <sup>1</sup>H and <sup>13</sup>C NMR data.<sup>[47]</sup>

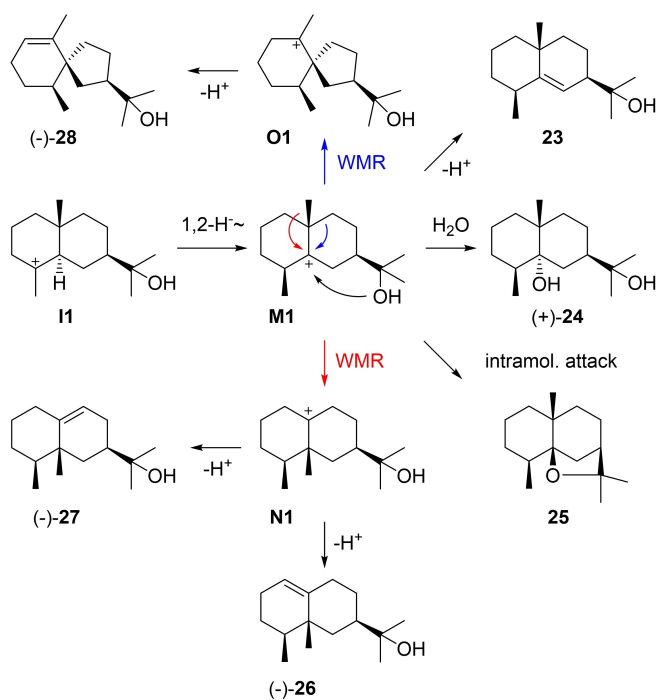
Through the attack of water to the cationic centre in I1 two diastereoisomeric diols, cryptomeridiol (12) and 4-*epi*-cryptomeridiol (21), can be formed. Cryptomeridiol (12) was first isolated from *Widdringtonia dracomontana*, but first only reported as a “diol” of negative optical rotation ( $[\alpha]_D = -24$ ).<sup>[87]</sup> It was subsequently reisolated from *Fokienia hodginsii*, shown to be identical to 12 from *W. dracomontana* by IR spectroscopy and an unchanged melting point upon admixture of an authentic sample, and its structure identified albeit with unspecified configuration at C4. The structural identification mainly relied on the conversion into (+)-17 with gaseous HCl and correlated the compound to the same enantiomeric series as the eudesmols.<sup>[88]</sup> After a third isolation from *Cryptomeria japonica* 12 was named cryptomeridiol and its structure fully

assigned by correlation with  $\beta$ -eudesmol (15) that was converted into 12 by epoxidation with monopero-phthalic acid and treatment with LiAlH<sub>4</sub> (Scheme 7B).<sup>[89]</sup> A more modern version of this synthesis using *m*CPBA for the epoxidation step was published in 1994.<sup>[90]</sup> Its identity with 12 from *W. dracomontana* and from *F. hodginsii* was not immediately recognised, possibly because of a typographical error in the given name for 12 as “selina-4,7-diol”<sup>[88]</sup> that should read “selina-4,11-diol”, but subsequently shown by IR and mixed melting point.<sup>[91]</sup> Also proximiadiol, the anti-spasmodic principle from *Cymbopogon proximus*,<sup>[92,93]</sup> was later shown to be identical to (-)-12.<sup>[94,95]</sup> Another interesting transformation that secures the absolute configuration of cryptomeridiol is the conversion of (-)-2 into (-)-12 by oxymercuration and reductive workup (Scheme 7C).<sup>[96]</sup>

The diol 12 is fairly widespread in the plant kingdom and has additionally been isolated from *Artemisia pygmaea*,<sup>[97]</sup> *Magnolia obovata*,<sup>[98]</sup> *Drymis winteri*,<sup>[99]</sup> *Hedychium spicatum*,<sup>[100]</sup> *Thujopsis dolabrata*,<sup>[101]</sup> *Carissa edulis*,<sup>[102]</sup> *Chenopodium graveolens*,<sup>[103]</sup> *Chamaecyparis pisifera*,<sup>[104]</sup> *Juglans mandshurica*<sup>[105]</sup> and *Achillea clypeolata*,<sup>[106]</sup> in all cases with a reported negative sign for the optical rotation. Compound (-)-12 was also obtained in a biotransformation of synthetic (+)-1 with a mortared root suspension of chicory.<sup>[70]</sup> A terpene synthase for 12 (of undetermined absolute configuration) is known from *Tripterygium wilfordii* (TwCS).<sup>[107]</sup> However, the surprisingly widespread occurrence of this compound in many plants may also point to a non-enzymatic formation from (+)-1 in an acid catalysed reaction e.g. during chromatographic purifications, especially if water is present,<sup>[70]</sup> or during steam distillation. This was impressively shown by steam distillation of plant leaves containing (+)-1 in the presence of H<sub>2</sub><sup>18</sup>O, leading to incorporation of the <sup>18</sup>O-label into 12 and its epimer 21.<sup>[108]</sup> Fully assigned <sup>1</sup>H- and <sup>13</sup>C NMR data were reported for 12 from the plant *Blumea balsamifera*. For unclear reasons this paper shows the enantiomer of (-)-12.<sup>[109]</sup>

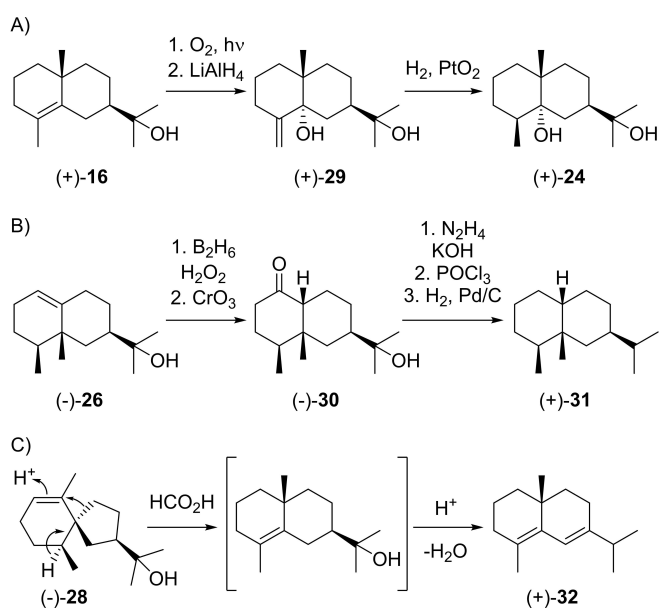
The epimer 4-*epi*-cryptomeridiol (21) was first isolated from *Amanoa oblongifolia* ( $[\alpha]_D = +3.8$ ,<sup>[110]</sup> in comparison to  $[\alpha]_D^{25} = +26.1$  for the synthetic compound obtained from (+)-15).<sup>[90]</sup> The same enantiomer (+)-21 was later reisolated from *Chamaecyparis pisifera*,<sup>[104]</sup> *Canarium ovatum*,<sup>[111]</sup> *Cryptomeria japonica*<sup>[48]</sup> and *Citrus hystrix*.<sup>[112]</sup> Fully assigned <sup>13</sup>C NMR data have been reported for synthetic 21.<sup>[90]</sup>

Cation I1 can undergo a 1,2-hydride shift to M1 that can either react by deprotonation to eudesm-5-en-11-ol (23), by capture with water to (+)-eudesmane-5 $\alpha$ ,11-diol (24), by intramolecular attack of the alcohol function to 4-*epi*-*cis*-dihydroagarofuran (25), by Wagner-Meerwein rearrangement (WMR) to N1 and deprotonation to (-)-eremoligenol (26) or its isomer 27, or by WMR to O1 and deprotonation to (-)-hinesol (28, Scheme 8). Only few reports are available for 23 that was first isolated from *Helichrysum italicum*<sup>[113]</sup> and later from *Bulnesia sarmientoi*.<sup>[114]</sup> Unfortunately, both studies did not report on the optical rotation of 23 and its absolute configuration has not formally been established, while fully assigned NMR data were given in both cases.<sup>[113,114]</sup> The diol 24 was first obtained synthetically from (+)- $\gamma$ -eudesmol (16) by photochemical



**Scheme 8.** Eudesmols derived from I1 and 1,2-hydride shift to M1.

oxidation and reduction of the allyl hydroperoxide, followed by catalytic hydrogenation (Scheme 9A), establishing its positive optical rotation ( $[\alpha]_D = +41.9$ ).<sup>[115]</sup> The same enantiomer was later reported with completely assigned NMR data from *Cryptomeria japonica*.<sup>[48]</sup> The epimer of **24** with  $5\beta$ -hydroxy group has only been obtained by synthesis,<sup>[62]</sup> but not from natural sources. The ether **25** was reported from *Cedrelopsis*



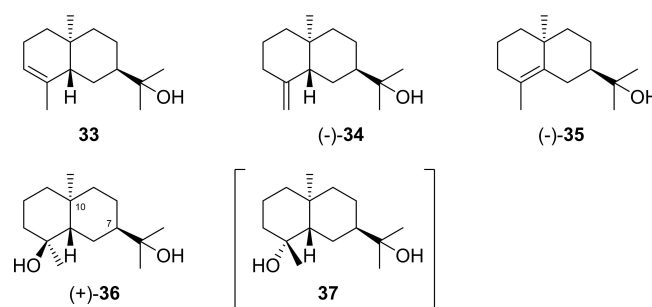
**Scheme 9.** Chemical correlations. A) Synthesis of (+)-**24** from (+)-**16**. B) Synthesis of (+)-**31** from (-)-**26**. C) Formic acid-catalysed rearrangement and dehydration of (-)-**28** to (+)-**32**.

*grevei*<sup>[116]</sup> and from *Pseuduvaria froggattii*, from which it was named froggatt ether.<sup>[117]</sup> Both studies gave fully assigned NMR data, but neither reported the optical rotation nor established the absolute configuration.<sup>[116,117]</sup>

The rearranged compound eremoligenol (**26**) was first isolated from *Ligularia fischeri* ( $[\alpha]_D = -93.5$ ) and its absolute configuration was established by correlation to (+)-eremophilane (**31**) through a sequence of hydroboration and oxidation to the ketole **30**, followed by Huang-Minlon reduction, dehydration and catalytic hydrogenation (Scheme 9B).<sup>[118]</sup> The compound was later reisolated from *Euryops sulcatus*<sup>[119]</sup> and *Oreodaphne porosa*.<sup>[120]</sup> The isomer **27** was first obtained as a synthetic material<sup>[121]</sup> followed by its isolation from *Alpinia japonica* ( $[\alpha]_D = -14.9$ ).<sup>[122]</sup> (-)-Hinesol (**28**) was first reported from *Atractylodes lancea* ( $[\alpha]_D = -40.2$ ) and shown to be a constituent of „attractylol“ that was initially believed to be a pure compound.<sup>[123]</sup> Its structure was initially wrongly assigned,<sup>[124]</sup> but later corrected with a suggested absolute configuration based on its co-occurrence with (+)- $\beta$ -eudesmol (**15**).<sup>[125]</sup> This assignment was later confirmed by a correlation with (+)- $\delta$ -selinene (**32**) that was obtained from **28** by formic acid catalysed rearrangement and dehydration (Scheme 9C), albeit not in pure form,<sup>[126]</sup> and by an enantioselective synthesis of (-)-**28**.<sup>[127]</sup> Hinesol shows an antitrypanosomal activity against *Trypanosoma brucei*.<sup>[128]</sup>

### 3.3. Eudesmols from cation I2

Cation **I2** could potentially lead to the alcohols **33–35** by deprotonation or to the diols **36** and **37** by addition of water (Scheme 10). For **33** only a synthesis of the racemate has been reported,<sup>[129]</sup> while **34** ( $[\alpha]_D^{25} = -17.5$ ) has been synthesised enantioselectively from (+)-intermedeol,<sup>[130]</sup> but both compounds are not known from natural sources. Also 10-*epi*- $\gamma$ -eudesmol (**35**) was first obtained by synthesis from dihydrocarvone (+)-**19**, unfortunately without reporting the optical rotation of **35**,<sup>[131]</sup> but the first isolation paper mentions the identity of (-)-**35** from vetiver oil (*Vetiveria zizanioides*) and the synthetic material.<sup>[132]</sup> The compound was also isolated from *Amyris balsamifera*,<sup>[38]</sup> *Aquilaria malaccensis* ( $[\alpha]_D = -68.8$ ).<sup>[133]</sup> *Alpinia japonica*,<sup>[122]</sup> *Hedychium spicatum*<sup>[134]</sup> and *Bursera graveolens*.<sup>[135]</sup>

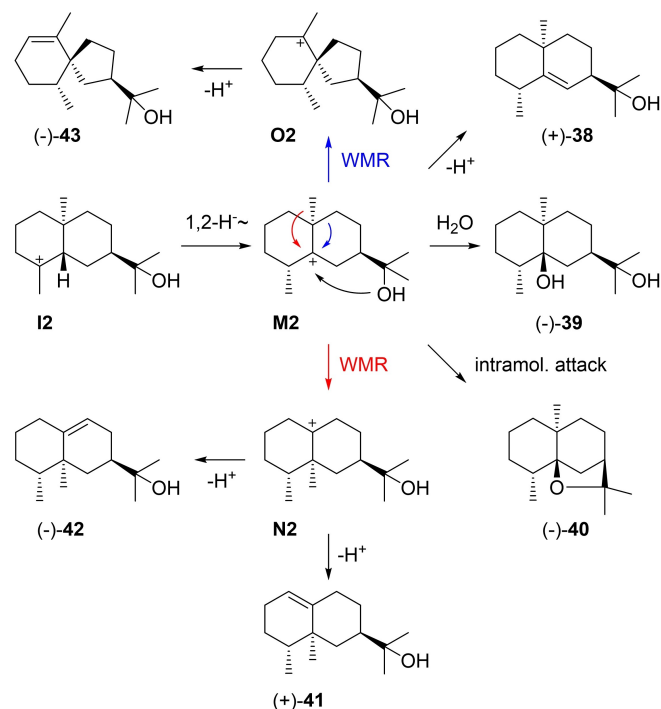


**Scheme 10.** Eudesmols derived from **I2**. Compound **37** in brackets is unknown.

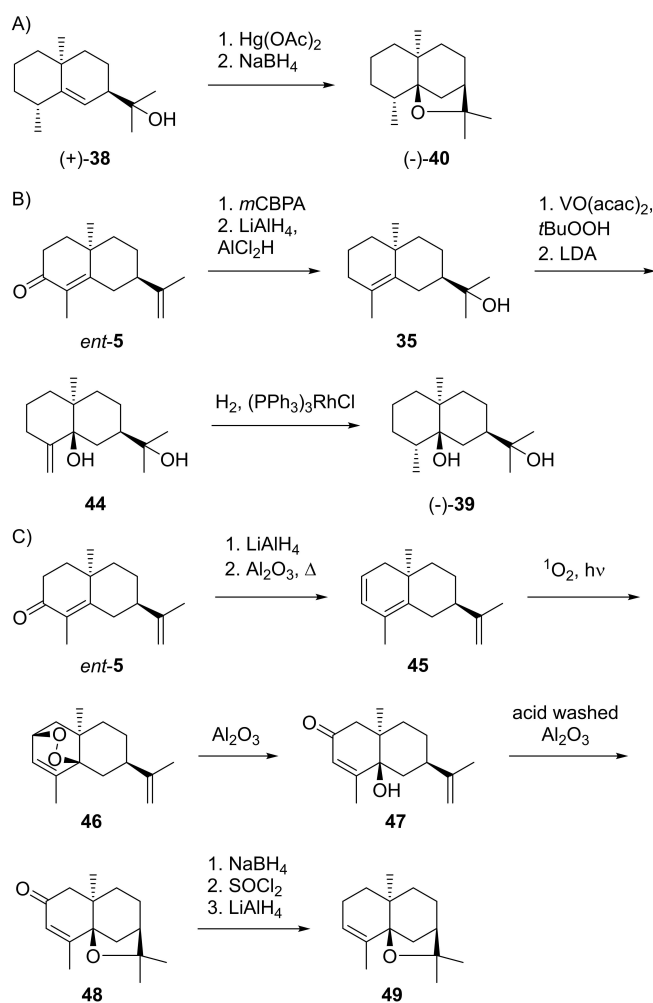
The diol **36** was also first synthesised,<sup>[136]</sup> followed by an isolation from *Ursinia trifida*,<sup>[137]</sup> in both cases without mentioning the optical rotation. At the same time the isolation of a compound from *Pluchea arguta* with same <sup>13</sup>C NMR data (apart from C4, this is likely a typographical error), but with a *cis*-decalin structure (10-*epi*-**36**) was reported ( $[\alpha]_D^{29} = +66.66$ ).<sup>[138]</sup> This erroneous structural assignment was later corrected based on a total synthesis of (+)-**36** ( $[\alpha]_D^{29} = +73.3$ ) from (+)-dihydrocarvone (**19**).<sup>[139]</sup> Pterodondioid from *Laggera pterodonta* for which initially a structure with 7S configuration was published,<sup>[140,141]</sup> is identical to **36** (with its 7R configuration), as was later demonstrated by X-ray crystallography.<sup>[142]</sup> Compound **36** is additionally known from *Goniiothalamus tapisoides*.<sup>[143]</sup> <sup>13</sup>C NMR data of **36** have been published in CDCl<sub>3</sub><sup>[137]</sup> and in C<sub>5</sub>D<sub>5</sub>N.<sup>[140]</sup> Compound **37** is unknown.

Rearranged compounds from **12** (Scheme 11) can be accessed by a 1,2-hydride shift to **M2**, from which a deprotonation leads to (+)-rosifoliol (**38**), a capture with water to (–)-**39**, and the intramolecular attack of the hydroxy function to (–)-dihydro-β-agarofuran (**40**). A methyl migration to N2 and deprotonation can result in (+)-valerianol (**41**) or (–)-jinkoheremol (**42**), while ring contraction to **O2** and deprotonation lead to (–)-agarospirol (**43**). Most of these compounds are fairly widespread.

Rosifoliol (**38**),  $[\alpha]_D = +105$ , was first isolated from *Rubus rosifolius*,<sup>[144]</sup> after its possible formation along the lines of Scheme 11 had been proposed.<sup>[145]</sup> Its structure and absolute configuration were established by correlation with (–)-**40** (Scheme 12A),<sup>[51]</sup> and also the X-ray crystal structure has been obtained.<sup>[146]</sup> The alcohol **38** was also found in *Phonus arborescens*, but this time with a reported negative optical



**Scheme 11.** Eudesmols derived from **12** and 1,2-hydride shift to **M2**.



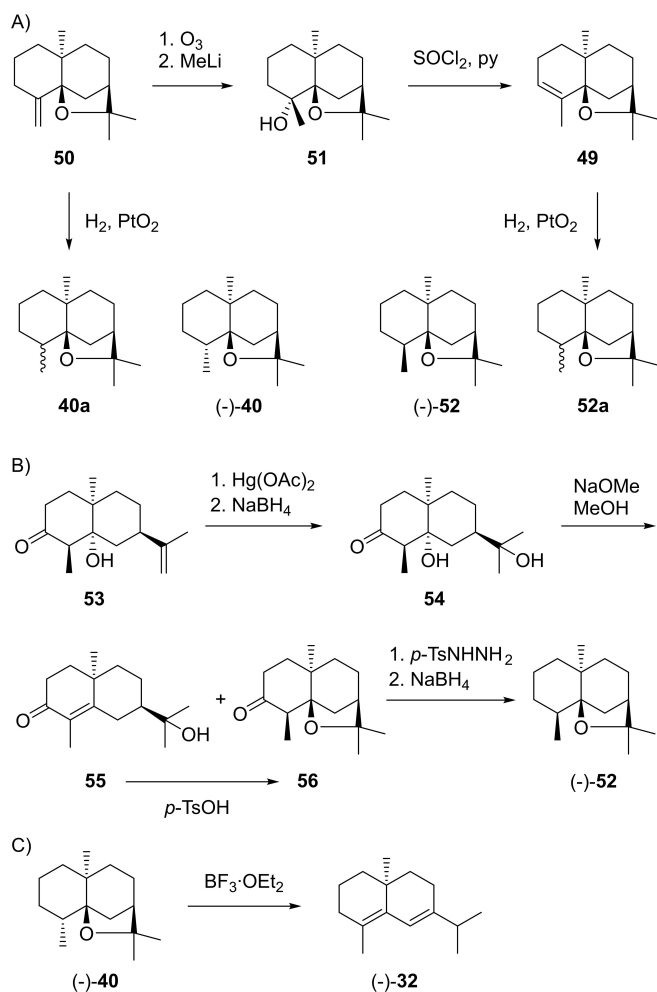
**Scheme 12.** Chemical correlations. A) Synthesis of (–)-**40** from (–)-**38**. B) Synthesis of (–)-**39** from (–)-*ent*-**5**. C) Synthesis of α-agarofuran (**49**) from (–)-10-*epi*-α-cyperone (*ent*-**5**).

rotation that was not commented on ( $[\alpha]_D^{20} = -17.1$ ).<sup>[147]</sup> Also the <sup>13</sup>C NMR data differ substantially,<sup>[144,147]</sup> leaving doubt if the material from *P. arborescens* is indeed identical to the originally isolated rosifoliol. The diol **39** was so far only isolated from *Alpinia japonica* ( $[\alpha]_D = -21.8$ )<sup>[148]</sup> and its structure was secured by synthesis from (–)-10-*epi*-α-cyperone (*ent*-**5**) that proceeded by epoxidation with *m*CPBA and epoxide opening with ketone reduction using LiAlH<sub>4</sub> and AlCl<sub>2</sub>H to yield 10-*epi*-γ-eudesmol (**35**, Scheme 12B). Selective β-epoxidation with VO(acac)<sub>2</sub> and *t*BuOOH followed by epoxide opening with LDA gave **44** that was catalytically hydrogenated with Wilkinson's catalyst to obtain (–)-**39** ( $[\alpha]_D^{10} = -46.2$ ).<sup>[149]</sup>

Dihydro-β-agarofuran (**40**,  $[\alpha]_D^{30} = -77.01$ ) was first isolated from fungus-infected agarwood (*Aquillaria agallocha*) with unknown configuration at C4 and the configurations at C5 and C7 determined wrongly.<sup>[150]</sup> The structure was later revised based on a synthesis from *ent*-**5** that gave the diene **45** upon reduction with LiAlH<sub>4</sub> and pyrolysis in the presence of basic alumina (Scheme 12C). Photosensitised oxygenation to peroxide **46** was followed by isomerisation to the hydroxy

ketone **47** under mildly basic conditions. Treatment with acid-washed  $\text{Al}_2\text{O}_3$  resulted in ring closure to **48**, that upon reduction to a stereoisomeric mixture of allyl alcohols with  $\text{NaBH}_4$ , conversion into the allyl chlorides with  $\text{SOCl}_2$  and reduction with  $\text{LiAlH}_4$  gave  $\alpha$ -agarofuran (**49**).<sup>[151]</sup>

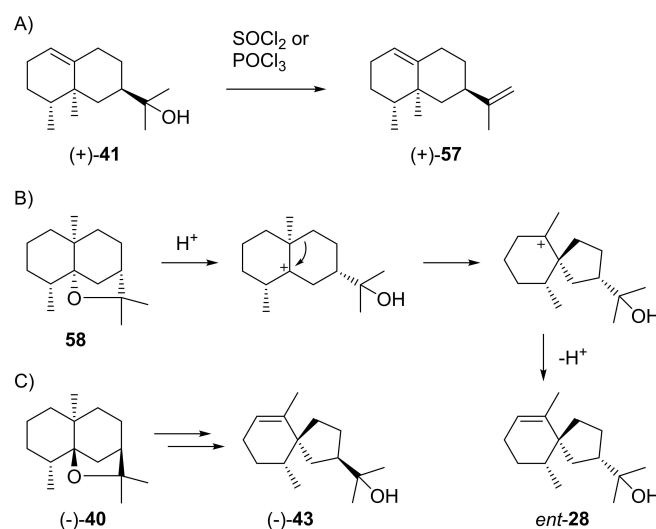
At this stage the previous work had shown that **49** can be obtained from  $\beta$ -agarofuran (**50**) by ozonolysis and addition of  $\text{MeLi}$  to **51**, followed by dehydration with  $\text{SOCl}_2$  in pyridine (Scheme 13A).<sup>[150,152]</sup> It was also known that the catalytic hydrogenation of **49** and **50** leads to materials with slightly different properties, with the compound obtained from **50** being identical to natural (–)-**40**. The two compounds **40a** and **52a** were suggested to be stereoisomers, but their configurations at C4 were unclear.<sup>[150]</sup> A later erroneous correlation with valencene through biotransformation resulted in a confusion of these stereoisomers,<sup>[153,154]</sup> but the situation was ultimately resolved by a synthesis of (–)-isodihydroagarofuran (**52**) from **53** (Scheme 13B).<sup>[155]</sup> This route proceeded through oxymercuration to **54**. Treatment with  $\text{NaOMe}$  in  $\text{MeOH}$  gave a mixture of mainly **55** and small amounts of **56**, with **55** being convertible into **56** under acid catalysis with *p*- $\text{TsOH}$ . Reduction with *p*-



**Scheme 13.** Chemical correlations. A) Conversion of **50** into **49** and catalytic hydrogenations. B) Synthesis of (–)-**52** from **53**. C) Absolute configuration of (–)-**40** by correlation with (–)- $\delta$ -selinene (**32**).

toluenesulfonyl hydrazine and  $\text{NaBH}_4$  resulted in (–)-**52** that was identical to the product obtained by catalytic hydrogenation of **49**, and consequently also the structure of **40** (=4-*epi*-**52**) was secured. The absolute configuration of (–)-**40** was evident from its correlation to (–)- $\delta$ -selinene formed upon treatment with  $\text{BF}_3$  etherate (Scheme 13C).<sup>[150]</sup> The ether (–)-**40** was also isolated from *Galbanum* resin,<sup>[156]</sup> *Alpinia japonica*,<sup>[122]</sup> *Lagera alata*<sup>[157]</sup> and *Vetiveria zizanioides*.<sup>[158]</sup>

(+)-Valerianol (**41**) was first isolated from *Valeriana officinalis* ( $[\alpha]_D^{20} = +134$ ) and its absolute configuration was established by dehydration with  $\text{SOCl}_2$  or  $\text{POCl}_3$ , yielding a hydrocarbon that was identical with (+)-valencene (**57**, Scheme 14A).<sup>[159]</sup> It is also known from *Amyris balsamifera*<sup>[38]</sup> and agarwood,<sup>[160]</sup> and is the main product of the G411 A enzyme variant of *Zea mays* eudesmanediol synthase (ZmEDS).<sup>[71]</sup> Kusunol that was reported from *Cinnamomum camphora* is identical to (+)-**41**.<sup>[161]</sup> (–)-Jinkoheremol (**42**) was first isolated from agarwood and its structure was determined by NMR spectroscopy. Further proof for the assigned structure was given by catalytic hydrogenation that yielded a mixture of the same epimeric dihydro-compounds as obtained from **41**. The absolute configuration was tentatively assigned by comparison of its optical rotation ( $[\alpha]_D = -66$ ) to values for structurally similar compounds,<sup>[160]</sup> but has not been formally established by chemical correlation. (–)-Agarospinol (**43**) was first isolated from *Aquilaria agallocha* ( $[\alpha]_D^{27} = -5.7$ ) with a suggested structure of *ent*-hinesol (*ent*-**28**), based on a biosynthetic relation to dihydro- $\beta$ -agarofuran with the at that time assumed structure of **58** (Scheme 14B). The same paper suggested **43** as an alternative stereochemical representation.<sup>[162]</sup> Notably, after the structural revision of dihydro- $\beta$ -agarofuran to **40**<sup>[151,155]</sup> an analogous biosynthetic relation can indeed explain **43** (Scheme 14C). A synthesis of (*rac*)-**28** also excluded this structure for agarospinol,<sup>[163]</sup> while later syntheses of (*rac*)- and (–)-**43** confirmed its structure and



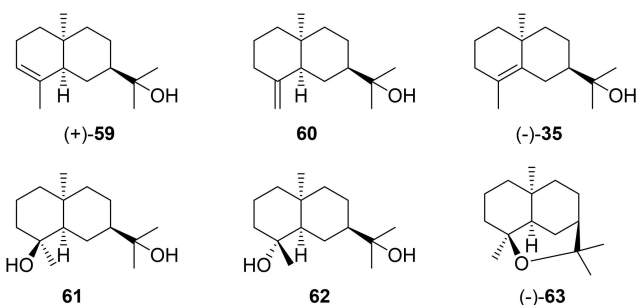
**Scheme 14.** Chemical correlations. A) Dehydration of (+)-**41** to (+)-**57**. B) Hypothetical structure for agarospinol (*ent*-**28**) based on an assumed biosynthetic relation to dihydro- $\beta$ -agarofuran with the initially reported structure of **58**. C) Revised structure of **40** for dihydro- $\beta$ -agarofuran and analogous biosynthetic relation to the correct structure **43** of agarospinol.

absolute configuration.<sup>[164,165]</sup> A later report about agarwood constituents claims a reisolation of (–)-**43**, but shows the structure of *ent*-**28**.<sup>[160]</sup> Neuroleptic properties have been described for **42** and **43** in mice which may be responsible for the sedative effects of agarwood.<sup>[166]</sup>

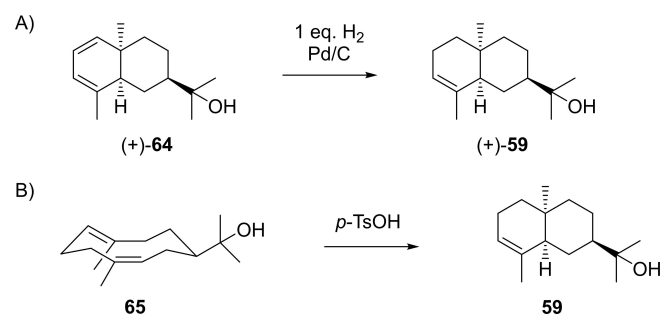
### 3.4. Eudesmols from cation **13**

The structures of the eudesmols that can directly be formed from **13** by deprotonation (**59**, **60** and **35**), capture with water (**61** and **62**) or intramolecular attack of the alcohol to the cation (**63**) are shown in Scheme 15. Compound **35** has already been discussed above as a deprotonation product from **12** (Scheme 10).

(+)-Dihydrooccidentalol (**59**),  $[\alpha]_D^{24} = +59.2$ , is not known as a natural product, but was obtained by catalytic hydrogenation from (+)-occidentalol (**64**, Scheme 16A), a constituent of *Thuja occidentalis*<sup>[167]</sup> for which the structure was assigned by detailed analysis of coupling constants in the <sup>1</sup>H NMR spectrum.<sup>[168]</sup> The compound is also formed from (*Z,E*)-hedycaryol (**65**) upon acid catalysed transannular reaction (Scheme 16B).<sup>[169]</sup> 10-*epi*-β-Eudesmol (**60**) has been isolated from *Bulnesia sarmientoi* with fully established structure by 2-dimensional NMR techniques,<sup>[114]</sup> but neither the optical rotation has been reported nor the absolute configuration has been assigned. The diols **61** and **62** are unknown from natural sources and have only been obtained by synthesis of their racemates.<sup>[170]</sup> The ether (–)-4,11-epoxy-*cis*-eudesmane (**63**,  $[\alpha]_D^{28} = -22$ )<sup>[171]</sup> is a major constituent of the frontal gland



Scheme 15. Eudesmols derived from **13**.



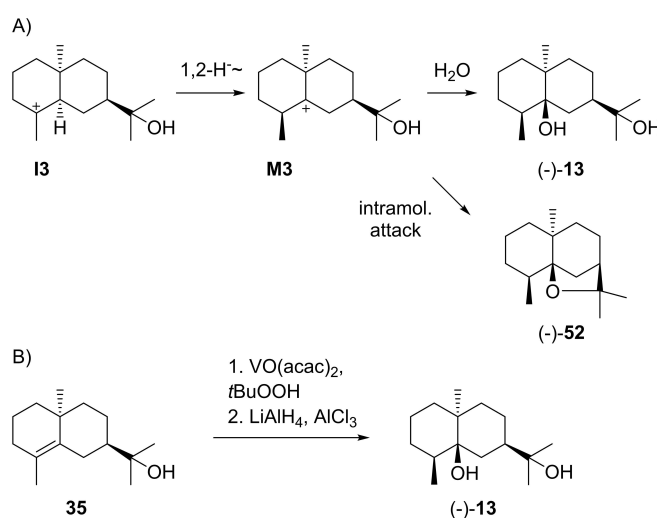
Scheme 16. Chemical correlations. A) Catalytic hydrogenation of (+)-**64**. B) Acid-catalysed conversion of (*Z,E*)-hedycaryol (**65**).

secretions of the termite *Amitermes evuncifer*.<sup>[172]</sup> Its structure was first correctly assigned based on a series of microreactions<sup>[172]</sup> and later confirmed by an enantioselective synthesis from (–)-carvone (**20**).<sup>[171]</sup> Compound **63** was later also isolated from *Amitermes excellens*<sup>[173]</sup> and from *A. minimus*, in which case the paper erroneously shows the opposite absolute configuration, but still reports a negative optical rotation ( $[\alpha]_D^{26} = -34$ ).<sup>[174]</sup> Interestingly, (–)-**63** has a repellent activity against the ant *Crematogaster californica*.<sup>[174]</sup> The same ether **63** is also known from the plant *Phonus arborescens*.<sup>[147]</sup>

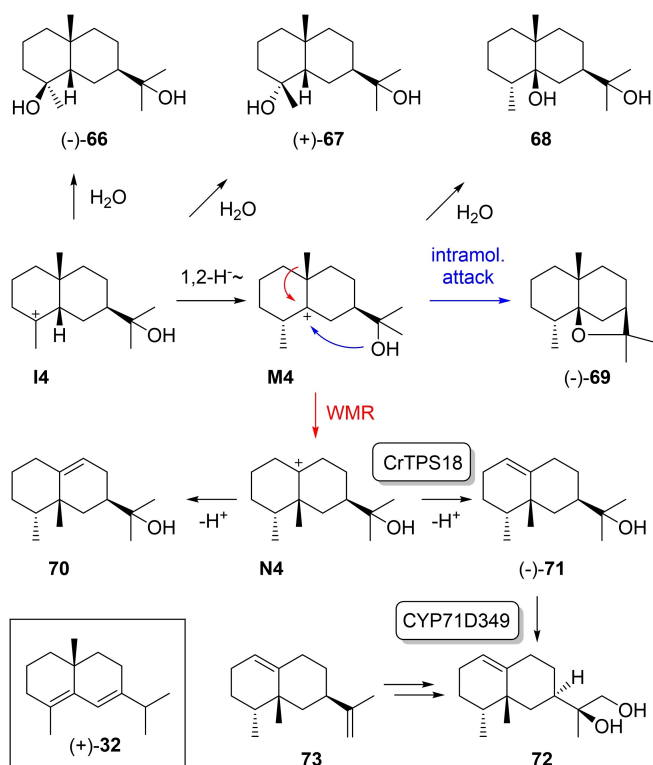
Further compounds from **13** (Scheme 17A) can be reached by a 1,2-hydride shift to **M3** and capture with water to **13** or intramolecular attack of the alcohol to (–)-**52** for which structure elucidation has already been discussed above. The diol **13** ( $[\alpha]_D^{25} = -9.0$ ) was so far only isolated from *Cymbopogon distans* with structure elucidation based on NMR spectroscopy and X-ray crystallography,<sup>[175]</sup> and is the main product of *Zea mays* eudesmanediol synthase (ZmEDS).<sup>[176]</sup> The absolute configuration was evident through a synthesis from **35** (prepared as shown in Scheme 12) by epoxidation and reductive epoxide opening (Scheme 17B).<sup>[149]</sup> Isodihydroagarofuran (**52**), also named α-dihydroagarofuran, was isolated from *Phonus arborescens*,<sup>[147]</sup> *Bursera graveolens*,<sup>[135]</sup> *Bulnesia sarmientoi*,<sup>[177]</sup> and identified in the cyanobacterium *Calothrix* by GC/MS in comparison to standards of **52** and its stereoisomer **40**, albeit without determination of absolute configuration.<sup>[178]</sup>

### 3.5. Eudesmols from cation **14**

Little is known about eudesmols from cation **14** (Scheme 18). The alcohols **66** ( $[\alpha]_D^{20} = -41.1$ ) and **67** ( $[\alpha]_D^{20} = +21.16$ ) were only obtained by synthesis.<sup>[90]</sup> The erroneous assignment of structure **66** to a sesquiterpene diol from *Pluchea arguta* and its structural revision to **36** have been discussed above.<sup>[138,139]</sup> Compounds that are accessible after 1,2-hydride shift to **M4**



Scheme 17. A) Eudesmols derived from **13** and 1,2-hydride shift to **M3**. B) Synthesis of (–)-**13**.



Scheme 18. Eudesmols derived from 14.

include the diol **68** that is unknown from natural sources, but has been obtained by synthesis together with its C4 epimer without further structural assignment regarding the stereochemistry at C4.<sup>[115]</sup> Intramolecular attack of the alcohol function to the cation in **M4** gives access to (–)-*cis*-dihydroagarofuran (**69**) that was so far only isolated from *Prostanthera ovalifolia* ( $[\alpha]_{\text{D}}^{25} = -87.6$ ). Its relative configuration was determined by 2-dimensional NMR techniques and direct comparison to its stereoisomers **40** and **52**, while the absolute configuration was evident from its dehydration to (+)- $\delta$ -selinene (**32**, boxed in Scheme 18).<sup>[179]</sup>

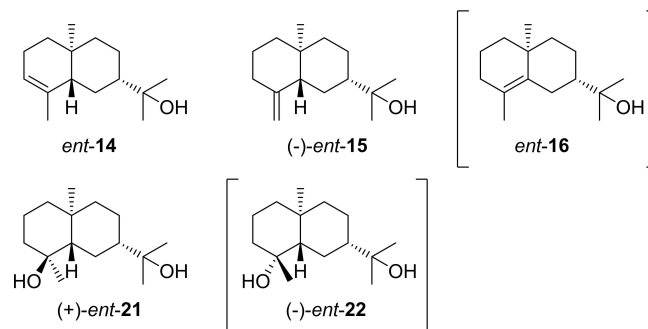
Methyl group migration from **M4** to **N4** and deprotonation gives access to (–)-5-*epi*-jinkoheremol (**71**,  $[\alpha]_{\text{D}}^{25} = -15$ ) for which recently a terpene synthase from *Catharanthus roseus* (CrTPS18) was discovered.<sup>[180]</sup> The absolute configuration of **71** was determined by a comparison of measured to calculated ECD curves. Notably, **71** was shown to be the biosynthetic precursor of debneyol (**72**) by a genetically clustered cytochrome P450 monooxygenase (CYP71D349),<sup>[180]</sup> which is in contrast to the earlier findings for the biosynthesis of **72** that showed incorporation of radioactivity from the sesquiterpene hydrocarbon 5-*epi*-aristolochene (**73**).<sup>[181]</sup> Alternatively, **N4** can be deprotonated to **70**, which is unknown as a natural product, but the racemic compound has been synthesised.<sup>[182]</sup>

### 3.6. Eudesmols from cation 15

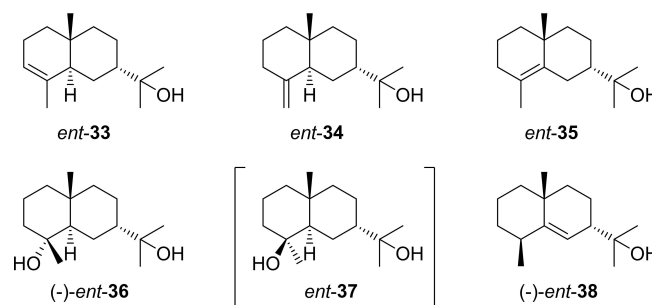
Generally, the number of reports on compounds from the enantiomeric series derived from (–)-hedycaryol through cations **15** - **18** is much lower than those discussed above for (+)-hedycaryol derivatives. Compounds that could biosynthetically directly arise from **15** (Scheme 19) include *ent*- $\alpha$ -eudesmol (*ent*-**14**) for which only one synthetic report is available. Herein, the absolute configuration was secured by MoK $\alpha$  X-ray crystallography of the *p*-bromobenzoate-epoxide of *ent*-**14** (Flack parameter: 0.030(3)) and the optical rotation of *ent*-**14** was found to be positive ( $[\alpha]_{\text{D}}^{25} = +6.4$ )<sup>[183]</sup> which supports the suggested revision of the signs of optical rotation for the enantiomers of **14**.<sup>[84]</sup> The freshwater fungus *Beltriana rhombica* is a source of *ent*-**15** ( $[\alpha]_{\text{D}}^{29} = -37.9$ ),<sup>[184]</sup> and (+)-cryptomeridiol (*ent*-**21**) has been reported from the cypress *Chamaecyparis obtusa*,<sup>[185]</sup> while *ent*-**16** and *ent*-**22** are unknown. No natural products obtained from **15** through 1,2-hydride shift and eventually skeletal rearrangement are known.

### 3.7. Eudesmols from cation 16

Compounds that can directly arise from **16** are summarised on Scheme 20. The sesquiterpene alcohol 7-*epi*- $\alpha$ -eudesmol (*ent*-**33**) was first claimed from *Amyris balsamifera*. The absolute configuration was concluded from the positive optical rotation ( $[\alpha]_{\text{D}} = +10$ ),<sup>[186]</sup> but since at that time no reference data of



Scheme 19. Eudesmols that can directly arise from 15. Compounds in brackets are unknown.

Scheme 20. Eudesmols derived from 16. Compound *ent*-**37** in brackets is unknown.

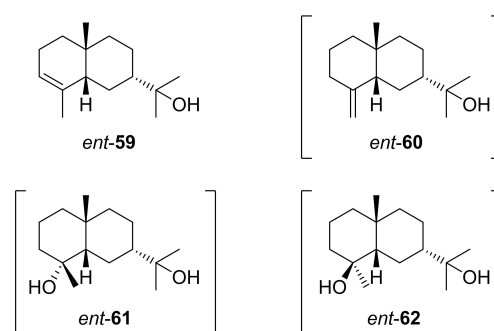


either enantiomer had been reported, the reason for this assignment is unclear. Notably, all other related compounds from this plant have the usual *7R* configuration.<sup>[38]</sup> *7-epi-γ-Eudesmol* (*ent-35*) was first reported with a negative optical rotation ( $[\alpha]_D^{25} = -15$ ) from *Cryptomeria japonica*.<sup>[48]</sup> This work describes structure elucidation by NMR, but does also not explain the reasoning for the assignment of absolute configuration. Subsequently, *ent-33* was also reported from *Laggera alata* without stating the optical rotation, together with *ent-34* and *ent-35* for which again negative optical rotations were given.<sup>[157]</sup> However, this conflicts previous assignments based on enantioselective syntheses of (–)-**34** and, from (+)-dihydrocarvone, of (–)-**35** (cf. Section 3.3.).<sup>[130–132]</sup> The situation becomes even more confusing, because a later synthesis study reported the transformation of (–)-dihydrocarvone into (–)-*ent-35* ( $[\alpha]_D^{10} = -30.1$ ).<sup>[187]</sup> Taken together, the assignments of optical rotations especially to the enantiomers of **35** are doubtful and await future clarification. *7-epi-α-Eudesmol* (**33**) has also been observed as the product of a bacterial sesquiterpene synthase from *Streptomyces viridochromogenes*.<sup>[58,188]</sup> Homologs of this enzyme can be found in many streptomycetes.<sup>[189]</sup> The absolute configuration of **33** from *7-epi-α-eudesmol* synthase is undetermined, but the enantiomer *ent-33* would possibly fit best for a bacterial compound as bacteria often produce the opposite enantiomer as observed in plants.

For isodonsesquitin A from *Isodon grandifolia* the structure of *ent-36* was assigned, but the positive optical rotation ( $[\alpha]_D^{26} = +24.6$ ) is in conflict with this assignment,<sup>[190]</sup> because a total synthesis of both enantiomers gave  $[\alpha]_D^{29} = -66.7$  for *ent-36* and  $[\alpha]_D^{29} = +73.3$  for **36**. The measurements also revealed a strong concentration dependency of these data, but always gave the same sign of optical rotation for the same enantiomer.<sup>[139]</sup> Unfortunately, the isolation paper from *I. grandifolia* did not further discuss the problem of absolute configuration assignment,<sup>[190]</sup> and thus the assignment may likely be in error in this study. After a first assignment of the structure of **67** to a diol from *Pluchea arguta*<sup>[138]</sup> a revision based on synthetic work suggested the compound to be *ent-36*,<sup>[90]</sup> but after synthesis of both enantiomers it was ultimately demonstrated that **36** is the correct structure.<sup>[139]</sup> *Pluchea quitoc* is also a reported source of *ent-36*,<sup>[191]</sup> giving a references to its isolation and first structural revision.<sup>[90,138]</sup> With the correction of the absolute configuration for the compound from *P. arguta*<sup>[139]</sup> it must be concluded that also *P. quitoc* is a producer of **36**. Taken together, despite some discussions about *ent-36* from natural sources in the literature, it seems that this compound is not known as a natural product. Also no reports are available for its C4 epimer *ent-37*. (–)-*ent-Rosifoliol* (*ent-38*) can arise from **16** by 1,2-hydride shift and deprotonation and has been described from the liverwort *Calypogeia muelleriana*.<sup>[192]</sup>

### 3.8. Eudesmols from cation 17

Eudesmols potentially arising from cation **17** are shown in Scheme 21. Starting with a report about the composition of the essential oil from *Elionurus elegans*,<sup>[193]</sup> compound *ent-59* (“5-

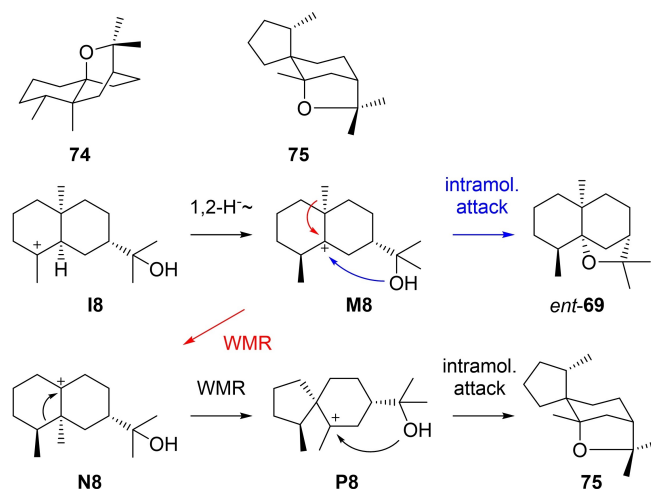


**Scheme 21.** Eudesmols derived from **17**. Compounds in brackets are unknown.

*epi-7-epi-α-eudesmol*) is mentioned in several GC/MS based studies, but has never been isolated, which leaves doubt about the absolute configuration assignment and most if not all these studies may indeed have detected **59** instead. This view is in line with the fact that also neither *ent-60*, *ent-61* and *ent-62* nor any compounds arising from **17** by 1,2-hydride shift and eventually skeletal rearrangement have ever been reported. In summary, no secure reports about natural products from **17** are available.

### 3.9. Eudesmols from cation 18

Only very little is known about eudesmol derivatives arising through cation **18** (Scheme 22). The knowledge is basically limited to the fungal phytotoxin hypodoratoxide. After the initially assigned structure of **74**<sup>[194]</sup> was corrected to that of **75**,<sup>[195]</sup> the biosynthesis was investigated through feeding experiments with isotopically labelled precursors. Starting from **18**, a 1,2-hydride shift leads to **M8** that can be deprotonated to *ent-69*, a cometabolite of **75** in *Hypomyces odoratus*. A methyl migration to **N8**, skeletal rearrangement to **P8** and intra-



**Scheme 22.** Eudesmols derived from **18**.

molecular attack of the alcohol function to the cation result in **75**.<sup>[195]</sup> The absolute configurations of **69** and **75** in *H. odoratus* have not firmly been established.

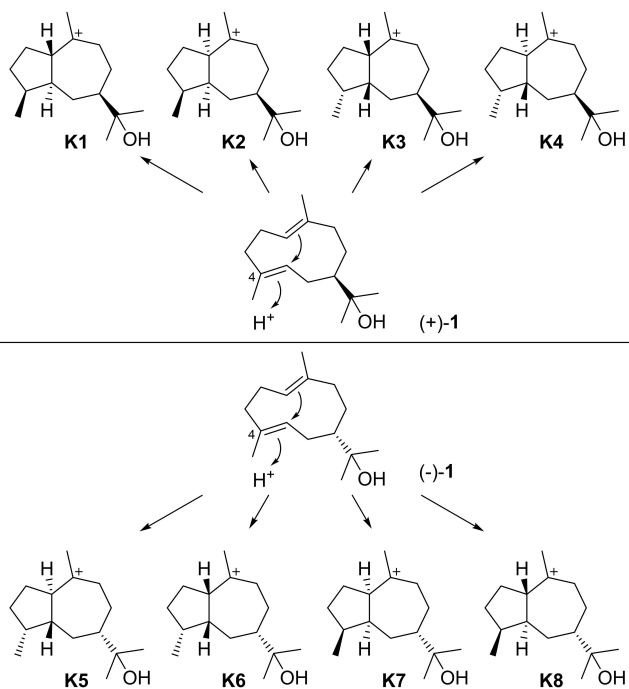
## 4. Guaiols

### 4.1. Cyclisation of hedycaryol by protonation at C4

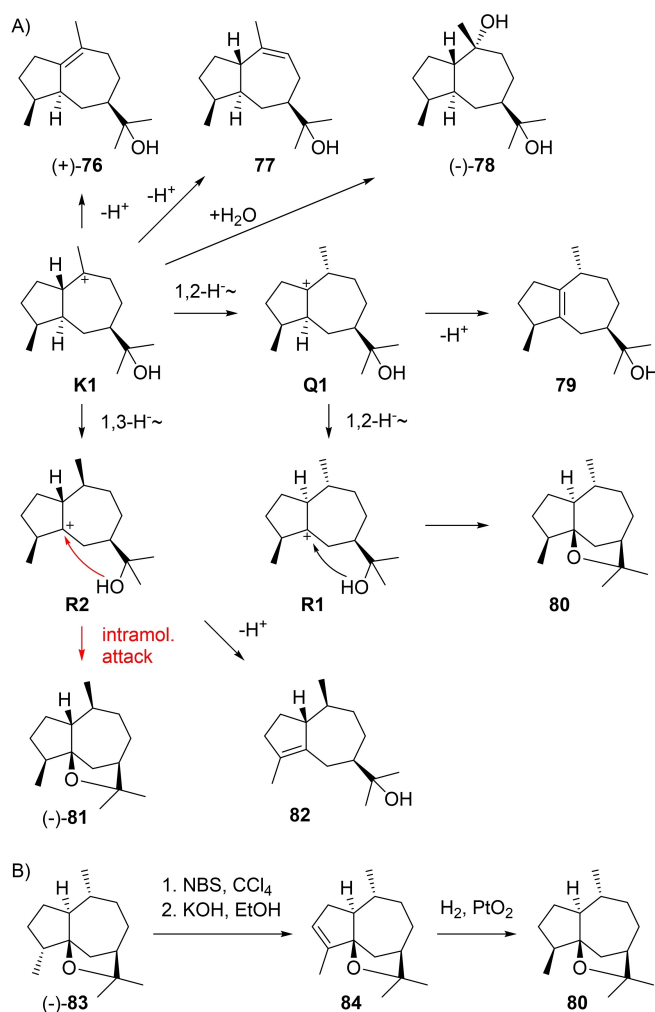
Hedycaryol (+)-**1** can undergo cyclisations through protonation at C4 towards four stereoisomeric intermediates **K1–K4** (Scheme 23). The series of opposite enantiomers **K5–K8** is analogously accessible through protonation induced cyclisations from (–)-**1**, but no natural products with unequivocally established absolute configurations from these intermediates with **7S** configuration appear in the literature. In all cases H5 and Me15 are *trans* to each other because the addition to the *E* configured C4=C5 double bond of hedycaryol is necessarily *anti*. The following sections discuss all known natural products that can be formed from the **K** stereoisomers either directly by deprotonation, capture with water or intramolecular attack of the alcohol function, or after hydride shifts.

### 4.2. Guaiols from cation K1

Guaiols that can be formed directly from **K1** are shown in Scheme 24A. (+)-Bulnesol (**76**) from guaiacwood oil ( $[\alpha]_D^{20} = +3.8$ )<sup>[196]</sup> is one of the most important representatives of the class of guaiols. Its structure was elucidated by Sorm in a correlation to guaiol (**89**, Scheme 25A) that yielded the same hydro-

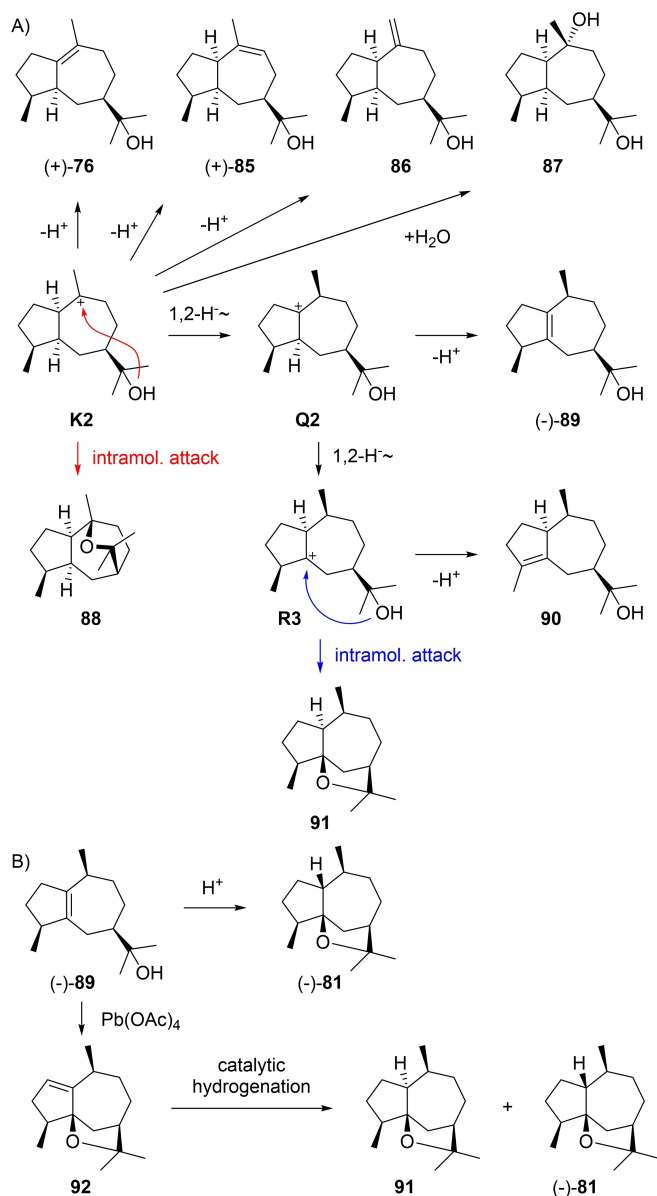


Scheme 23. Cyclisation reactions of **1** induced by reprotonation at C4 towards intermediates **K1–8**.



Scheme 24. A) Guaiols derived from **K1**. B) Conversion of the natural product **83** into its epimer **80**.

genation product as **76**.<sup>[196–198]</sup> It was later also isolated from *Galbanum resin*<sup>[199]</sup> and *Neocallitropsis pancheri*,<sup>[47]</sup> and a sesquiterpene synthase from *Thapsia laciniata* for the production of **76** and **89** as main products (TITPS509) with compound isolation by preparative GC and NMR based structure elucidation was described.<sup>[200]</sup> The alcohol 5 $\alpha$ H-guai-9-en-11-ol (**77**) was recently reported from guaiacwood oil,<sup>[114]</sup> while the diol (–)-**78** ( $[\alpha]_D^{25} = -25.0$ ) is known from the extremophilic fungus *Pithomyces* isolated from a mine waste pit.<sup>[201]</sup> The absolute configuration of **78** has not formally been established yet. Starting from **K1** a 1,2-hydride shift to **Q1** and deprotonation explain **79** that has also recently been found in guaiacwood oil.<sup>[114]</sup> The ether **80** can arise from **Q1** by a second 1,2-hydride shift to **R1** and intramolecular attack of the alcohol function, but is only known as a synthetic compound that was obtained from its 4-epimer (–)-**83**, a known natural product from *Ligularia* ( $[\alpha]_{578} = -45$ , Scheme 24B).<sup>[202]</sup> Bromination at C4 with NBS and elimination gave **84** that upon catalytic hydrogenation yielded **80**,<sup>[202]</sup> thereby completing the set of all eight stereoisomers with **7R** configuration (for discussion of other stereo-



Scheme 25. A) Guaiols derived from K2. B) Chemical correlations of 89 with 81 and 91.

isomers see below). A 1,3-hydride shift from K1 to R2 and deprotonation yield the alcohol 82 from guaiacwood oil,<sup>[114]</sup> while ring closure gives guaioxide (81) that will be discussed in detail in the next section.

#### 4.3. Guaiols from cation K2

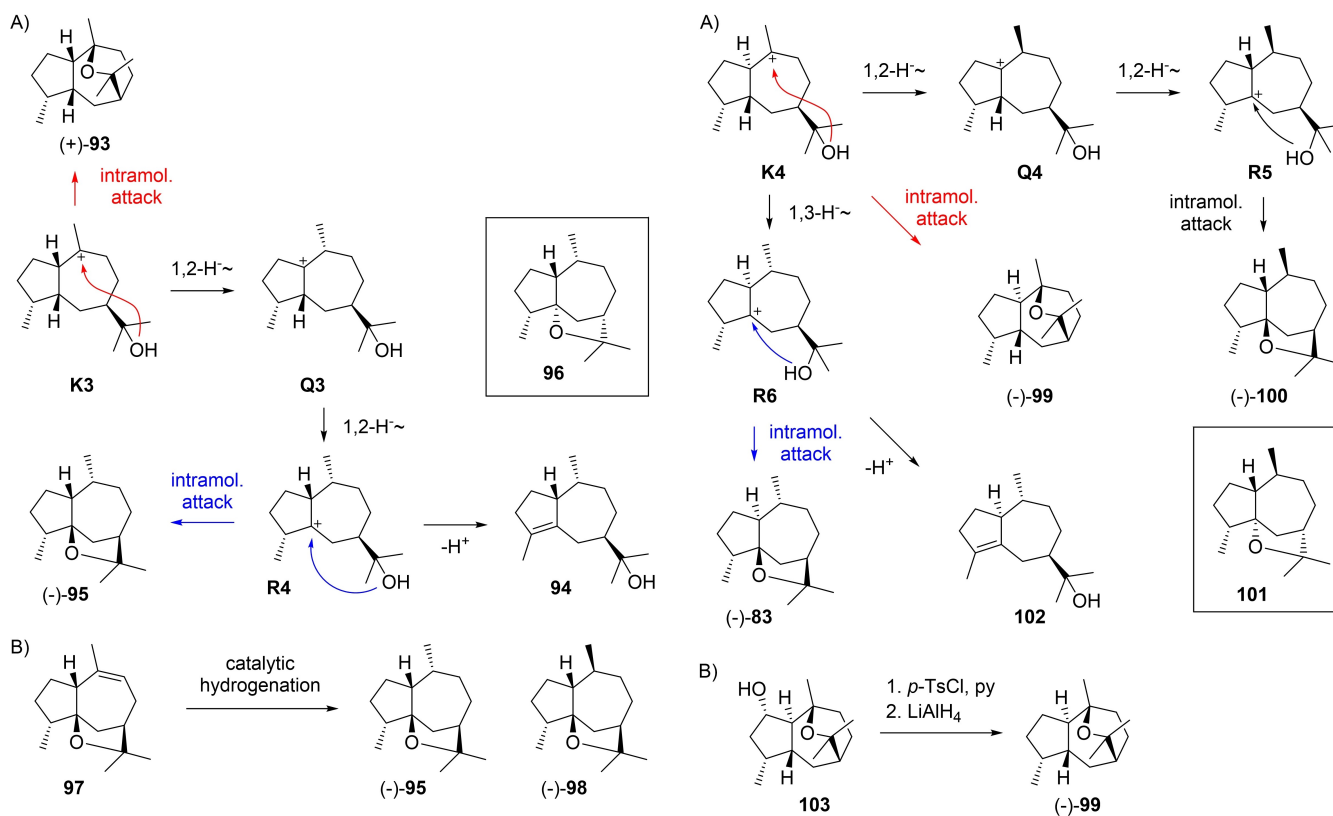
Compounds from K2 are summarised in Scheme 25A. As an alternative to its formation from K1, bulnesol (76) could also be formed from K2 by deprotonation, which may better explain its co-occurrence with guaiol (89), the lead compound from the class of hedycaryol derived 5–7 membered bicyclic sesquiterpene alcohols, that can also be formed from K2 by 1,2-hydride

shift to Q2 and deprotonation. Guaiol was first described from guaiacwood by Gandurin ( $[\alpha]_D^{25} = -26.64$ ) as a bicyclic tertiary alcohol with one double bond.<sup>[203]</sup> The compound is widespread and has also been isolated from *Callitris intratropica*,<sup>[204]</sup> *Eucalyptus maculata*,<sup>[205]</sup> *Drimys lanceolata*,<sup>[206]</sup> *Cinnamomum camphora*,<sup>[207]</sup> *Callitris columellaris*,<sup>[208]</sup> *Guillonea scabra*,<sup>[209]</sup> *Thapsia villosa*,<sup>[210]</sup> *Canarium luzonicum* (Manila elemi),<sup>[211]</sup> *Murraya glenei*,<sup>[212]</sup> *Neocallitropsis pancheri*,<sup>[213]</sup> *Eriostemon fitzgeraldii*,<sup>[214]</sup> *Ferula feruloides*<sup>[215]</sup> and *Uvaria puguensis*,<sup>[216]</sup> and is a product of the above mentioned terpene synthase TITPS509 from *Thapsia laciniata*.<sup>[200]</sup> After establishment of its constitution,<sup>[217]</sup> the absolute configuration was clarified by chemical correlation.<sup>[196,198,218,219]</sup>

Other known compounds that can directly arise from K2 include *cis*-guai-9-en-11-ol (85) from *Galbanum* resin ( $[\alpha]_D^{20} = +4.9$ )<sup>[156]</sup> and from guaiacwood oil that is also a source of 1 $\alpha$ H,5 $\alpha$ H-guai-10(14)-en-11-ol (86) and 10,11-epoxyguaiane (88).<sup>[114,177]</sup> The diol 87 was first isolated from *Leuceria floribunda* with the relative configuration secured by NOE experiments,<sup>[220]</sup> and later reported again from *Jatropha curcas*.<sup>[221]</sup> Starting from Q2, a second 1,2-hydride shift to R3 and deprotonation leads to 90. This compound is known from guaiacwood oil<sup>[114]</sup> and has been synthesised from guaiol (89).<sup>[222]</sup> (-)-Guaioxide (81,  $[\alpha]_D^{24} = -38.2$ ) is easily formed by acid treatment of 89 (Scheme 25B).<sup>[223,224]</sup> It has also been isolated from guaiacwood oil, but may have been formed during the isolation process.<sup>[177]</sup> Its hypothetical biosynthesis requires a 1,3-hydride shift from K1 to R2 and intramolecular attack of the alcohol function (Scheme 24A). The stereoisomer 1-*epi*-guaioxide (91) can arise analogously from R3, but is not known as a natural product (Scheme 25A). Both compounds have been synthesised from 89 by oxidation with Pb(OAc)<sub>4</sub> to yield 92, followed by catalytic hydrogenation to 91 and (-)-81 (Scheme 25B).<sup>[225]</sup> Guaioxide (81) has also been correlated to dihydroguaiol, the hydrogenation product of 89, by a combination of microbial and chemical transformations.<sup>[226]</sup>

#### 4.4. Guaiols from cation K3

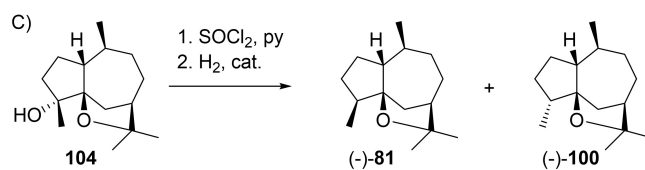
Guaiols from K3 include (+)-isokessane (93) by intramolecular attack of the alcohol (Scheme 26A). This compound has been isolated from *Rubus rosifolius* ( $[\alpha]_D = +19.2$ ) and its structure was elucidated by one and two-dimensional NMR spectroscopy.<sup>[227]</sup> The alcohol 94 is known from guaiacwood oil<sup>[114]</sup> and can arise through a sequence of two 1,2-hydride shifts to Q3 and R4, followed by deprotonation. Alternatively, R4 can react by ring closure to (-)-10-*epi*-liguloxide (95) that has been isolated from *Ligularia* ( $[\alpha]_D = -3.5$ ).<sup>[228]</sup> For this compound initially the structure of 96 (box in Scheme 26A) was assigned, but a later structural revision of liguloxide (98) showed the requirement of a structural revision also of 95,<sup>[229]</sup> because the two compounds are epimers as they are simultaneously formed by catalytic hydrogenation of 97 (Scheme 26B).<sup>[228]</sup>



**Scheme 26.** A) Guaiols derived from **K3**. B) Catalytic hydrogenation of **97** yields the epimers **95** and **98**.

#### 4.5. Guaiols from cation **K4**

Guaiols from **K4** are given in Scheme 27A. A direct ring closure explains the formation of (–)-kessane (**99**) that is known from the roots of several Japanese *Valeriana* species (kesso,  $[\alpha]_D = -7.2$ ).<sup>[230]</sup> Its structure including absolute configuration was established by correlation with known  $\alpha$ -kessyl alcohol (**103**)<sup>[231]</sup> that was converted into **99** by tosylation and treatment with  $\text{LiAlH}_4$  (Scheme 27B),<sup>[230]</sup> and by enantioselective synthesis from (+)-aromadendrene.<sup>[232]</sup> Kessane (**99**) was later isolated again from *Senecio*,<sup>[233–235]</sup> *Bothriochloa intermedia*,<sup>[40]</sup> *Prostanthera ovalifolia*,<sup>[179]</sup> *Olearia phlogopappa*<sup>[236]</sup> and *Machaerium multiflorum*.<sup>[237]</sup> Two sequential 1,2-hydride shifts via **Q4** to **R5** and ring closure give rise to (–)-liguloxide (**100**) from *Ligularia* ( $[\alpha]_D = -52.8$ ).<sup>[228]</sup> Initially, the structure of **101** was assigned to this compound, but elimination of water from **104** and catalytic hydrogenation yielded guaioxide (**81**) and liguloxide (**100**), showing that these compounds must be C4 epimers (Scheme 27C).<sup>[229]</sup> A 1,3-hydride shift from **K4** to **R6** and deprotonation lead to **102** that is observed in guaiacwood oil,<sup>[114]</sup> while intramolecular attack of the alcohol to the cation in **R6** offers an explanation for the biosynthesis of **83** from *Ligularia* (Scheme 27A).<sup>[202]</sup>

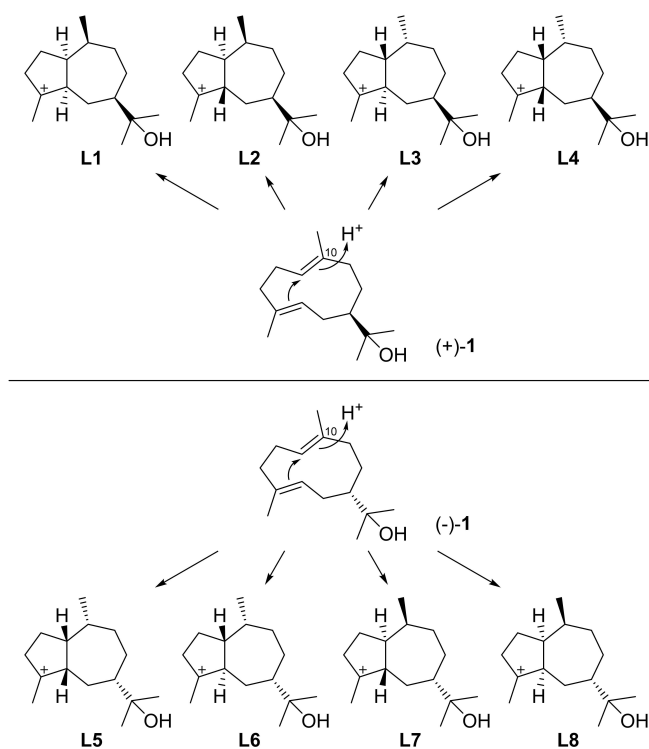


**Scheme 27.** A) Guaiols derived from **K4**. B) Correlation of **103** with **99**. C) Correlation of **104** with **81** and **100**.

#### 4.6. Cyclisation of hedycaryol by protonation at C10

The cyclisation of hedycaryol can also be initiated by protonation at C10 (Scheme 28), leading to the two enantiomeric series of cationic intermediates **L1–L4** from (+)-**1** and **L5–L8** from (–)-**1**. Again, no examples of natural products for the series from (–)-**1** with unambiguously determined absolute configuration are available, and thus the further discussion will be limited to the compounds derived from (+)-**1**.

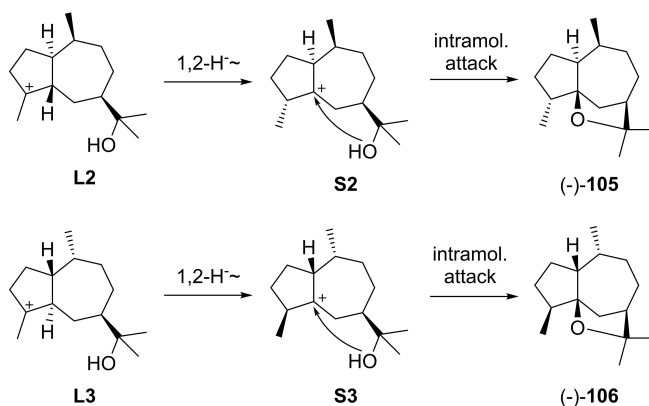
It is interesting to note that subsequent hydride transfers in some cases lead to the same intermediates as discussed above (Scheme 29). Specifically, 1,2-hydride migrations from **L1–L4** result in **S1–S4** and then **T1–T4**. Herein, **S1** and **T1** are equal to **R3** and **Q2** (Scheme 25), while **S4** and **T4** are equal to **R4** and **Q3**, respectively (Scheme 26). Compounds that were already discussed above and could have an alternative biosynthesis along these lines will not be presented here again. Furthermore, **L2** and **L3** can react in 1,3-hydride migrations to **T5** and **T6**, respectively. Analogous steps are sterically not possible for **L1** and **L4**, as was also shown by DFT calculations.<sup>[18]</sup>



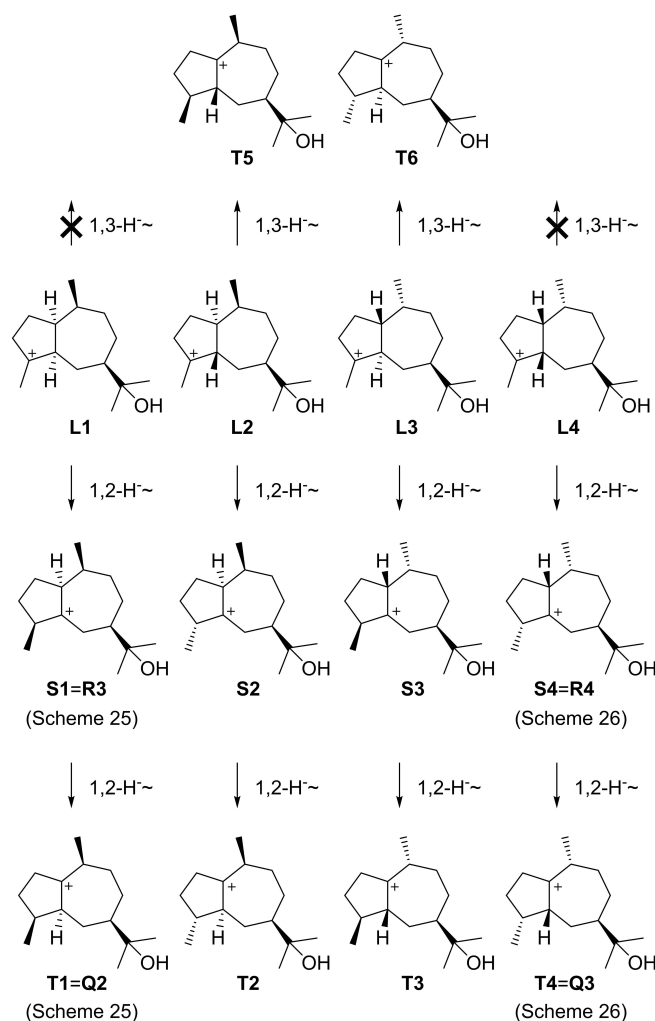
**Scheme 28.** Cyclisation reactions of **1** induced by reprotonation at C10 towards intermediates L1–L8.

#### 4.7. Guaiols potentially arising from hedycaryol by C10 protonation

Notably, most bicyclic 5–7 membered compounds from (+)-**1** can be rationalised through a cyclisation induced by protonation at C4. While the biosynthesis in many cases has not been studied in detail and it is often unknown, whether compounds are formed from (+)-**1** by C4 or C10 protonation, only two more compounds exist whose biosynthesis cannot be easily understood by C4 protonation (Scheme 30). In these cases C10 protonation could more reasonably explain their direct biosynthesis, which could lead to the only two remaining compounds



**Scheme 30.** Compounds **105** and **106** that may arise by C10 protonation of **1**.



**Scheme 29.** Downstream steps from L1–L4 by 1,2- and 1,3-hydride migrations.

(–)-*1-epi*-liguloxide (**105**) and (–)-bulnesoxide (**106**) that will be discussed here.

Starting from **L2**, a 1,2-hydride shift to **S2** and intramolecular attack of the alcohol can give rise to (–)-**105** ( $[\alpha]_D = -25.6$ ),<sup>[238]</sup> while similar reactions from **L3** via **S3** can lead to (–)-**106** ( $[\alpha]_D = -8.2$ ).<sup>[239]</sup> In fact, both compounds were so far only obtained by synthesis,<sup>[238,239]</sup> which questions whether a protonation of (+)-**1** at C10 in a terpene synthase catalysed reaction is relevant for any natural product, as it seems that the formation of all compounds that were isolated from natural sources can be explained through cyclisation of (+)-**1** by C4 protonation and the subsequent reactions discussed above.

## 5. Conclusions

Many natural products are known that biosynthetically arise from hedycaryol (**1**). Plants generally make the compounds derived from (+)-**1**, while bacteria and fungi produce compounds derived from (–)-**1**, and because significantly more

research has been done on plants than on bacteria and fungi, most known compounds originate from (+)-1 and thus have 7*R* configuration. For many compounds, the absolute configurations have been secured by chemical correlations including total synthesis, but sometimes the situation is not fully resolved or even confusing. Particularly the assignments of optical rotations can be erroneous, which can easily happen if impure materials have been measured and the minor contaminants may have large optical rotations of opposite sign in comparison to the investigated compound. Especially the cases of the enantiomers 5-*epi*-10-*epi*- $\gamma$ -eudesmol and 7-*epi*- $\gamma$ -eudesmol that were both synthesised from the enantiomers of dihydrocarvone,<sup>[131,187]</sup> but then both reported to have negative optical rotations, and eventually of  $\alpha$ -eudesmol for which the old work consistently reported a positive optical rotation, while new data support a negative value, deserve a revision.

Cyclisations of hedycaryol can either give a 6–6 membered bicyclic system, which represents the majority of cases. These cyclisations are always induced by protonation at C1, leading to a tertiary cationic intermediate, and not at C4 that would give a less stable and disfavoured secondary cation. Alternatively, a 5–7 membered bicyclic system can be formed for which protonations of 1 at C4 or C10 could potentially be relevant. As we demonstrated here, all compounds can be explained through protonation at C4, with only two remaining cases whose biosynthesis would need C10 protonation, but these compounds are only known as synthetic materials. Therefore, it seems that C4 protonation may serve as the general mechanistic model towards 5–7 bicyclic compounds, and we argue that this is because protonations at the C1=C10 double bond may preferentially happen at C1 to result in the 6–6 membered bicyclic systems. This reflects the situation that we have recently summarised for compounds derived from germacrene A for which the analysis of all known compounds also suggested that protonations of the C1=C10 double bond preferentially happen at C1 with formation of 6–6 membered bicyclic compounds, while protonations at the opposite C4=C5 double bond are directed toward C4 and induce formation of 5–7 membered bicyclic sesquiterpenes.<sup>[17]</sup> Taken together, hedycaryol and germacrene A show – not surprisingly – the same intrinsic reactivity, and the question of forming a 6–6 versus a 5–7 bicyclic ring system is a question of which of the two double bonds in the macrocycle becomes reprotonated. Notably, for patchoulol synthase different mechanisms with C4 and C10 protonation of germacrene A were discussed in the literature,<sup>[240–242]</sup> and a recent mechanistic study from our laboratories has shown that C4 protonation is relevant for this molecule.<sup>[243]</sup> However, clearly more research is required to further confirm the general hypothesis outlined here, because for most compounds the biosynthesis has not been studied experimentally.

## Acknowledgements

We thank Prof. Yoshinori Asakawa (Tokushima Bunri University) for fruitful discussions on the absolute configurations of

eudesmols and their assigned optical rotations. Open Access funding enabled and organized by Projekt DEAL.

## Conflict of Interest

The authors declare no conflict of interest.

## Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

**Keywords:** biosynthesis · configuration determination · hedycaryol · sesquiterpenes · structure elucidation

- [1] J. Gershenzon, N. Dudareva, *Nat. Chem. Biol.* **2007**, *3*, 408–414.
- [2] J. Degenhardt, T. G. Köllner, J. Gershenzon, *Phytochemistry* **2009**, *70*, 1621–1637.
- [3] F. Chen, D. Tholl, J. Bohlmann, E. Pichersky, *Plant J.* **2011**, *66*, 212–229.
- [4] A. Minami, T. Ozaki, C. Liu, H. Oikawa, *Nat. Prod. Rep.* **2018**, *35*, 1330–1346.
- [5] J. S. Dickschat, *Nat. Prod. Rep.* **2016**, *33*, 87–110.
- [6] M. B. Quin, C. M. Flynn, C. Schmidt-Dannert, *Nat. Prod. Rep.* **2014**, *31*, 1449–1473.
- [7] X. Chen, T. G. Köllner, Q. Jia, A. Norris, B. Santhanam, P. Rabe, J. S. Dickschat, G. Shaulsky, J. Gershenzon, F. Chen, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 12132–12137.
- [8] J. S. Dickschat, *Angew. Chem. Int. Ed.* **2019**, *58*, 15964–15976; *Angew. Chem.* **2019**, *131*, 16110–16123.
- [9] H. Xu, J. S. Dickschat, *Synthesis* **2022**, *54*, 1551–1565.
- [10] V. Harms, A. Kirschning, J. S. Dickschat, *Nat. Prod. Rep.* **2020**, *37*, 1080–1097.
- [11] R. Chen, Q. Jia, X. Mu, B. Hu, X. Sun, Z. Deng, F. Chen, G. Bian, T. Liu, *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2023247118.
- [12] V. J. J. Martin, D. J. Pitera, S. T. Withers, J. D. Newman, J. D. Keasling, *Nat. Biotechnol.* **2003**, *21*, 796–802.
- [13] P. Baer, P. Rabe, K. Fischer, C. A. Citron, T. A. Klapschinski, M. Groll, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2014**, *53*, 7652–7656; *Angew. Chem.* **2014**, *126*, 7783–7787.
- [14] Y.-H. Wang, H. Xu, J. Zou, X.-B. Chen, Y.-Q. Zhuang, W.-L. Liu, E. Celik, G.-D. Chen, D. Hu, H. Gao, R. Wu, P.-H. Sun, J. S. Dickschat, *Nat. Catal.* **2022**, *5*, 128–135.
- [15] J. B. Hendrickson, *Tetrahedron* **1959**, *7*, 82–89.
- [16] D. Arigoni, *Pure Appl. Chem.* **1975**, *41*, 219–245.
- [17] H. Xu, J. S. Dickschat, *Chem. Eur. J.* **2020**, *26*, 17318–17341.
- [18] H. Xu, B. Goldfuss, J. S. Dickschat, *Chem. Eur. J.* **2021**, *27*, 9758–9762.
- [19] F. W. Semmler, F. Liao, *Chem. Ber.* **1916**, *49*, 794–798.
- [20] V. Sykora, V. Herout, F. Sorm, *Collect. Czech. Chem. Commun.* **1955**, *20*, 220–226.
- [21] R. O. Hellyer, *Aust. J. Chem.* **1962**, *15*, 157.
- [22] R. V. H. Jones, M. D. Sutherland, *Aust. J. Chem.* **1968**, *21*, 2255–2264.
- [23] R. V. H. Jones, M. D. Sutherland, *J. Chem. Soc. Chem. Commun.* **1968**, 1229–1230.
- [24] A. D. Wagh, S. K. Paknikar, S. C. Bhattacharyya, *Tetrahedron* **1964**, *20*, 2647–2654.
- [25] T. G. Halsall, D. W. Theobald, K. B. Walshaw, *J. Chem. Soc.* **1964**, 1029–1037.
- [26] E. von Rudloff, *Can. J. Chem.* **1963**, *41*, 2876–2881.
- [27] E. von Rudloff, F. M. Couchman, *Can. J. Chem.* **1964**, *42*, 1890–1895.
- [28] Y. Fujise, I. Maruta, S. Ito, T. Nozoe, *Chem. Pharm. Bull.* **1964**, *12*, 991–994.
- [29] G. L. K. Hunter, M. G. Moshonas, *Anal. Chem.* **1965**, *37*, 378–380.
- [30] T.-S. Wu, H. Furukawa, *Chem. Pharm. Bull.* **1983**, *31*, 901–906.
- [31] T.-S. Wu, *Phytochemistry* **1987**, *26*, 3094–3095.
- [32] A. S. Bawdekar, G. R. Kelkar, S. C. Bhattacharyya, *Tetrahedron* **1967**, *23*, 1993–1996.

- [33] H. Shuichi, Y. Katsumi, H. Nanao, M. Tamon, *Bull. Chem. Soc. Jpn.* **1968**, *41*, 1465–1466.
- [34] H.-P. Korthals, D. Merkel, M. Mühlstädt, *Liebigs Ann. Chem.* **1971**, *745*, 39–58.
- [35] R. P. Collins, A. F. Halim, *Planta Med.* **1971**, *20*, 241–243.
- [36] S. G. Agarwal, V. N. Vashist, C. K. Atal, *Phytochemistry* **1974**, *13*, 2024–2025.
- [37] H. Irie, K. Ohno, Y. Ito, S. Uyeo, *Chem. Pharm. Bull.* **1975**, *23*, 1892–1894.
- [38] M. Rohmer, A.-C. Schwartz, R. Anton, *Phytochemistry* **1977**, *16*, 773–774.
- [39] W. M. Bandaranayake, *Phytochemistry* **1980**, *19*, 255–257.
- [40] L. H. Zalkow, J. T. Baxter, R. J. McClure, M. M. Gordon, *J. Nat. Prod.* **1980**, *43*, 598–608.
- [41] C. H. Brieskorn, P. Noble, *Planta Med.* **1982**, *44*, 87–90.
- [42] J. de Pascual, T. S. Vicente, M. S. Gonzalez, I. S. Bellido, *Phytochemistry* **1983**, *22*, 2235–2238.
- [43] M. H. Elgamal, P. Wolff, *Planta Med.* **1987**, *53*, 293–294.
- [44] E. L. Ghisalberti, P. R. Jefferies, H. T. N. Vu, *Phytochemistry* **1990**, *29*, 316–318.
- [45] N. Ruangrunsi, S. Prathanururug, G. L. Lange, M. G. Organ, *Phytochemistry* **1992**, *31*, 2397–2400.
- [46] H. Achenbach, D. Frey, *Phytochemistry* **1992**, *31*, 4263–4274.
- [47] P. Raharivelomanana, J.-P. Bianchini, A. Cambon, M. Azzaro, R. Faure, *Magn. Reson. Chem.* **1995**, *33*, 233–235.
- [48] W.-C. Su, J.-M. Fang, Y.-S. Cheng, *Phytochemistry* **1995**, *39*, 603–607.
- [49] A. L. Marsaro, R. C. Souza, T. M. C. Della Lucia, J. B. Fernandes, M. F. G. F. Silva, P. C. Vieira, *J. Chem. Ecol.* **2004**, *30*, 1771–1780.
- [50] I. A. Southwell, *Phytochemistry* **1970**, *9*, 2243–2245.
- [51] I. A. Southwell, *Aust. J. Chem.* **1978**, *31*, 2527–2538.
- [52] S. Hasegawa, Y. Hirose, *Phytochemistry* **1981**, *20*, 508–510.
- [53] E. Stahl, *Planta Med.* **1984**, *50*, 157–160.
- [54] S. Nagahama, M. Tazaki, H. Kobayashi, M. Sumimoto, *Phytochemistry* **1993**, *33*, 879–882.
- [55] T. Hieda, M. Tazaki, Y. Morishita, T. Aoki, S. Nagahama, *Phytochemistry* **1996**, *42*, 159–162.
- [56] H. Xu, N. Lackus, T. G. Köllner, J. S. Dickschat, *Org. Lett.* **2022**, *24*, 587–591.
- [57] H. Xu, J. Rinkel, J. S. Dickschat, *Org. Chem. Front.* **2021**, *8*, 1177–1184.
- [58] P. Rabe, T. Schmitz, J. S. Dickschat, *Beilstein J. Org. Chem.* **2016**, *12*, 1839–1850.
- [59] L. Ding, H. Goerls, K. Dornblut, W. Lin, A. Maier, H.-H. Fiebig, C. Hertweck, *J. Nat. Prod.* **2015**, *78*, 2963–2967.
- [60] Y. Terada, S. Yamamuro, *Tetrahedron Lett.* **1979**, *20*, 3303–3306.
- [61] J. A. Faraldos, S. Wu, J. Chappell, R. M. Coates, *Tetrahedron* **2007**, *63*, 7733–7742.
- [62] A. J. Minnaard, J. B. P. A. Wijnberg, A. de Groot, *Tetrahedron* **1994**, *50*, 4755–4764.
- [63] S. Irmsch, Y. Jiang, F. Chen, J. Gershenzon, T. G. Köllner, *BMC Plant Biol.* **2014**, *14*, 270.
- [64] J. Hattan, K. Shindo, T. Ito, Y. Shibuya, A. Watanabe, C. Tagaki, F. Ohno, T. Sasaki, J. Ishii, A. Kondo, N. Misawa, *Planta* **2016**, *243*, 959–972.
- [65] L. Chuang, C.-S. Wen, Y.-R. Lee, Y.-L. Lin, L.-R. Hsu, S.-Y. Wang, F.-H. Chu, *J. Nat. Prod.* **2018**, *81*, 1162–1172.
- [66] P. Baer, P. Rabe, C. A. Citron, C. C. de Oliveira Mann, N. Kaufmann, M. Groll, J. S. Dickschat, *ChemBioChem* **2014**, *15*, 213–216.
- [67] Z. Li, Y. Jiang, X. Zhang, Y. Chang, S. Li, X. Zhang, S. Zheng, C. Geng, P. Men, L. Ma, Y. Yang, Z. Gao, Y.-J. Tang, S. Li, *ACS Catal.* **2020**, *10*, 5846–5851.
- [68] P. Rabe, J. Rinkel, E. Dolja, T. Schmitz, B. Nubbemeyer, T. H. Luu, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2017**, *56*, 2776–2779; *Angew. Chem.* **2017**, *129*, 2820–2823.
- [69] K. Schriever, P. Saenz-Mendez, R. S. Rudraraju, N. M. Hendrikse, E. P. Hudson, A. Biundo, R. Schnell, P.-O. Syren, *J. Am. Chem. Soc.* **2021**, *143*, 3794–3807.
- [70] D. P. Piet, A. J. Minnaard, K. A. van der Heyden, M. C. R. Franssen, J. B. P. A. Wijnberg, A. de Groot, *Tetrahedron* **1995**, *51*, 243–254.
- [71] J. Liang, L. Wang, J. Liu, Q. Shen, J. Fu, R. J. Peters, Q. Wang, *Biochemistry* **2020**, *59*, 2660–2666.
- [72] P. S. Wharton, C. E. Sundin, D. W. Johnson, H. C. Kluender, *J. Org. Chem.* **1972**, *37*, 34–38.
- [73] M. Kodama, S. Yokoo, Y. Matsuki, S. Ito, *Tetrahedron Lett.* **1979**, *20*, 1687–1690.
- [74] L. Ruzicka, A. H. Wind, D. R. Koolhaas, *Helv. Chim. Acta.* **1931**, *14*, 1132–1151.
- [75] F. J. McQuillin, J. D. Parrack *J. Chem. Soc.* **1956**, 2973–2978.
- [76] B. Riniker, J. Kalvoda, D. Arigoni, A. Fürst, O. Jeger, A. M. Gold, R. B. Woodward, *J. Am. Chem. Soc.* **1954**, *76*, 313–314.
- [77] A. R. Pinder, R. A. Williams, *J. Chem. Soc.* **1963**, 2773–2778.
- [78] D. C. Humber, A. R. Pinder, R. A. Williams, *J. Org. Chem.* **1967**, *32*, 2335–2340.
- [79] R. O. Hellyer, H. H. G. McKern, *Aust. J. Chem.* **1963**, *16*, 515–519.
- [80] R. O. Hellyer, H. H. G. McKern, *Aust. J. Chem.* **1966**, *19*, 1541–1543.
- [81] E. von Rudloff, G. V. Nair, *Can. J. Chem.* **1964**, *42*, 421–425.
- [82] R. B. Bates, E. K. Hendrickson, *Chem. Ind.* **1962**, 1759–1760.
- [83] M. K. Seikel, J. W. Rowe, *Phytochemistry* **1964**, *3*, 27–32.
- [84] M. Toyota, Y. Yonehara, I. Horibe, K. Minagawa, Y. Asakawa, *Phytochemistry* **1999**, *52*, 689–694.
- [85] Y. Aoyama, Y. Araki, T. Konoike, *Synlett* **2001**, *9*, 1452–1454.
- [86] N. Kumar, B. Ravindranath, T. R. Seshadri, *Phytochemistry* **1974**, *13*, 633–636.
- [87] H. Erdtman, B. R. Thomas, *Acta Chem. Scand.* **1958**, *12*, 267–273.
- [88] L. Dolejs, V. Herout, *Collect. Czech. Chem. Commun.* **1961**, *26*, 2045–2049.
- [89] M. Sumimoto, H. Ito, H. Hirai, K. Wada, *Chem. Ind.* **1963**, *19*, 780–781.
- [90] M. Ando, K. Arai, K. Kikuchi, K. Isogai, *J. Nat. Prod.* **1994**, *57*, 1189–1199.
- [91] L. Sumimoto, *Chem. Ind.* **1963**, *34*, 1436.
- [92] F. M. Abdel-Moneim, Z. F. Ahmed, M. B. E. Fayed, H. Ghaleb, *Planta Med.* **1969**, *17*, 209–216.
- [93] A. S. Radwan, *Planta Med.* **1975**, *27*, 93–97.
- [94] F. E. Evans, D. W. Miller, T. Cairns, G. V. Baddeley, E. Wenkert, *Phytochemistry* **1982**, *21*, 937–938.
- [95] H. D. Locksley, M. B. E. Fayed, A. S. Radwan, V. M. Chari, G. A. Cordell, H. Wagner, *Planta Med.* **1982**, *45*, 20–22.
- [96] W. Renold, G. Ohloff, T. Norin, *Helv. Chim. Acta* **1979**, *62*, 985–993.
- [97] M. A. Irwin, T. A. Geissman, *Phytochemistry* **1973**, *12*, 849–852.
- [98] M. Fujita, H. Itokawa, Y. Sashida, *Yakugaku Zasshi* **1973**, *93*, 422–428.
- [99] A. Cruz, M. Silva, P. G. Sammes, *Phytochemistry* **1973**, *12*, 2549–2550.
- [100] S. C. Sharma, Y. N. Shukla, J. S. Tandon, *Phytochemistry* **1975**, *14*, 578–579.
- [101] S. Hasegawa, Y. Hirose, *Phytochemistry* **1981**, *20*, 508–510.
- [102] H. Achenbach, R. Waibel, I. Addae-Mensah, *Phytochemistry* **1985**, *24*, 2325–2328.
- [103] R. Mata, A. Navarrete, L. Alvarez, R. Pereda-Miranda, G. Delgado, A. Romo de Vivar, *Phytochemistry* **1987**, *26*, 191–193.
- [104] M. Yatagai, T. Takahashi, T. Sato, *J. Essent. Oil Res.* **1994**, *6*, 459–461.
- [105] K.-S. Lee, G. Li, S. H. Kim, C.-S. Lee, M.-H. Woo, S.-H. Lee, Y.-D. Jhang, J.-K. Son, *J. Nat. Prod.* **2002**, *65*, 1707–1708.
- [106] I. Werner, P. Mucaji, A. Presser, S. Glasl, *Z. Naturforsch.* **2007**, *62b*, 267–271.
- [107] Y. Tong, P. Su, H. Guan, T. Hu, J. Chen, Y. Zhang, Y. Zhao, L. Gao, X. Zhang, L. Huang, W. Gao, *Biochem. J.* **2018**, *475*, 2713–2725.
- [108] C. P. Cornwell, N. Reddy, D. N. Leach, S. G. Wyllie, *Flavour Fragrance J.* **2000**, *15*, 421–431.
- [109] C. Y. Ragasa, A. L. K. C. Co, J. A. Rideout, *Nat. Prod. Res.* **2005**, *19*, 231–237.
- [110] N. P. D. Nanayakkara, A. D. Kinghorn, N. R. Farnsworth, *J. Chem. Res.* **1986**, *12*, 454–455.
- [111] T. Kikuchi, K. Watanabe, Y. Tochigi, A. Yamamoto, M. Fukatsu, Y. Ezaki, R. Tanaka, T. Akihisa, *Chem. Biodiversity* **2012**, *9*, 1500–1507.
- [112] J. Youkwan, S. Sutthivaiyakit, P. Sutthivaiyakit, *J. Nat. Prod.* **2010**, *73*, 1879–1883.
- [113] A. Bianchini, F. Tomi, P. Richomme, A.-F. Bernardini, J. Casanova, *Magn. Reson. Chem.* **2004**, *42*, 983–984.
- [114] L. Tissandie, S. Viciano, H. Brevard, U. J. Meierhenrich, J.-J. Filippi, *Phytochemistry* **2018**, *149*, 64–81.
- [115] I. Kitagawa, Y. Yamazoe, H. Shibuya, R. Takeda, H. Takeno, I. Yosioka, *Chem. Pharm. Bull.* **1974**, *22*, 2662–2674.
- [116] J.-F. Cavalli, F. Tomi, A.-F. Bernardini, J. Casanova, *Magn. Reson. Chem.* **2004**, *42*, 709–711.
- [117] J. J. Brophy, R. J. Goldsack, J. M. Hook, C. J. R. Fookes, P. I. Forster, *J. Essent. Oil Res.* **2004**, *16*, 362–366.
- [118] H. Ishii, T. Tozayo, H. Minato, *J. Chem. Soc. C* **1966**, 1545–1548.
- [119] F. Bohlmann, C. Zdero, *Phytochemistry* **1978**, *17*, 1135–1153.
- [120] P. Weyerstahl, H. Marschall-Weyerstahl, H.-C. Wahlburg, *Liebigs Ann. Chem.* **1989**, 307–308.
- [121] L. H. Zalkow, G. L. Chetty, *J. Org. Chem.* **1975**, *40*, 1833–1834.
- [122] H. Itokawa, H. Morita, K. Watanabe, S. Mihashi, Y. Iitaka, *Chem. Pharm. Bull.* **1985**, *33*, 1148–1153.
- [123] I. Yosioka, S. Takahashi, H. Hikino, Y. Sasaki, *Chem. Pharm. Bull.* **1959**, *7*, 319–323.

- [124] I. Yosioka, Y. Sasaki, H. Hikino, *Chem. Pharm. Bull.* **1961**, *9*, 84–85.
- [125] I. Yosioka, T. Kimura, *Chem. Pharm. Bull.* **1969**, *17*, 856–857.
- [126] H. Itokawa, H. Nakanishi, S. Mihashi, *Chem. Pharm. Bull.* **1983**, *31*, 1991–1997.
- [127] Y. Du, X. Lu, *J. Org. Chem.* **2003**, *68*, 6463–6465.
- [128] K. Otoguro, M. Iwatsuki, A. Ishiyama, M. Namatame, A. Nishihara-Tukushima, H. Kiyohara, T. Hashimoto, Y. Asakawa, S. Omura, H. Yamada, *Phytochemistry* **2011**, *72*, 2024–2030.
- [129] M. A. Schwartz, A. M. Willbrand, *J. Org. Chem.* **1985**, *50*, 1359–1365.
- [130] L. H. Zalkow, M. Smith, G. L. Chetty, A. W. Shaligram, P. Ingwalson, *J. Org. Chem.* **1976**, *41*, 3710–3714.
- [131] J. A. Marshall, M. T. Pike, *J. Org. Chem.* **1968**, *33*, 435–437.
- [132] R. Kaiser, P. Naegeli, *Tetrahedron Lett.* **1972**, *20*, 2009–2012.
- [133] T. Nakanishi, E. Yamagata, K. Yoneda, T. Nagashima, I. Kawasaki, T. Yoshida, H. Mori, I. Miura, *Phytochemistry* **1984**, *23*, 2066–2067.
- [134] A. T. Bottini, D. J. Garfagnoli, L. S. Delgado, V. Dev, S. T. Duong, C. G. Kelley, R. Keyer, R. Raffel, P. Joshi, C. S. Mathela, *J. Nat. Prod.* **1987**, *50*, 732–734.
- [135] C. Yukawa, H. Iwabuchi, T. Kamikawa, S. Komemushi, A. Sawabe, *Flavour Fragrance J.* **2004**, *19*, 565–570.
- [136] J. W. Huffman, C. A. Miller, A. R. Pinder, *J. Org. Chem.* **1976**, *41*, 3705–3710.
- [137] J. Jakupovic, U. Ganzer, P. Pritschow, L. Lehmann, F. Bohlmann, R. M. King, *Phytochemistry* **1992**, *31*, 863–880.
- [138] V. U. Ahmad, T. A. Farooqui, K. Fizza, A. Sultana, R. Khatoon, *J. Nat. Prod.* **1992**, *55*, 730–735.
- [139] F. Shimoma, H. Kondo, S. Yuuya, T. Suzuki, H. Hagiwara, M. Ando, *J. Nat. Prod.* **1998**, *61*, 22–28.
- [140] S.-L. Li, J.-K. Ding, *Acta Bot. Yunnan.* **1993**, *15*, 303–305.
- [141] Y. Zhao, J. Yue, Z. Lin, J. Ding, H. Sun, *Phytochemistry* **1997**, *44*, 459–464.
- [142] W.-M. Zhu, Q. Zhao, S.-L. Li, X.-J. Hao, *J. Asian Nat. Prod. Res.* **2007**, *9*, 277–283.
- [143] R. P. T. Kim, V. Bihud, K. bin Mohamad, K. H. Leong, J. bin Mohamad, F. bin Ahmad, H. Hazni, N. Kasim, S. N. A. Halim, K. Awang, *Molecules* **2013**, *18*, 128–139.
- [144] I. A. Southwell, *Tetrahedron Lett.* **1977**, *18*, 873–876.
- [145] D. F. MacSweeney, R. Ramage, R. Sattar, *Tetrahedron Lett.* **1970**, *11*, 557–560.
- [146] B. Beagley, R. G. Pritchard, R. Ramage, I. A. Southwell, *Acta Crystallogr.* **1982**, *B38*, 1391–1393.
- [147] A. F. Barrero, P. Artega, J. F. Quilez, I. Rodriguez, M. M. Herrador, *J. Nat. Prod.* **1997**, *60*, 1026–1030.
- [148] H. Itokawa, H. Morita, K. Watanabe, *Chem. Pharm. Bull.* **1987**, *35*, 1460–1463.
- [149] Z. Xiong, G. Zhou, J. Yang, Y. Chen, Y. Li, *Tetrahedron: Asymmetry* **1998**, *9*, 1525–1530.
- [150] M. L. Maheshwari, T. C. Jain, R. B. Bates, S. C. Bhattacharyya, *Tetrahedron* **1963**, *19*, 1079–1090.
- [151] H. C. Barrett, G. Buechi, *J. Am. Chem. Soc.* **1967**, *89*, 5665–5667.
- [152] M. L. Maheshwari, K. R. Varma, S. C. Bhattacharyya, *Tetrahedron* **1963**, *19*, 1519–1525.
- [153] S. K. Paknikar, C. G. Naik, *Tetrahedron Lett.* **1975**, *16*, 1293–1294.
- [154] A. F. Thomas, M. Ozainne, *Tetrahedron Lett.* **1976**, *17*, 1717–1718.
- [155] A. Asselin, M. Mongrain, P. Deslongchamps, *Can. J. Chem.* **1968**, *46*, 2817–2820.
- [156] A. F. Thomas, M. Ozainne, *Helv. Chim. Acta* **1978**, *61*, 2874–2880.
- [157] P. Raharivelomanana, J.-P. Bianchini, A. R. P. Ramanoelina, J. R. E. Rasoarahona, R. Faure, A. Cambon, *Phytochemistry* **1998**, *47*, 1085–1088.
- [158] P. Weyerstahl, H. Marschall, U. Splittgerber, D. Wolf, H. Surburg, *Flavour Fragrance J.* **2000**, *15*, 395–412.
- [159] G. Jommi, J. Krepinsky, V. Herout, F. Sorm, *Tetrahedron Lett.* **1967**, *8*, 677–681.
- [160] T. Nakanishi, E. Yamagata, K. Yoneda, *J. Chem. Soc. Perkin Trans. 1* **1983**, 601–604.
- [161] H. Hikino, N. Suzuki, T. Takemoto, *Chem. Pharm. Bull.* **1968**, *16*, 832–838.
- [162] K. R. Varma, M. L. Maheshwari, S. C. Bhattacharyya, *Tetrahedron* **1965**, *21*, 115–138.
- [163] J. A. Marshall, S. F. Brady, *Tetrahedron Lett.* **1969**, *10*, 1387–1390.
- [164] M. Mongrain, J. Lafontaine, A. Belanger, P. Deslongchamps, *Can. J. Chem.* **1970**, *48*, 3273–3274.
- [165] M. Deighton, C. R. Hughes, R. Ramage, *J. Chem. Soc. Chem. Commun.* **1975**, 662–663.
- [166] H. Okugawa, R. Ueda, K. Matsumoto, K. Kawanishi, A. Kato, *Planta Med.* **1996**, *62*, 2–6.
- [167] T. Nakatsuka, Y. Hirose, *Bull. Agric. Chem. Soc. Jpn.* **1956**, *20*, 215–218.
- [168] A. G. Hortmann, J. B. de Roos, *J. Org. Chem.* **1969**, *34*, 736–738.
- [169] M. Kodama, S. Yokoo, Y. Matsuki, S. Ito, *Tetrahedron Lett.* **1979**, *20*, 1687–1690.
- [170] T. Hatsui, K. Li, A. Mori, H. Takeshita, *Heterocycles* **1970**, *54*, 765–776.
- [171] R. Baker, D. A. Evans, P. G. McDowell, *J. Chem. Soc. Chem. Commun.* **1977**, 111.
- [172] L. J. Wadhams, R. Baker, P. E. Howse, *Tetrahedron Lett.* **1974**, *15*, 1697–1700.
- [173] Y. Naya, G. D. Prestwich, S. G. Spanton, *Tetrahedron Lett.* **1982**, *23*, 3047–3050.
- [174] R. H. Scheffrahn, J. J. Sims, L. K. Gaston, M. K. Rust, *Experientia* **1984**, *40*, 1136–1137.
- [175] C. S. Mathela, A. B. Melkani, A. Pant, V. Dev, T. E. Nelson, H. Hope, A. T. Bottini, *Phytochemistry* **1989**, *28*, 936–938.
- [176] J. Liang, J. Liu, R. Brown, M. Jia, K. Zhou, R. J. Peters, Q. Wang, *Plant J.* **2018**, *94*, 847–856.
- [177] L. Tissandie, M. Gaysinski, H. Brevard, U. J. Meierhenrich, J.-J. Filippi, *J. Nat. Prod.* **2017**, *80*, 526–537.
- [178] C. Höckelmann, P. G. Becher, S. H. von Reuß, F. Jüttner, *Z. Naturforsch.* **2009**, *64c*, 49–55.
- [179] I. A. Southwell, D. J. Tucker, *Phytochemistry* **1993**, *33*, 857–862.
- [180] J. Liang, T. An, J.-X. Zhu, S. Chen, J.-H. Zhu, R. J. Peters, R. Yu, J. Zi, *J. Nat. Prod.* **2021**, *84*, 2709–2716.
- [181] I. M. Whitehead, D. R. Threlfall, D. F. Ewing, *Phytochemistry* **1989**, *28*, 775–779.
- [182] Y. Zhao, D. J. Schenk, S. Takahashi, J. Chappell, R. M. Coates, *J. Org. Chem.* **2004**, *69*, 7428–7435.
- [183] K. Ota, K. Kamaike, H. Miyaoka, *Synthesis* **2022**, *54*, 689–696.
- [184] V. Rukachaisirikul, C. Kaewbumrung, S. Phongpaichit, Z. Hajiwangoh, *Chem. Pharm. Bull.* **2005**, *53*, 238–240.
- [185] M. Kuroyanagi, R. Ikeda, H. Y. Gao, N. Muto, K. Otaki, T. Sano, N. Kawahara, T. Nakane, *Chem. Pharm. Bull.* **2008**, *56*, 60–63.
- [186] T. A. van Beek, R. Kleis, M. A. Posthumus, A. van Veldhuizen, *Phytochemistry* **1989**, *28*, 1909–1911.
- [187] Z. Zhang, G. Zhou, X. Gao, Y. Li, R. Liao, *Synth. Commun.* **2001**, *31*, 847–851.
- [188] P. Rabe, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2013**, *52*, 1810–1812; *Angew. Chem.* **2013**, *125*, 1855–1857.
- [189] L. Martin-Sanchez, K. S. Singh, M. Avalos, G. P. van Wezel, J. S. Dickschat, P. Garbeva, *Beilstein J. Org. Chem.* **2019**, *15*, 1181–1193.
- [190] W. Shun-Hua, J.-M. Fang, L. Zhong-Wen, S. Han-Dong, *Phytochemistry* **1993**, *34*, 1176–1178.
- [191] G. M. S. P. Guilhon, A. H. Müller, *Phytochemistry* **1998**, *47*, 227–229.
- [192] U. Warmers, K. Wihstutz, N. Bülow, C. Fricke, W. A. König, *Phytochemistry* **1998**, *49*, 1723–1731.
- [193] J. P. Mevy, J.-M. Bessiere, M. Dherbomez, J. Viano, *J. Agric. Food Chem.* **2002**, *50*, 4240–4243.
- [194] B. Kühne, H.-P. Hanssen, W.-R. Abraham, V. Wray, *Phytochemistry* **1991**, *30*, 1463–1465.
- [195] L. Barra, K. Ibrom, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2015**, *54*, 6637–6640; *Angew. Chem.* **2015**, *127*, 6737–6740.
- [196] L. Dolejs, A. Mironov, F. Sorm, *Collect. Czech. Chem. Commun.* **1961**, *26*, 1015–1020.
- [197] L. Dolejs, A. Mironov, F. Sorm, *Tetrahedron Lett.* **1960**, *1*, 18–21.
- [198] E. J. Eisenbraun, T. George, B. Riniker, C. Djerassi, *J. Am. Chem. Soc.* **1960**, *82*, 3648–3652.
- [199] P. Pesnelle, P. Teisseire, M. Wichtl, *Planta Med.* **1964**, *12*, 403–405.
- [200] T. B. Andersen, S. A. Rasmussen, S. B. Christensen, H. T. Simonsen, *Phytochemistry* **2019**, *157*, 168–174.
- [201] A. A. Stierle, D. B. Stierle, E. Goldstein, K. Parker, T. Bugni, C. Baarson, J. Gress, D. Blake, *J. Nat. Prod.* **2003**, *66*, 1097–1100.
- [202] H. Hirota, Y. Tanahashi, T. Takahashi, *Tetrahedron Lett.* **1975**, *16*, 4579–4582.
- [203] A. Gandurin, *Chem. Ber.* **1908**, *41*, 4359–4363.
- [204] V. M. Trikojus, D. E. White, *J. Proc. R. Soc. N. S. W.* **1934**, *68*, 177–183.
- [205] H. H. G. McKern, M. C. Spies, J. L. Willis, *J. Proc. R. Soc. N. S. W.* **1954**, *88*, 15–21.
- [206] J. W. Loder, *Aust. J. Chem.* **1962**, *15*, 389–390.
- [207] H. Shuichi, Y. Katsumi, H. Nanao, M. Tamon, *Bull. Chem. Soc. Jpn.* **1968**, *41*, 1465–1466.
- [208] Y. Yazaki, W. E. Hillis, *Holzforschung* **1977**, *31*, 188–191.
- [209] M. Pinar, M. Rico, B. Rodriguez, *Phytochemistry* **1982**, *21*, 735–737.



- [210] E. Lemmich, B. Jensen, U. Rasmussen, *Phytochemistry* **1984**, *23*, 809–811.
- [211] C. H. Brieskorn, G. Krauss, *Planta Med.* **1986**, *52*, 305–307.
- [212] V. Kumar, J. Reisch, D. B. M. Wickremaratne, R. A. Hussain, K. S. Adesina, S. Balasubramaniam, *Phytochemistry* **1987**, *26*, 511–514.
- [213] P. Raharivelomanana, J.-P. Bianchini, A. Cambon, M. Azzaro, R. Faure, *Magn. Reson. Chem.* **1995**, *33*, 233–235.
- [214] S. D. Sarker, J. A. Armstrong, P. G. Waterman, *Phytochemistry* **1995**, *40*, 1159–1162.
- [215] K. Kojima, K. Isaka, O. Purev, G. Jargalsaikhan, D. Suran, H. Mizukami, Y. Ogihara, *Chem. Pharm. Bull.* **1998**, *46*, 1781–1784.
- [216] J. J. Makangara, S. A. Jonker, M. H. H. Nkunya, *Nat. Prod. Lett.* **2002**, *16*, 267–272.
- [217] P. A. Plattner, G. Magyar, *Helv. Chim. Acta.* **1941**, *24*, 1163–1166.
- [218] K. Takeda, H. Minato, *Tetrahedron Lett.* **1960**, *22*, 33–37.
- [219] H. Minato, *Tetrahedron Lett.* **1961**, *8*, 280–284.
- [220] M. Bittner, M. Silva, Z. Rozas, F. Papastergiou, J. Jakupovic, *Phytochemistry* **1994**, *36*, 695–698.
- [221] X.-C. Wang, S.-P. Ma, J.-H. Liu, L.-H. Hu, *Nat. Prod. Commun.* **2008**, *3*, 1649–1652.
- [222] A.-C. Huang, C. J. Sumbly, E. R. T. Tieking, D. K. Taylor, *J. Nat. Prod.* **2014**, *77*, 2522–2536.
- [223] R. B. Bates, R. C. Slagel, *Chem. Ind.* **1962**, 1715–1716.
- [224] M. V. Kadival, M. S. R. Nair, S. C. Bhattacharyya, *Tetrahedron* **1967**, *23*, 1241–1249.
- [225] C. Ehret, G. Ourisson, *Bull. Soc. Chim. Fr.* **1968**, 2629.
- [226] H. Ishii, T. Tozyo, M. Nakamura, H. Minato, *Tetrahedron* **1970**, *26*, 2751–2757.
- [227] I. A. Southwell, *J. Essent. Oil Res.* **1996**, *8*, 143–148.
- [228] H. Ishii, T. Tozyo, H. Minato, *Chem. Commun.* **1968**, 106–108.
- [229] H. Ishii, T. Tozyo, M. Nakamura, H. Minato, *Tetrahedron* **1970**, *26*, 2911–2918.
- [230] H. Hikino, Y. Hikino, Y. Takeshita, K. Shirata, T. Takemoto, *Chem. Pharm. Bull.* **1963**, *11*, 547–548.
- [231] S. Ito, M. Kodama, T. Nozoe, H. Hikino, Y. Hikino, Y. Takeshita, T. Takemoto, *Tetrahedron Lett.* **1963**, *4*, 1787–1792.
- [232] H. J. M. Gijzen, J. B. P. A. Wijnberg, G. A. Stork, A. de Groot, *Tetrahedron* **1991**, *47*, 4409–4416.
- [233] F. Bohlmann, K.-H. Knoll, C. Zdero, P. K. Mahanta, M. Grenz, A. Suwita, D. Ehlers, N. L. Van, W.-R. Abraham, A. A. Natu, *Phytochemistry* **1977**, *16*, 965–985.
- [234] F. Bohlmann, A. Suwita, C. Zdero, *Phytochemistry* **1978**, *17*, 1763–1767.
- [235] F. Bohlmann, C. Zdero, D. Berger, A. Suwita, P. Mahanta, C. Jeffrey, *Phytochemistry* **1979**, *18*, 79–93.
- [236] C. Dragar, V. A. Dragar, R. C. Menary, *J. Essent. Oil Res.* **1993**, *5*, 507–511.
- [237] I. Muhammad, X.-C. Li, D. C. Dunbar, M. A. ElSohly, I. A. Khan, *J. Nat. Prod.* **2001**, *64*, 1322–1325.
- [238] H. Ishii, T. Tozyo, M. Nakamura, E. Funke, *Chem. Pharm. Bull.* **1972**, *20*, 203–205.
- [239] H. Ishii, T. Tozyo, M. Nakamura, *Tetrahedron* **1971**, *27*, 4263–4268.
- [240] R. Croteau, S. L. Munck, C. C. Akoh, H. J. Fisk, D. M. Satterwhite, *Arch. Biochem. Biophys.* **1987**, *256*, 56–68.
- [241] A. Akhila, P. K. Sharma, R. S. Thakur, *Phytochemistry* **1988**, *27*, 2105–2108.
- [242] J. A. Faraldos, S. Wu, J. Chappell, R. M. Coates, *J. Am. Chem. Soc.* **2010**, *132*, 2998–3008.
- [243] H. Xu, B. Goldfuss, G. Schnakenburg, J. S. Dickschat, *Beilstein J. Org. Chem.* **2022**, *18*, 13–24.

---

Manuscript received: February 8, 2022

Accepted manuscript online: March 3, 2022

Version of record online: March 21, 2022

## Appendix C

### **Germacrene B – A Central Intermediate in Sesquiterpene Biosynthesis**

*Beilstein J. Org. Chem.* **2023**, 19, 186

DOI: 10.3762/bjoc.19.18



# Germacrene B – a central intermediate in sesquiterpene biosynthesis

Houchao Xu and Jeroen S. Dickschat\*

## Review

Open Access

Address:  
Kekulé-Institute of Organic Chemistry and Biochemistry, University of Bonn, Gerhard-Domagk-Straße 1, 53121 Bonn, Germany

Email:  
Jeroen S. Dickschat\* - dickschat@uni-bonn.de

\* Corresponding author

Keywords:  
biosynthesis; configuration determination; germacrene B; structure elucidation; terpenes

*Beilstein J. Org. Chem.* **2023**, *19*, 186–203.  
<https://doi.org/10.3762/bjoc.19.18>

Received: 06 January 2023  
Accepted: 10 February 2023  
Published: 20 February 2023

Associate Editor: D. Y.-K. Chen

© 2023 Xu and Dickschat; licensee Beilstein-Institut.  
License and terms: see end of document.

## Abstract

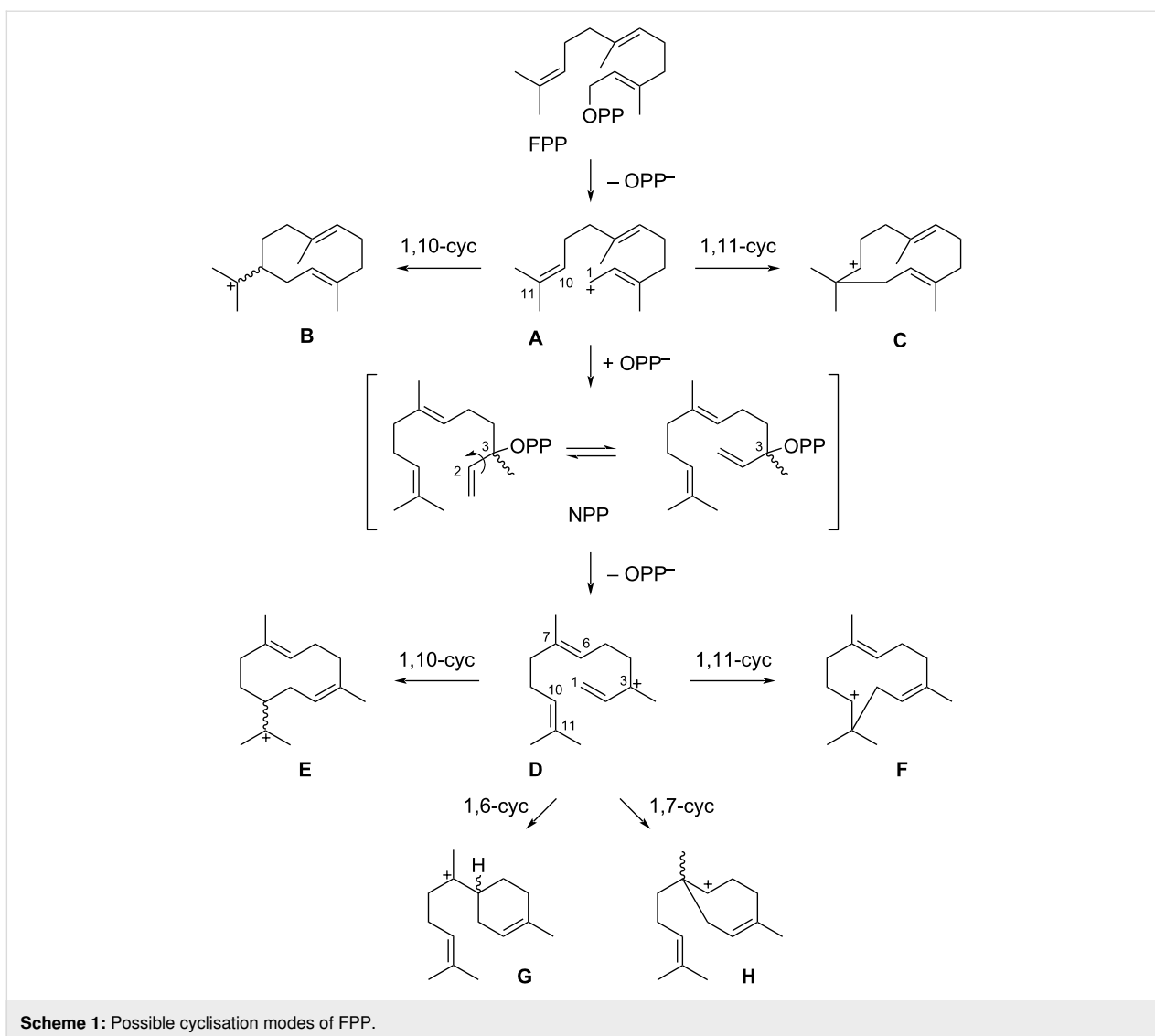
Germacrane are important intermediates in the biosynthesis of eudesmane and guaiane sesquiterpenes. After their initial formation from farnesyl diphosphate, these neutral intermediates can become reprotonated for a second cyclisation to reach the bicyclic eudesmane and guaiane skeletons. This review summarises the accumulated knowledge on eudesmane and guaiane sesquiterpene hydrocarbons and alcohols that potentially arise from the achiral sesquiterpene hydrocarbon germacrene B. Not only compounds isolated from natural sources, but also synthetic compounds are discussed, with the aim to give a rationale for the structural assignment for each compound. A total number of 64 compounds is presented, with 131 cited references.

## Introduction

Terpenoids constitute the largest class of natural products with ca. 100,000 known compounds. Biosynthetically, all terpenoids are derived from only a few acyclic precursors, including the monoterpene precursor geranyl diphosphate (GPP) [1], the precursor for sesquiterpenes farnesyl diphosphate (FPP) [2], geranylgeranyl diphosphate (GGPP) towards diterpenes [3], and the sesterterpene precursor geranylfarnesyl diphosphate (GFPP) [4]. It has been demonstrated recently, that even farnesylfarnesyl diphosphate (FFPP) can serve as a precursor to triterpenes [5], a compound class that was believed to be solely derived from squalene. Terpene synthases convert these linear precursors through cationic cascade reactions into terpene hydrocarbons or alcohols [6-8]. For type I terpene synthases this

multistep process is initiated by the abstraction of diphosphate to produce an allyl cation that subsequently undergoes typical cation reactions such as cyclisations by intramolecular attack of an olefin to the cationic centre, Wagner–Meerwein rearrangements, hydride or proton shifts. The process is terminated by deprotonation to yield a terpene hydrocarbon or by nucleophilic attack of water to generate a terpene alcohol.

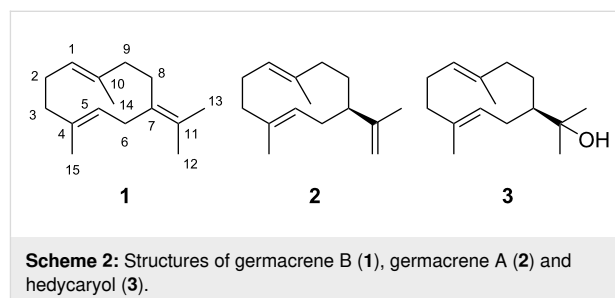
For the precursor of sesquiterpenes FPP six initial cyclisation modes are possible (Scheme 1). After ionisation to **A** either a 1,10-cyclisation to the (*E,E*)-germacradienyl cation (**B**) or a 1,11-cyclisation to the (*E,E*)-humulyl cation (**C**) is possible. Reattack of diphosphate at C-3 results in nerolidyl diphosphate



(NPP) that can undergo a conformational change by rotation around the C-2/C-3 single bond, which allows reionisation to **D**. This intermediate can react in a 1,10-cyclisation to the (*Z,E*)-germacradienyl cation (**E**) or a 1,11-cyclisation to the (*Z,E*)-humulyl cation (**F**), the *E/Z* stereoisomers of **B** and **C**. Furthermore, a 1,6-cyclisation to the bisabolyll cation (**G**) or a 1,7-cyclisation to **H** may follow, which is not possible from **A** because of its *2E* configuration (a hypothetical (*E*)-cyclohexene or (*E*)-cycloheptene would be too strained, the smallest possible ring with an *E* configuration is (*E*)-cyclooctene).

In some cases the initially formed neutral product can become reprotonated to initiate a second round of cyclisation reactions which usually leads to compounds of higher structural complexity. It was already noticed in the 1950s by Ruzicka [9] and Barton and de Mayo [10], followed by a more detailed elaboration by Hendrickson [11], that 10-membered sesquiterpenes

such as hedycaryol (**3**) can serve as neutral intermediates that can react upon reprotonation to 6-6- (selinane) or 5-7-bicyclic (guaiane) sesquiterpenes. We have recently summarised the accumulated knowledge about sesquiterpenes derived from germacrene A (**2**) [12] and hedycaryol (**3**) [13]. Now we wish to provide a review on the known chemical space of sesquiterpenes derived from germacrene B (**1**) (Scheme 2). Compounds



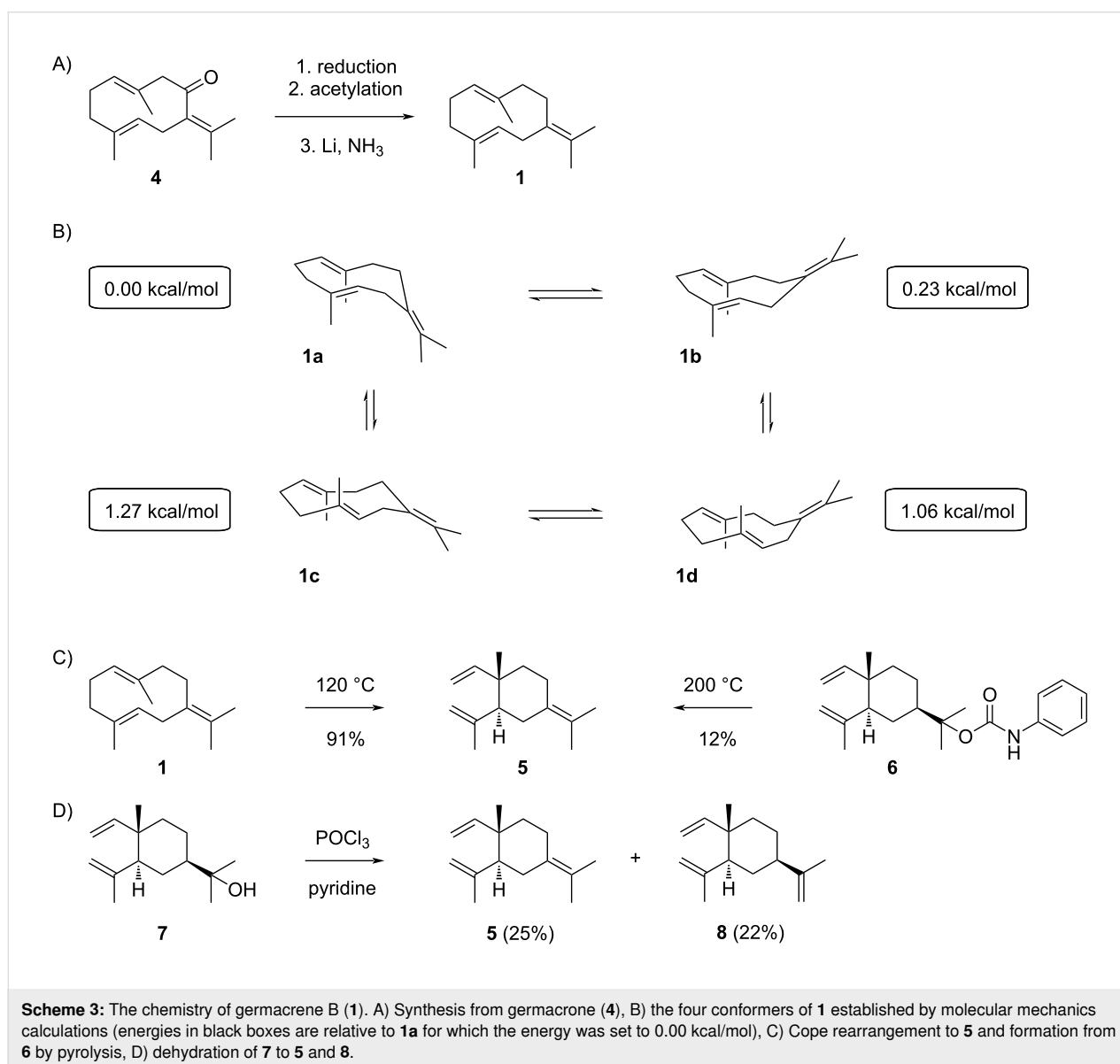
derived from **1** by oxidation will not be included in this article. The interested reader can find exemplary relevant information about this topic in references [14–18].

## Review

### Germacrene B

Germacrene B (**1**) was first prepared from germacrone (**4**), a compound identified by Šorm and co-workers [19], through a sequence of reduction to the alcohol, acetylation and reduction with lithium in ammonia (Scheme 3A) [20], and its structure was unambiguously assigned by X-ray crystallography of a silver nitrate adduct [21]. From natural sources, the compound was first obtained from *Humulus lupulus* by preparative gas chromatography [22] and from *Citrus junos* [23], followed by isolations from *Stenocalyx michelii* [24], *Citrus aurantifolia*

[25], and *Solidago canadensis* [26]. Germacrene B has been ascribed a warm, sweet, woody-spicy, geranium-like odour and is an important flavour constituent of lime peel oil [25]. Germacrene B is also one of the main constituents of the essential oils from different plants that have antibacterial activity [27–29]. Germacrene B synthases have been reported from *Solanum habrochaites* [30] and *Cannabis sativa* [31]. In addition, **1** is a minor product of the germacrene C synthase from *Lycopersicon esculentum* [32], the (+)-germacrene D synthase from *Zingiber officinalis* (17.1%) [33], the avermitilol synthase from *Streptomyces avermitilis* (5%) [34], and VoTPS1 from *Valeriana officinalis* [35]. For the bacterial selinadiene synthase (SdS) from *Streptomyces pristinaespiralis* **1** is an intermediate in the cyclisation of farnesyl diphosphate (FPP) to selina-4(15)-7(11)-diene [36]. Several SdS enzyme variants have been



constructed by site-directed mutagenesis, including the enzyme variants D83E, E159D and W304L, for which the product spectrum is shifted towards **1** as the main product [36].

Based on molecular mechanics calculations, four conformers **1a–d** have been described for **1** (Scheme 3B) [37]. The calculations revealed all four conformers are of similar stability, with **1a** being the most stable conformer. The fact that **1** shows a defined set of fifteen sharp signals in the  $^{13}\text{C}$  NMR spectrum [26] indicates that the interconversion between these conformers is a fast process at room temperature. This is in contrast to the findings for germacrene A (**2**) and hedycaryol (**3**) that show strong line broadening in the NMR spectra and multiple sets of peaks for different conformers [26,38–41], pointing to a higher energy barrier between their conformers in comparison to the barriers between the conformers of **1**. Like observed for germacrene A [40] and hedycaryol [41,42], **1** readily undergoes a Cope rearrangement to  $\gamma$ -elemene (**5**) above 120 °C (Scheme 3C), while the reaction of **1** with bis(benzonitrile) palladium chloride generates the palladium chloride complex of **5** from which **5** can be liberated by treatment with dimethyl sulfoxide [43]. Compound **5**, with tentatively assigned structure, was first obtained as a pyrolysis product of elemol phenylurethane (**6**) [44]. Its structure was subsequently secured by preparation from **1** through Cope rearrangement [20] and through dehydration of elemol (**7**) with  $\text{POCl}_3$  in pyridine yielding **5** and  $\beta$ -elemene (**8**) (Scheme 3D) [45]. Compound **5** has also frequently been reported from natural sources especially after heat treatment of the sample, and has been isolated from *Cryptotaenia japonica* [46], *Bunium cylindricum* [47], an unidentified *Pilocarpus* sp. [48], and *Aristolochia triangularis* [49].

Germacrene B (**1**) is also easily cyclised to selinanes. Percolation of **1** through alumina yields a 1:1 mixture of selina-3,7(11)-diene (**9**) and  $\gamma$ -selinene (**10**) (Scheme 4A) [43]. Interestingly, while racemic juniper camphor (**11**) is formed from **1** upon acid treatment [50], this reaction with diluted sulfuric acid in acetone results in (*rac*)-**11** quantitatively. This observation is explained by a protonation-induced cyclisation, successive addition of acetone and water to a hemiacetal that can decompose to **11** (Scheme 4B) [43]. Furthermore, **1** shows an interesting photochemistry (Scheme 4C). A [2 + 2] cycloaddition of the endocyclic double bonds yields **12** whose formation is understandable from conformers **1c** and **1d**. The all-*cis* stereoisomer **14** requires a photochemical *E/Z* isomerisation to **13** prior to [2 + 2] cycloaddition. Further photochemical products from **1** include **5**, **15** that may be formed through a biradical mechanism, and rearranged **16** [51]. Germacrene B (**1**) has planar chirality (Scheme 4D), but recovery of the starting material from an incomplete Sharpless epoxidation of its derivative

15-hydroxygermacrene (**17**) showed that this material was racemic, indicating a rapid interconversion between the enantiomers of **17**. Consequently, also the enantiomers of **1** may undergo a fast interconversion [52]. The  $^1\text{H}$  and  $^{13}\text{C}$  NMR data of **1** have been reported [26].

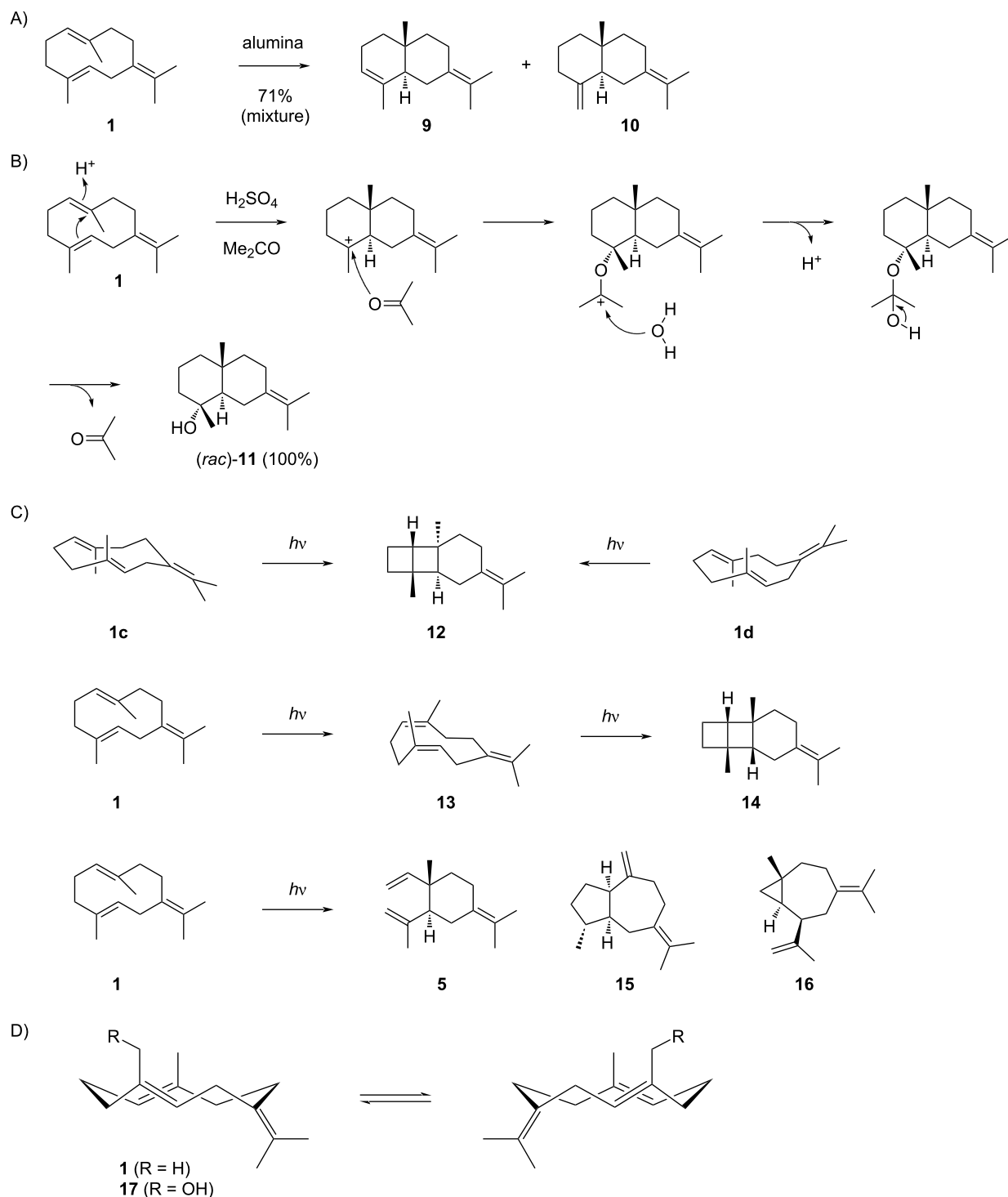
Upon reprotonation germacrene **B** (**1**) can in theory yield several cyclisation products with distinct skeletons. Eudesmanes can be obtained through reprotonation at C-1 and cyclisation to intermediate **I**, or through reprotonation at C-4 leading to cation **J** (Scheme 5A). Further cyclisation modes include a reprotonation at C-4 and cyclisation to **K** or reprotonation at C-10 and cyclisation to **L**, which represent possible precursors of guaianes (Scheme 5B). For all four intermediates **I–L** different stereochemistries may be realised. In principle, these reactions may be enzyme catalysed or proceed without enzyme catalysis, e.g., during chromatographic purifications of compounds from complex extracts. In the latter case, because of the achiral nature of **1**, racemic mixtures are expected, while enzyme products should usually be enantiomerically pure or enriched.

## Eudesmanes

The eudesmane skeleton can arise by reprotonation at C-1 of **1**, leading to four different stereoisomers of cation **I**, i.e., **I1** with a *trans*-decalin skeleton, its enantiomer **I2**, **I3** representing the *cis*-decalin skeleton, and its enantiomer **I4** (Scheme 6A). In principle, the eudesmane skeleton can also be formed through cyclisations induced by reprotonation at C-4. Assuming *anti* addition to the C-4/C-5 double bond, these reactions lead to four stereoisomers of the secondary cation **J**, two with a *trans*-decalin skeleton (**J1** and **J2**) and two with a *cis*-decalin skeleton (**J3** and **J4**). However, no natural products are known that may arise through any of these cations **J**, showing that a cyclisation of **1** induced by reprotonation at C-4 is not preferred. Also no compounds have been isolated with their structures rigorously elucidated that arise through cation **I4**. For compounds potentially generated through intermediates **I1–I3** the accumulated knowledge will be discussed in the following sections.

## Eudesmanes from **I1**

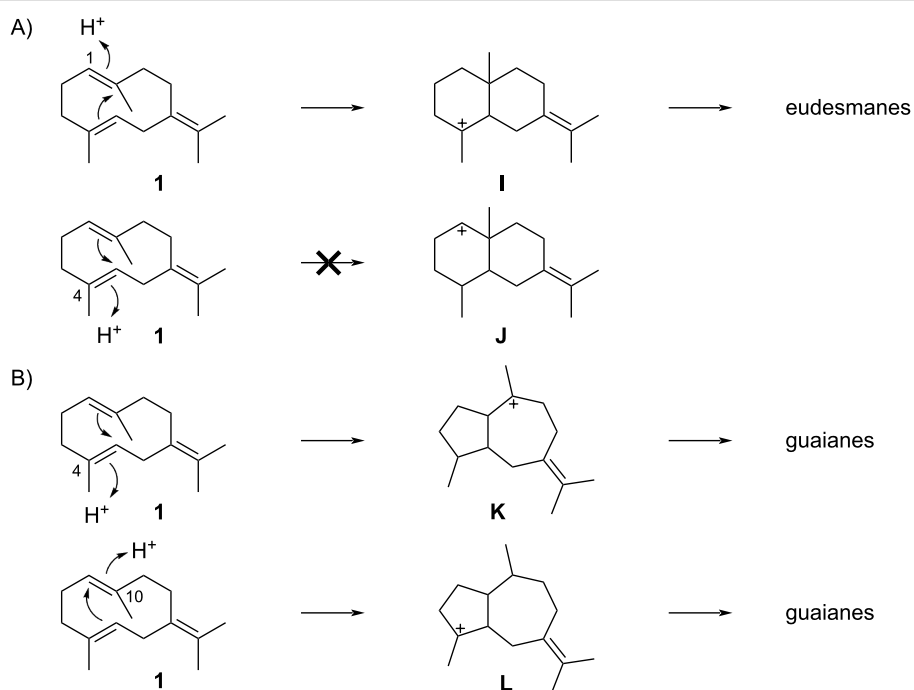
The eudesmane sesquiterpenes derived from cation **I1** are summarised in Scheme 7. Cation **I1** can either be deprotonated to yield selina-3,7(11)-diene (**9**), (+)- $\gamma$ -selinene (**10**) or (+)-selina-4,7(11)-diene (**18**), or captured by water resulting in juniper camphor (**11**) or 4-*epi*-juniper camphor (**19**).  $\gamma$ -Selinene (**10**) was first obtained by Šorm and co-workers from wormwood oil (*Artemisia absinthum*). Its positive optical rotation ( $[\alpha]_{\text{D}}^{25} = +2.8$ ) [53] suggests an enzymatic formation from **1** in this species. Compound **9**, along with **10**, was first isolated from *Humulus lupulus*, and the structures of both compounds were



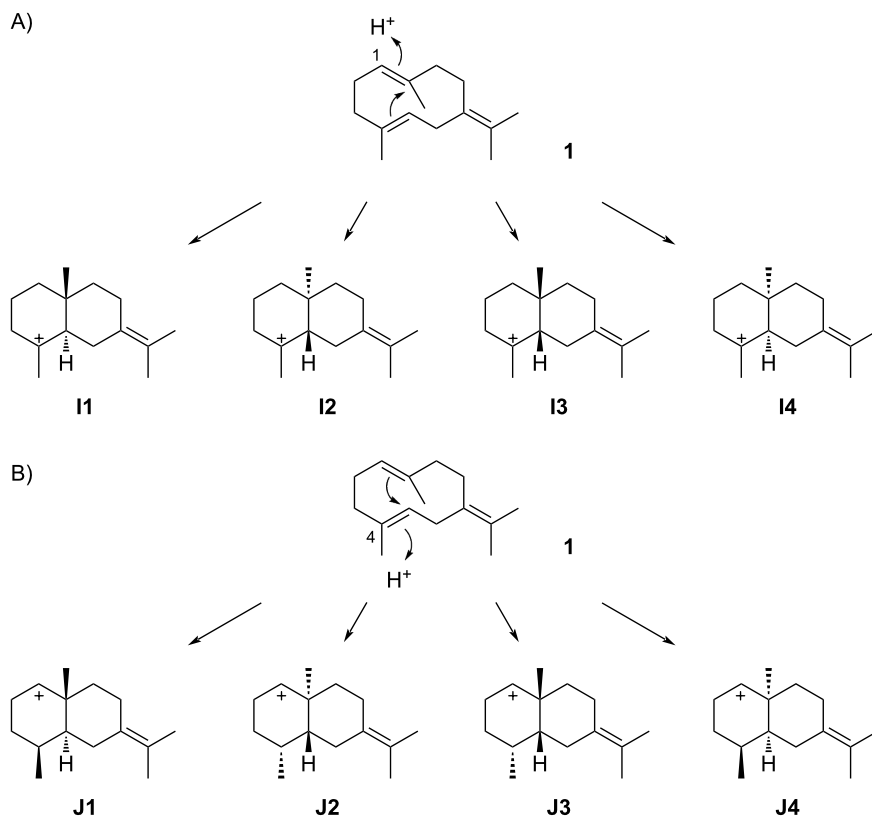
**Scheme 4:** The chemistry of germacrene B (**1**). A) Cyclisation of **1** to **9** and **10** upon treatment with alumina, B) conversion into (*rac*)-**11** by treatment with diluted sulfuric acid in acetone, C) photochemical products from **1**, and D) planar chirality of **1** and its derivative **17**.

elucidated by  $^1\text{H}$  NMR spectroscopy and catalytic hydrogenation, yielding the same compound selinane in both cases [54]. Both compounds were later also isolated from *Cannabis sativa*

[55]. Unfortunately, no optical rotations were given in these reports, so it remains unknown if the isolated materials arose from **1** by enzymatic or acid-catalysed reactions.

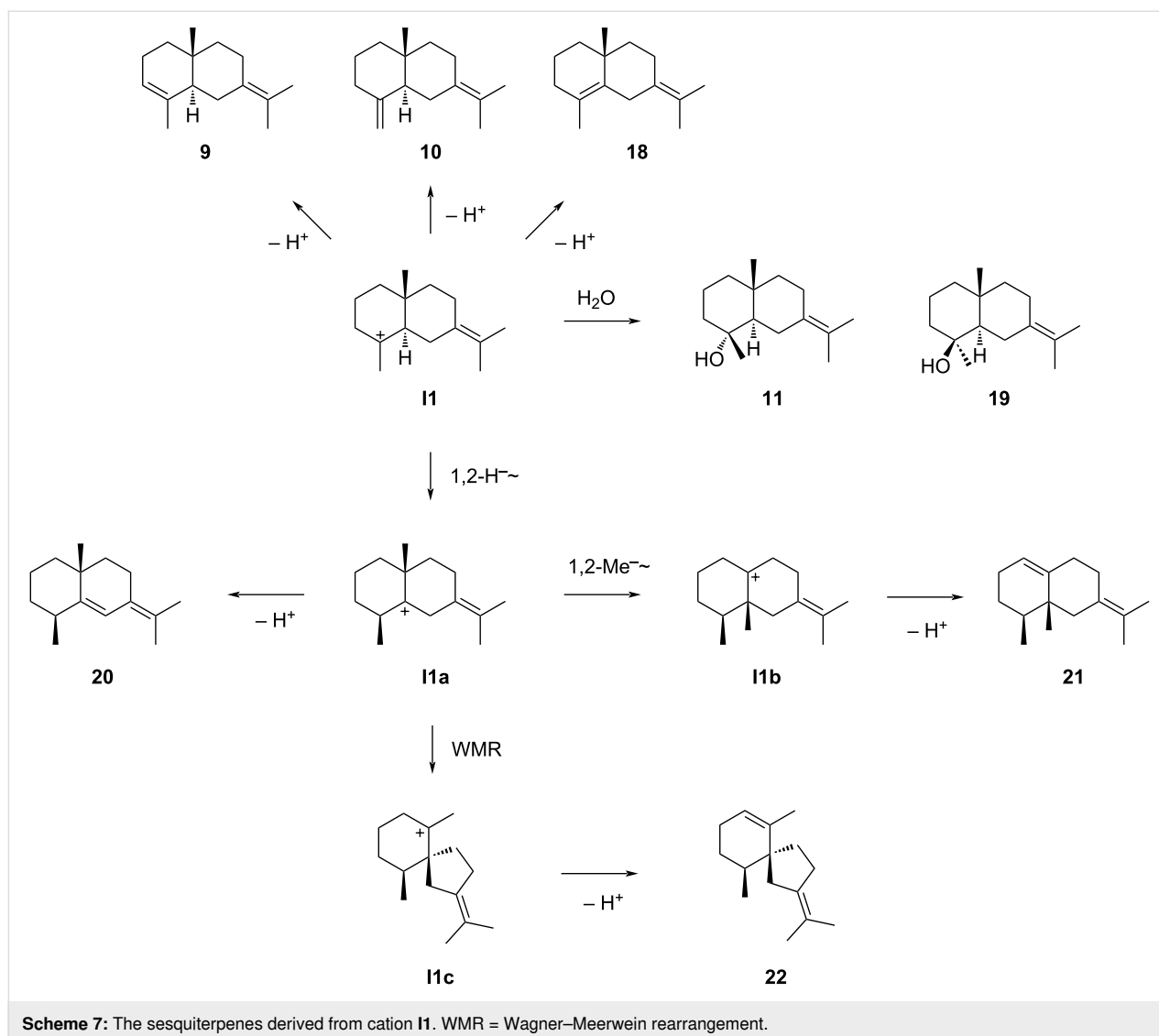


**Scheme 5:** Possible cyclisation reactions upon reprotonation of **1**. A) Cyclisations to eudesmane sesquiterpenes, B) cyclisations to guaiane sesquiterpenes.



**Scheme 6:** Cyclisation modes for **1** to the eudesmane skeleton. A) The reprotonation of **1** at C-1 potentially leads to four stereoisomers of cation **I**, B) reprotonation at C-4 potentially leads to four stereoisomers of **J**.

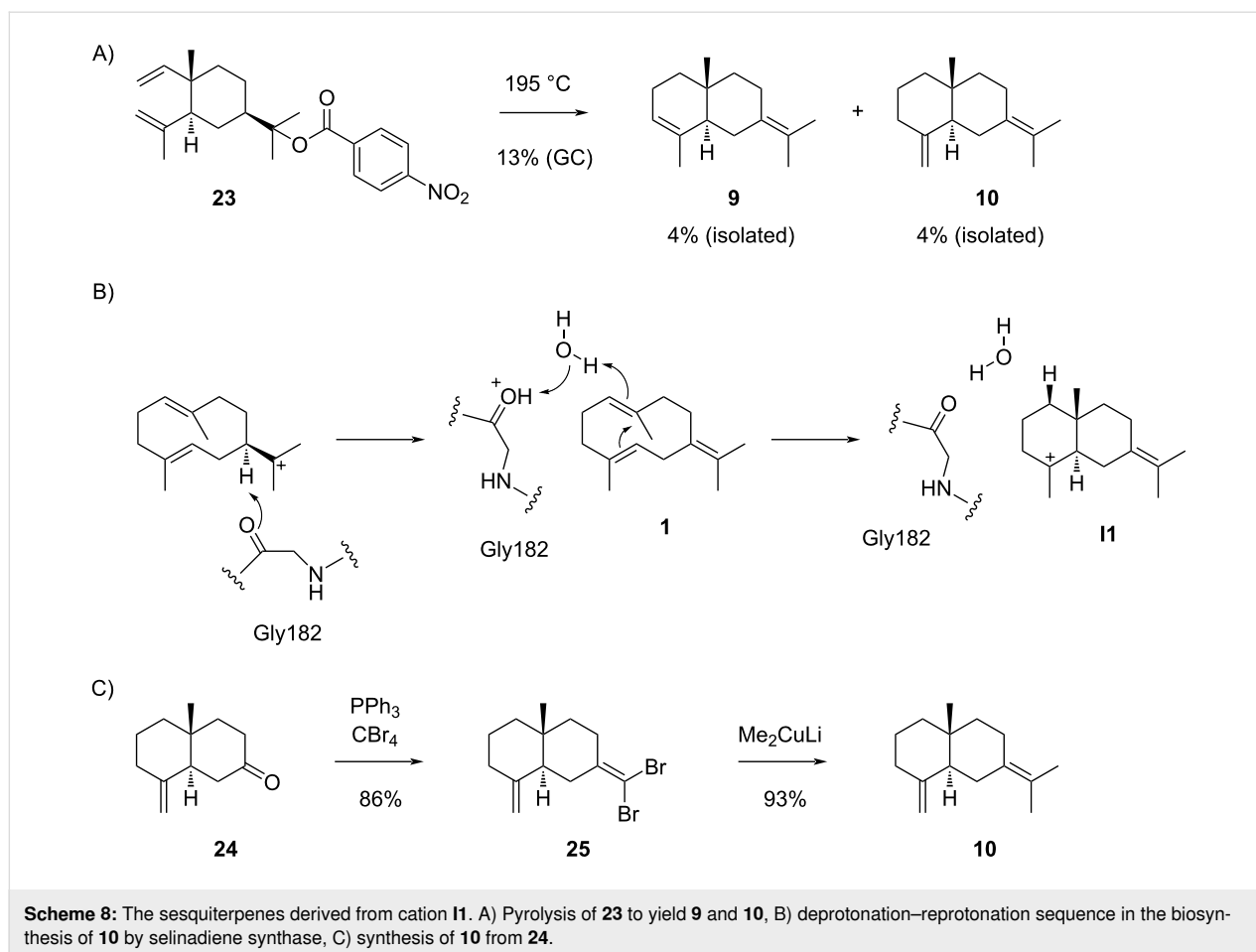




The sesquiterpenes **9** and **10**, besides several other products, were also prepared through pyrolysis of elemyl *p*-nitrobenzoate (**23**) (Scheme 8A) [56]. Because of the enantiomerically pure starting material, the products were obtained in enantiomerically pure form, showing an optical rotation of  $[\alpha]_{\text{D}} = -6.0$  ( $c$  0.484) for **10**, while no data were given for the optical rotation of **9**. Compound **9** was also isolated from *Asarum caulescens* ( $[\alpha]_{\text{D}}^{25} = -5.5$ ,  $c$  0.4, MeOH) [57]. Despite the opposite sign for the optical rotation as reported by Šorm and co-workers [53], the same absolute configuration of **9** is shown in this report. Furthermore, **9** has been described as a marker of the Lemberger variety of grapes (*Vitis vinifera*) [58]. Additional sources from which **10** has been isolated include *Persea japonica* [59], *Solidago canadensis* [60], *Citrus nobilis* ((+)-form) [61], *Zingiber officinalis* [62], *Myrica pensylvanica* and *M. macfarlanei* [63], *Trichogonia scottmorii* [64], and *Podocarpus spicatus* in which case a

high optical rotation was reported ( $[\alpha]_{\text{D}}^{20} = +82$ ,  $c$  2.9,  $\text{CHCl}_3$ ) [65].

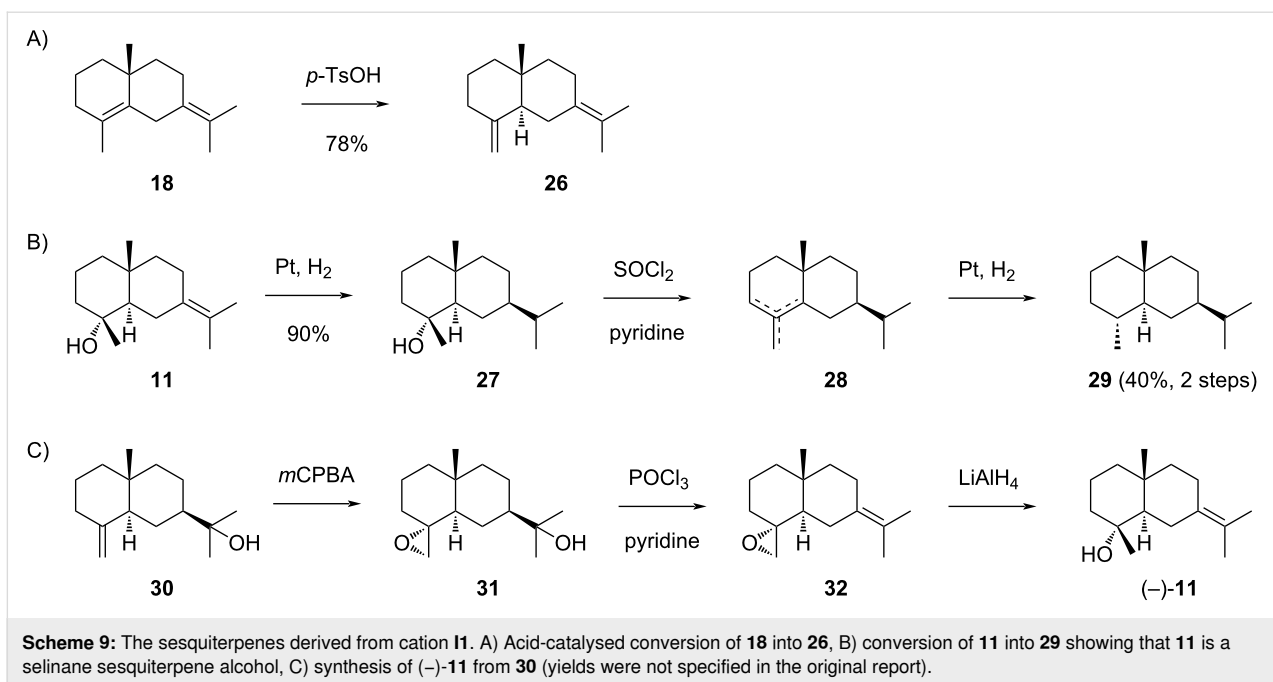
The sesquiterpene **9** is a side product of the  $\delta$ -selinene synthase (ag4) from *Abies grandis* [66] and a product of several terpene synthases from *C. sativa* (CsTPS7, CsTPS8 and CsTPS22) [67], while **10** is the main product of the bacterial selinadiene synthase from *Streptomyces pristinaespiralis* [36,68]. It has recently been shown by a combined computational and experimental approach that in this enzyme the main chain carbonyl oxygen of Gly182 near the helix G kink and an active site water are involved in the deprotonation–reprotonation sequence in the biosynthesis of **10** (Scheme 8B) [69].  $\gamma$ -Selinene (**10**) has been synthesised from ketone **24** through conversion into the dibromoalkene **25** with  $\text{PPh}_3$  and  $\text{CBr}_4$ , followed by treatment with  $\text{Me}_2\text{CuLi}$  (Scheme 8C) [70]. NMR data for **9** [71] and for **10** [59] have been published.



Selina-4,7(11)-diene (**18**),  $[\alpha]_D^{24} = +34$  ( $c$  0.90), was first isolated from the marine alga *Laurencia nidifica*. Its structure was determined by NMR spectroscopy and verified by the acid-catalysed conversion into  $\delta$ -selinene (**26**) (Scheme 9A) [72]. The same compound **18** was also reported from the closely related alga *Laurencia nipponica* [73] and from lime oil (*Citrus aurantifolia*) [74]. Fully assigned <sup>1</sup>H and <sup>13</sup>C NMR data were reported for **18** [72,74].

The structure elucidation of juniper camphor (**11**), a compound originally isolated by chemists at Schimmel, the world leading company of the late 19th and early 20th century dealing with essential oils and perfumes, was initiated by Šorm and co-workers [75]. From the sequence of catalytic hydrogenation to **27**, dehydration to a mixture of alkenes (**28**) and hydrogenation to selinane (**29**) it was concluded that **11** was a selinane sesquiterpene alcohol (Scheme 9B) [75]. Four years later, based on NMR data Bhattacharyya and co-workers suggested a *cis*-ring junction for **11** [76], but a synthesis from  $\beta$ -eudesmol (**30**) through epoxidation to **31**, dehydration to **32** and epoxide opening with LiAlH<sub>4</sub> yielded (–)-**11** (Scheme 9C) [77], contradicting this assignment.

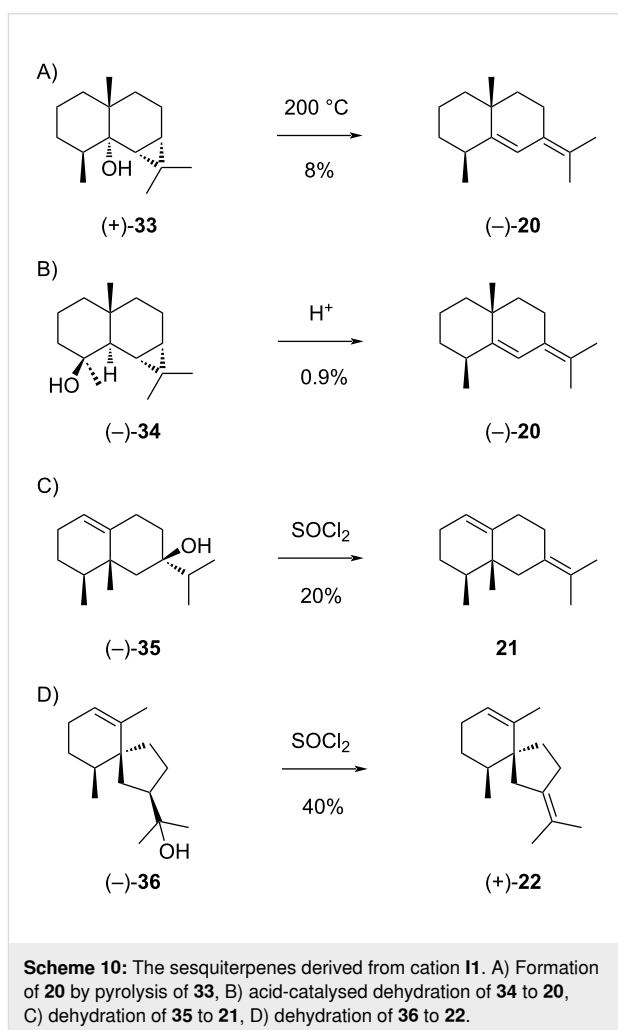
Notably, Šorm and co-workers noticed that **11** was racemic, because neither **11** nor any of its degradation products showed optical activity [75], suggesting that the compound they had isolated arose through acid-catalysed cyclisation of **1** rather than in an enzymatic process. Also the material isolated from *Platysace linearifolia* showed no optical rotation [78], while the optical activity of **11** isolated from *Bunium cylindricum* [47], *Acritopappus prunifolius* [79], *Aniba riparia* [80], *Juniperus oxycedrus* [81], and *Laggera alata* [82] has not been determined. The (–)-enantiomer of **11** with the structure as shown in Scheme 9C was reported from *Cabralea cangerana* ( $[\alpha]_D^{20} = -1.3$ ,  $c$  1.3, CDCl<sub>3</sub>) [83], *Zanthoxylum naranjillo* (no value specified) [84], and *Chiloscyphus polyanthos* ( $[\alpha]_D = -3.0$ ,  $c$  2.41, CHCl<sub>3</sub>) [85]. The (+)-enantiomer of **11** is known from *Cinnamomum camphora* ( $[\alpha]_D^{25} = +1.79$ ), representing the first isolated enantiomerically enriched material [86]. The low value of the optical rotation of **11** makes configurational assignments based on optical activity difficult, especially if minor contaminants falsify these data. Furthermore, the variability of the optical rotations given in the literature may be a consequence of mixed enantiomeric compositions arising from contaminations of enzymatically formed **11** with **11** generated upon acid cataly-



sis during compound isolations. The reporting of (-)-**11**, (+)-**11** and **11** of unspecified absolute configuration all under the same CAS number (473-04-1) adds to the confusion. Moreover, one report is available that mentions the isolation of **11** from *Atractylodes macrocephala* [87]. For unclear reason, this paper is assigned to CAS number 1647153-38-5 representing the structure of **19** (Scheme 7), which actually seems to be an unknown compound.

Compound **11** is a side product of ZmTPS7 from *Zea mays* [88] and  $^1\text{H}$  and  $^{13}\text{C}$  NMR data for **11** have been published [82,83]. A recent molecular docking study suggested that **11** can bind to the main protease  $\text{M}^{\text{Pro}}$  of the SARS-CoV-2 virus that is involved in viral reproduction, but experimental tests supporting this finding are lacking [89].

Selina-5,7(11)-diene (**20**) can arise from **11** through 1,2-hydride shift to **11a** and deprotonation (Scheme 7). This compound was first reported from olibanum oil, but only identified from its mass spectrum and GC retention time [90]. This structural assignment in the absence of a reference standard or at least literature data for **20** is likely erroneous. Compound (-)-**20** was later obtained by thermal degradation of (+)-maalian-5-ol (**33**) (Scheme 10A) and upon treatment of 4-*epi*-maaliol (**34**) with acid (Scheme 10B). Full  $^1\text{H}$  and  $^{13}\text{C}$  NMR data for **20** were reported [91]. Compound **21** can in theory be formed from **11a** by 1,2-methyl group shift to **11b** and deprotonation (Scheme 7). However, this compound was only obtained as synthetic material by dehydration of (-)-1(10)-valencen-7 $\beta$ -ol (**35**) (Scheme 10C) [92], but has not been isolated from natural



sources. Compound **22** could be formed from **11a** by Wagner–Meerwein rearrangement to **11c** and deprotonation (Scheme 7). This hydrocarbon ( $[\alpha]_D^{22} = +26$ ,  $c$  0.06) has been obtained as a dehydration product of (–)-hinesol (**36**) (Scheme 10D), but has never been isolated from natural sources.  $^1\text{H}$  NMR data have been reported [92].

## Eudesmanes from **12**

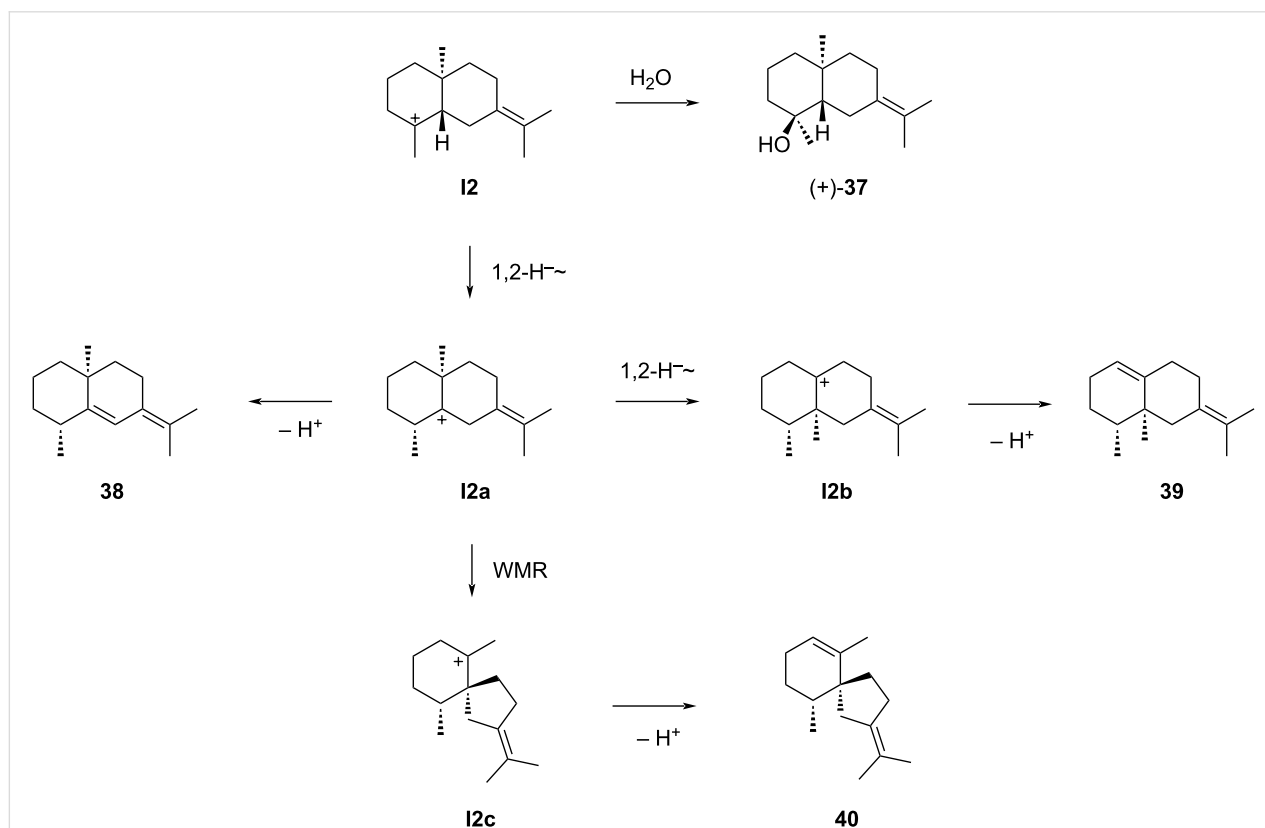
Much less is known about sesquiterpenes derived from cation **12** (Scheme 11). The compounds described in the literature include (+)-juniper camphor (**37**) that can be formed by attack of water to **12**. As mentioned above, this compound occurs in *Cinnamomum camphora* [86] and has later also been isolated from *Laggera pterodonta* ( $[\alpha]_D^{24} = +4$ ,  $c$  0.5, MeOH) [93]. Compound **38**, (+)-eudesma-5,7(11)-diene, could potentially arise from **12** by 1,2-hydride shift to **12a** and deprotonation, but has not been isolated from natural sources. This material was obtained by treatment of (+)-6,11-epoxyeudesmane (**41**) with acidic ion exchange resin (Scheme 12A) [94].

Also 4 $\beta$ H,5 $\alpha$ -eremophila-1(10),7(11)-diene (**39**), biosynthetically accessible from **12a** by 1,2-methyl shift to **12b** and deprotonation (Scheme 11), is only known as a synthetic compound. This hydrocarbon has first been obtained by dehydration of

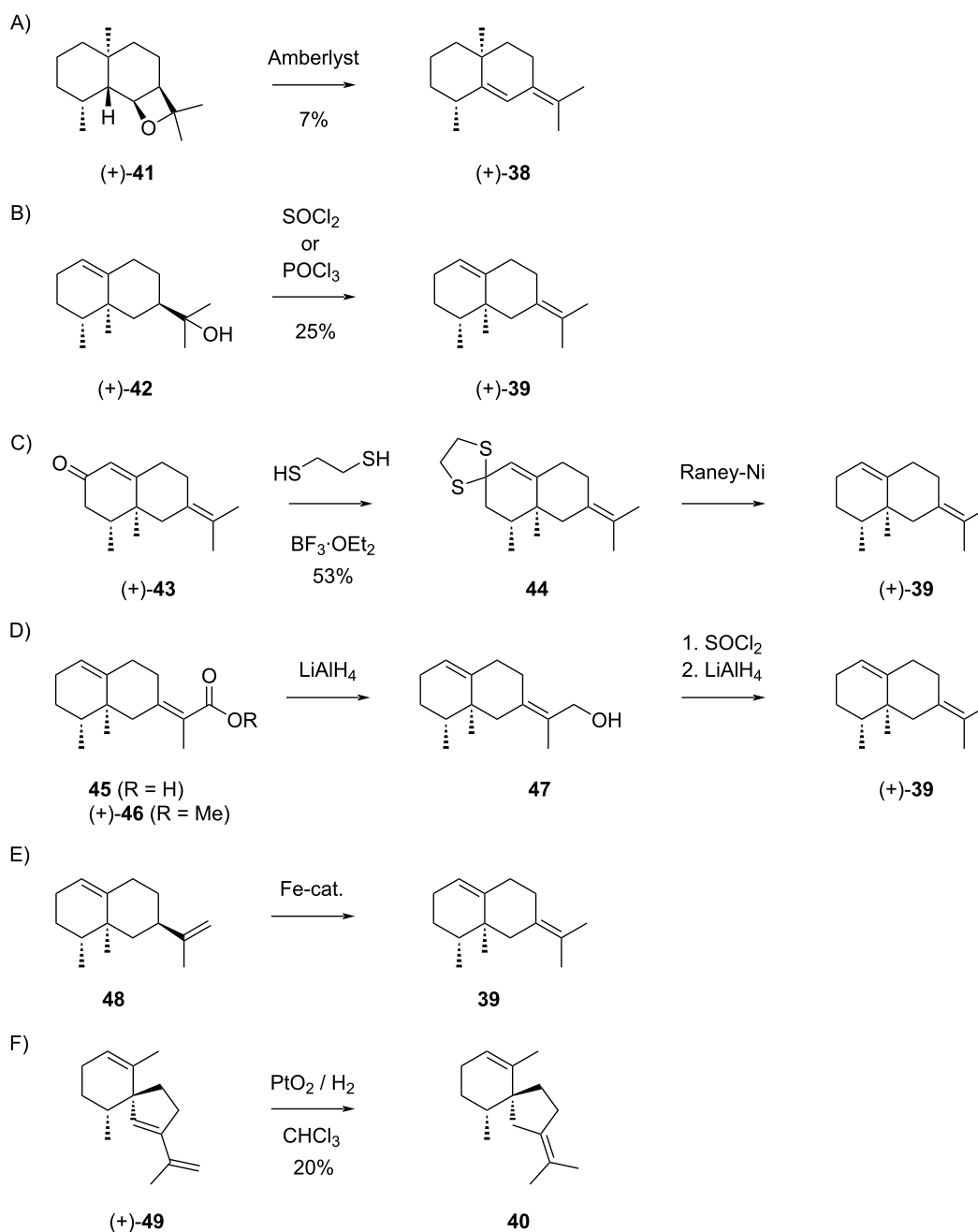
(+)-valerianol (**42**) with  $\text{SOCl}_2$  or  $\text{POCl}_3$ , yielding – besides the Hofmann product as main product (75%) – (+)-**39** (25%,  $[\alpha]_D^{20} = +167.5$ , neat) (Scheme 12B) [95]. After the first description of **39**, also (+)- $\alpha$ -vetivone (**43**) (Scheme 12C) [96,97] and isovalencenic acid (**45**) (Scheme 12D) [98] were correlated to this hydrocarbon. Recently, an iron catalyst has been developed that was applied in the isomerisation of valencene (**48**) to **39** (Scheme 12E) [99]. The biogenesis of **40** would be possible from **12a** through Wagner–Meerwein rearrangement to **12c** and deprotonation, but also this compound is not known as a natural product. This hydrocarbon has been obtained by partial hydrogenation of (+)- $\alpha$ -vetispirene (**49**) in a small scale reaction using  $\text{PtO}_2$  hydrate in  $\text{CHCl}_3$  as a catalyst (Scheme 12F). The amounts of isolated **40** (0.2 mg) were insufficient for a full spectroscopic characterisation [92].

## Eudesmanes from **13**

Also only a few compounds potentially arising from **13** are known (Scheme 13). Compound **18** was already discussed above and can be formed by deprotonation from **11** or **13**. Cation **11** seems to be the more likely precursor than **13**, because **11** is the intermediate towards structurally related natural products such as the widespread compounds **9** and **10** and a common biosynthesis of **18** through the same intermedi-



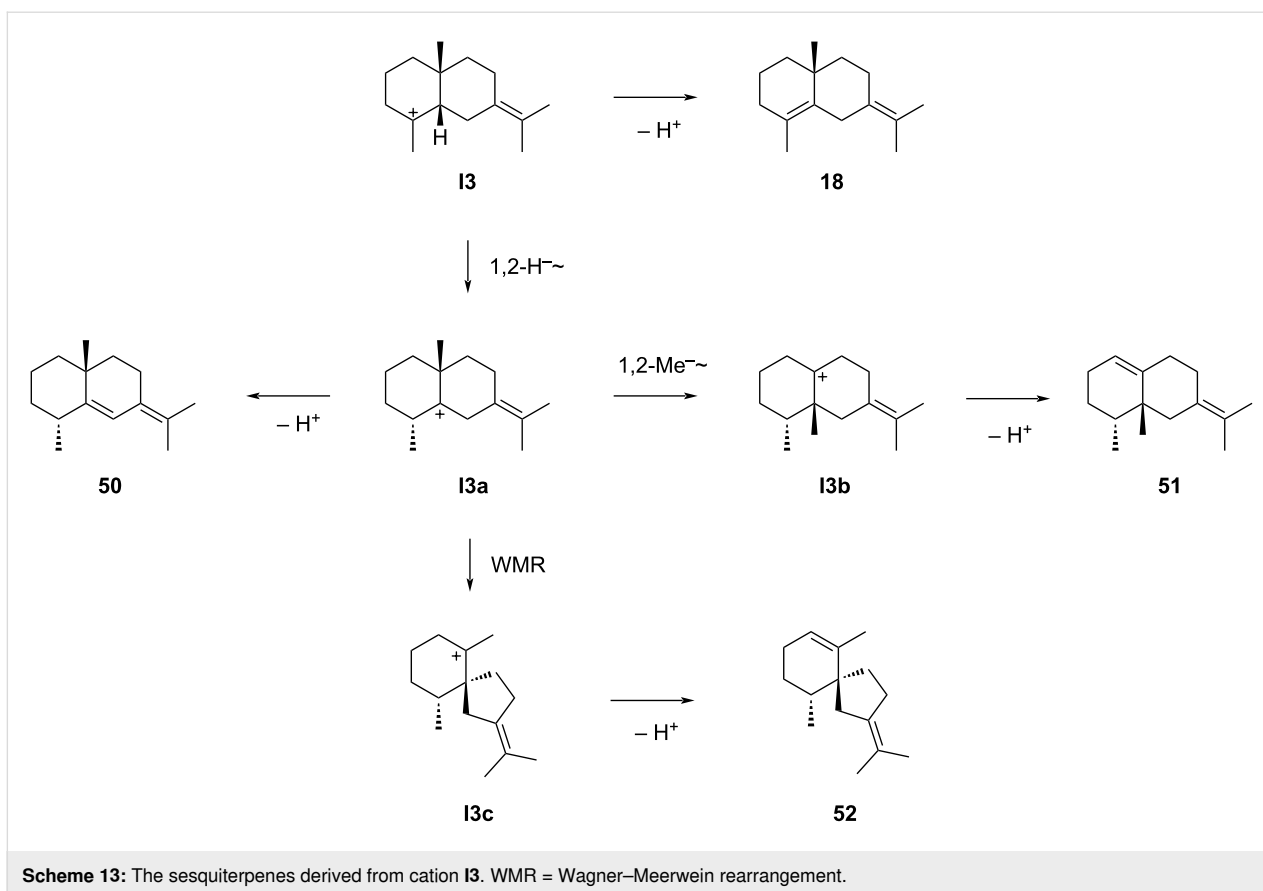
**Scheme 11:** The sesquiterpenes derived from cation **12**. WMR = Wagner–Meerwein rearrangement.



**Scheme 12:** The sesquiterpenes derived from cation **12**. A) Acid catalysed conversion of **41** into **38**, B) dehydration of **42** to **39**, C) chemical correlation of **43** with **39**, D) chemical correlation of **45** with **39** (no yields were given in the original report), E) isomerisation of **48** to **39** (product was not isolated), F) partial hydrogenation of **49** to **40**.

ate can be assumed (Scheme 7). A 1,2-hydride shift to **I3a** and deprotonation could give rise to **50**, a compound for which the situation in the literature is very confusing. There is no paper available describing the isolation and structure elucidation of a compound with the structure of **50**, and the first published paper that can be found under the CAS number of **50** (869998-21-0) does not mention this compound [100]. Several later reports

claim the detection of “eudesma-5,7(11)-diene”, a name assigned to CAS number 869998-21-0, but neither a structure is shown nor a reference to previous work is given in these reports, leaving doubt about the stereostructure the authors of this work had in mind [101-103]. One recent report mentions the detection of “eudesma-5,7(11)-diene”, but again no structure is shown, and the structural assignment is based on a com-



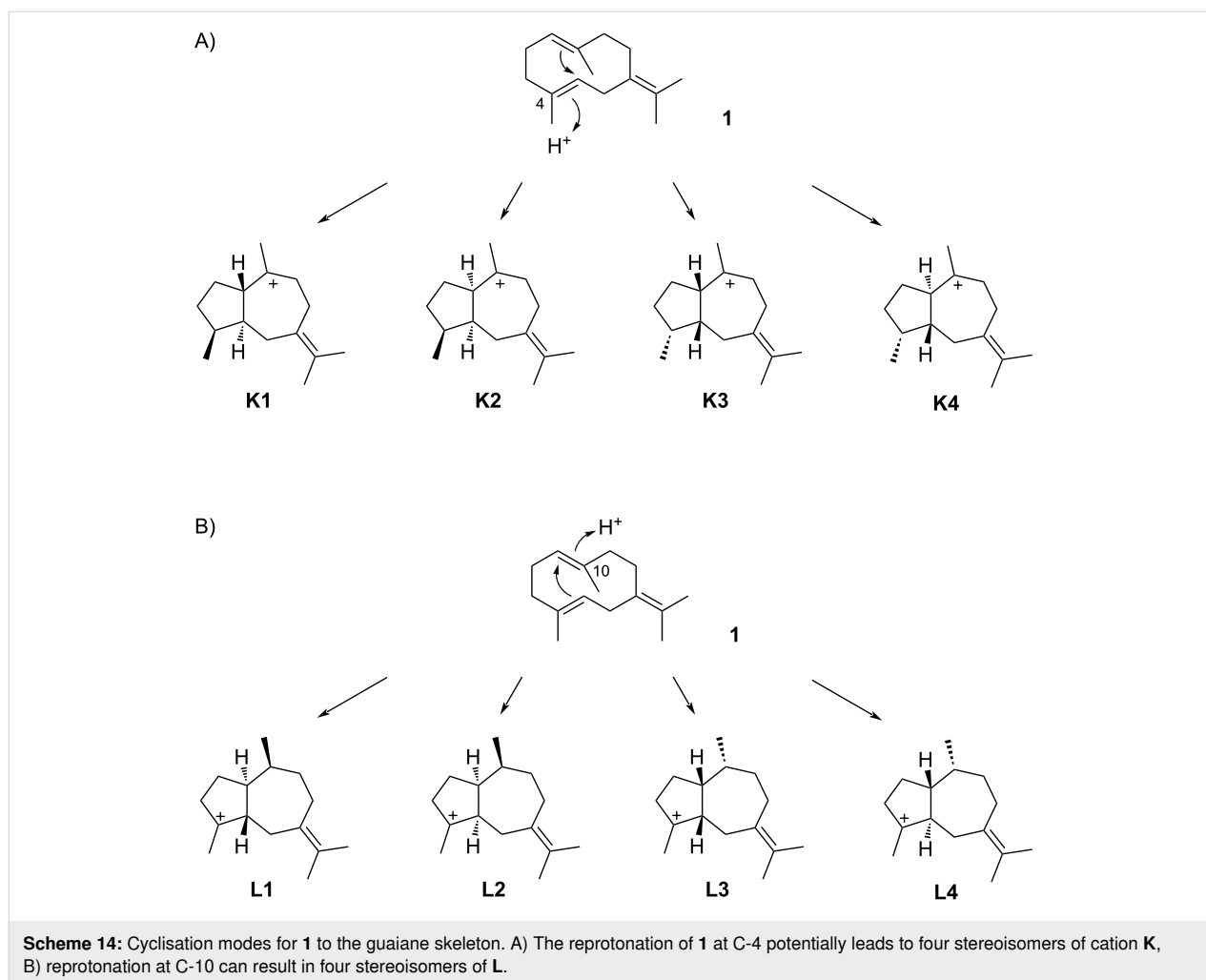
parison of retention indices [104]. However, the deviation between measured and reference retention index is quite large ( $I = 1572$  vs  $1543$ ), and the reference data originate from [103] in which the basis for structural assignment is unclear. Finally, one more paper assigned to CAS number 869998-21-0 mentions the detection of “eudesma-5,7(11)-diene”, but in this case the structure of **38** (Scheme 11) instead of **50** is shown, which based on a comparison of the measured to a database retention index may at least in terms of the relative configuration be a correct structural assignment [105]. Taken together, the confusing situation for **50** in the literature demonstrates impressively, how inaccurate data reporting can lead to unclear structural assignments and even error propagation, and shows the importance of structure elucidation by classical methods, i.e., isolation and compound characterisation by NMR spectroscopy and determination of optical rotation.

Compound **51** can be generated biosynthetically from **13a** through 1,2-methyl migration to **13b** and deprotonation. However, this hydrocarbon has not been isolated from natural sources and is only known as racemic synthetic material [106]. Similarly, **52** has only been described as a synthetic compound [107]. Its hypothetical biosynthesis is possible from **13a** by Wagner–Meerwein rearrangement to **13c** and deprotonation.

## Guaianes

As discussed above, the cyclisation of **1** induced by reprotonation at C-4 to the eudesmane skeleton encounters obstacles because of the formation of secondary cations. Preferentially, reprotonation at C-4 leads to the guaiane skeleton since the formed cations are tertiary. Alternatively, reprotonation of **1** at C-10 can also induce the formation of the guaiane skeleton. Assuming *anti* addition to the C-4/C-5 double bond in **1**, only four cationic intermediates (**K1–K4**) can be generated by reprotonation at C-4 (Scheme 14A). Similarly, reprotonation of **1** at C-10 leads by *anti* addition to the C-1/C-10 double bond to four cationic intermediates, **L1–L4** (Scheme 14B).

The guaiane sesquiterpenes derived from cationic intermediates **K1**, **K2** and **K4** are summarised in Scheme 15A, while no compounds are known whose formation could be explained from **K3**.  $\beta$ -Bulnesene (**53**), a product by the deprotonation of **K1** or **K2**, was first isolated from the guaiac wood oil of *Bulnesia sarmientoi* [108] and later also observed in *Pogostemon cablin* [109]. Bulnesol (**57**), a compound of known absolute configuration [110] that occurs in the same essential oil [108], has been converted through pyrolysis of its acetate **58** into **53** (Scheme 15B) [111], securing the relative configuration. This work did not comment on the question of absolute configura-



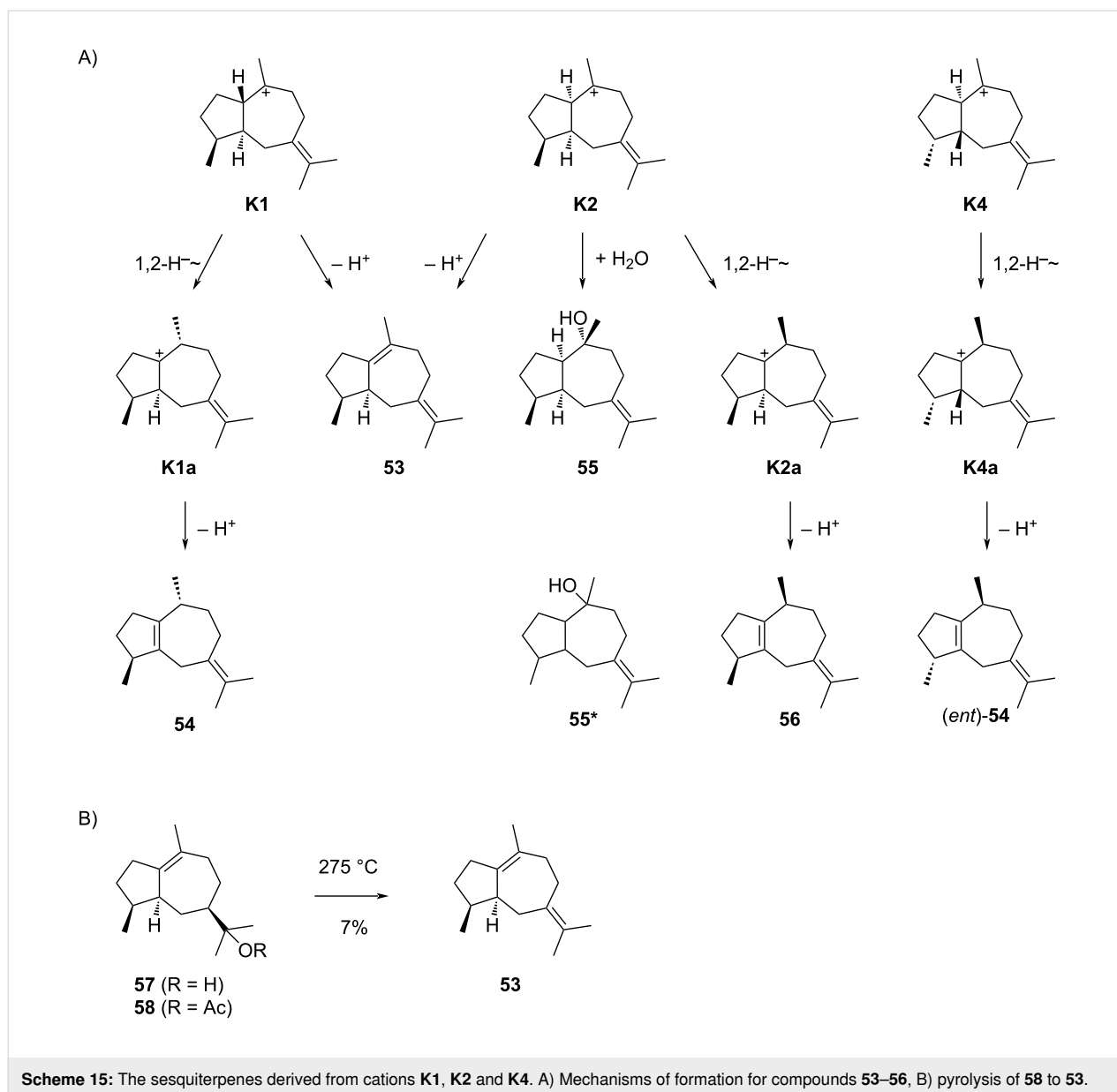
tion, but assuming a common biosynthesis of **53** and **57** analogous absolute configurations for these compounds are likely. Despite several reported syntheses of (*rac*)-**53** [112–116], no enantioselective synthesis is available. Full  $^1\text{H}$  and  $^{13}\text{C}$  NMR data of **53** (including 14 carbon signals) have been published [113].

The guaiane sesquiterpenes that are potentially derived from cationic intermediates **L1**–**L4** are summarised in Scheme 16A. *trans*- $\beta$ -Guaiene (**54**) can either be generated from **K1** undergoing a 1,2-hydride shift to **K1a** followed by deprotonation (Scheme 15A), or from **L4** through a similar sequence of steps (Scheme 16A). Its enantiomer *ent*-**54** could analogously arise from **K4** or **L1**. The first detection of this compound was claimed from *Aframomum alboviolaceum*, but this study did not report on the isolation and structure elucidation [117]. Rather the identification was only based on GC–MS data, without a reference to a previous identification through rigorous structure elucidation. Conclusively, this compound has not been described thoroughly and its identification is doubtful. Information about the mass spectrum and Kovats retention index have

been added to data bases such as the NIST Chemistry Webbook [118], which promoted the ambiguous detection of **54** in many other species, as described in more than 300 papers to date.

Compound **55** can be formed from **K2** through capture with water. A compound with the same planar structure of **55\*** named guai-7(11)-en-10-ol has been reported from *Zanthoxylum syncarpum* with fully assigned  $^1\text{H}$  and  $^{13}\text{C}$  NMR data, but unresolved relative and absolute configuration [119]. For unclear reason, this compound has been assigned to CAS number 461691-86-1, a molecule for which the relative and absolute configuration are shown. No other reports for this compound are available.

$\beta$ -Guaiene (**56**) is a well described compound that can biosynthetically arise from **K2** by a 1,2-hydride shift to **K2a** and deprotonation (Scheme 15A), or alternatively from **L2** through similar reactions, or from **L1** by 1,3-hydride shift to **L1b** and deprotonation (Scheme 16A). DFT calculations have shown that such 1,3-hydride shifts are only possible for *trans*-fused guaiane

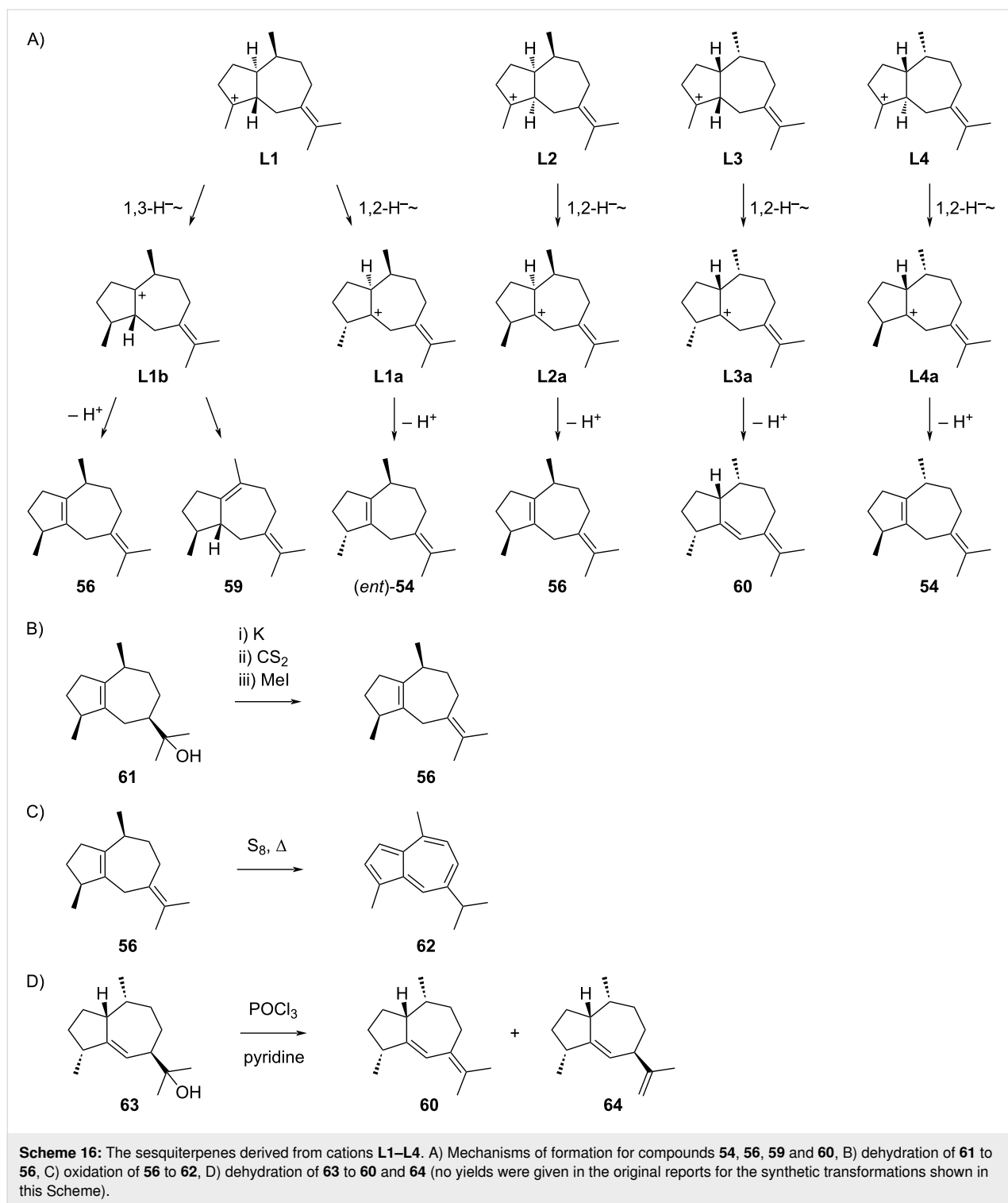


systems [120]. Without detailed knowledge about the structure,  $\beta$ -guaiene (**56**) was first obtained from guaiol (**61**) by Wallach in 1894 [121] and again prepared by Gandurin in 1908 by elimination of the instable methyl xanthogenate (Scheme 16B) [122], followed by an isolation from *Acorus calamus* ( $[\alpha]_D^{20} = +13$ ) by Šorm and co-workers [123]. It is well known that **56** can easily be dehydrogenated, e.g., by heating with sulphur, to the blue azulene derivative **62** (Scheme 16C) [121,122,124–126], but the structure elucidation of this compound was only completed in 1936 [127]. Based on a comparison of IR spectra of natural terpenes, their hydrogenation and dehydrogenation products, the correct planar structure of **56** was concluded by Pliva and Šorm [128]. After the absolute configuration of **61** was solved [129], the full stereostructure of **56** became known.

No total synthesis and no NMR data are available for **56**.  $\beta$ -Guaiene is one of the main constituents of the essential oil from *Achillea millefolium* that shows inhibitory activity against *Babesia canis*, a parasite transmitted by ticks that infects blood cells [130].

Compound **59** is accessible by deprotonation of **L1b**, but only known as synthetic racemic material [113–116]. Compound **60** can be produced by cationic intermediate **L3** through 1,2-hydride shift to **L3a** and deprotonation (Scheme 16A). However, this compound itself is not known as a natural product, but has been obtained together with  $\gamma$ -gurjunene (**64**) from guai-11-en-5-ol (**63**), a natural product isolated from gurjun wood oil, by elimination (Scheme 16D) [131].





## Conclusion

As summarised in this review, the biosynthesis of many sesquiterpene hydrocarbons and alcohols exhibiting the eudesmane or guaiane skeleton can be explained from the neutral intermediate germacrene B, although not all compounds known to literature have been isolated from natural sources; some com-

pounds are only known as synthetic materials. Compared to the known compounds arising from germacrene A or hedycaryol through similar reactions as discussed here [12,13], however, the number of terpenes derived from germacrene B is much lower. In this article we have explained the rationale for the structure elucidation including relative and, if known, absolute

configurations. Through a detailed analysis of the available information it also turned out that some of the assigned structures are doubtful. The importance of rigorous structure elucidation, historically usually performed by chemical correlations and today preferentially done by NMR spectroscopy or X-ray analysis, is clearly evident from the fact that wrongly reported structures or structures assigned without any comprehensible basis lead to error propagations and highly confusing situations in the literature. Today many reports are only based on tentative GC–MS assignments, often even without comparison to authentic standards, which results in a lot of information of questionable relevance. The large number of such papers published today makes it more and more difficult to find the relevant information in the literature. With this work we hope to help the interested reader to have an easier access to the knowledge about sesquiterpenes derived from germacrene B.

## ORCID® iDs

Houchao Xu - <https://orcid.org/0000-0002-4480-2035>

Jeroen S. Dickschat - <https://orcid.org/0000-0002-0102-0631>

## References

- Croteau, R.; Purkett, P. T. *Arch. Biochem. Biophys.* **1989**, *271*, 524–535. doi:10.1016/0003-9861(89)90304-4
- Lynen, F.; Eggerer, H.; Henning, U.; Kessel, I. *Angew. Chem.* **1958**, *70*, 738–742. doi:10.1002/ange.19580702403
- Nandi, D. L.; Porter, J. W. *Arch. Biochem. Biophys.* **1964**, *105*, 7–19. doi:10.1016/0003-9861(64)90230-9
- Tachibana, A. *FEBS Lett.* **1994**, *341*, 291–294. doi:10.1016/0014-5793(94)80475-3
- Tao, H.; Lauterbach, L.; Bian, G.; Chen, R.; Hou, A.; Mori, T.; Cheng, S.; Hu, B.; Lu, L.; Mu, X.; Li, M.; Adachi, N.; Kawasaki, M.; Moriya, T.; Senda, T.; Wang, X.; Deng, Z.; Abe, I.; Dickschat, J. S.; Liu, T. *Nature* **2022**, *606*, 414–419. doi:10.1038/s41586-022-04773-3
- Cane, D. E. *Chem. Rev.* **1990**, *90*, 1089–1103. doi:10.1021/cr00105a002
- Dickschat, J. S. *Nat. Prod. Rep.* **2016**, *33*, 87–110. doi:10.1039/c5np00102a
- Minami, A.; Ozaki, T.; Liu, C.; Oikawa, H. *Nat. Prod. Rep.* **2018**, *35*, 1330–1346. doi:10.1039/c8np00026c
- Ruzicka, L. *Experientia* **1953**, *9*, 357–367. doi:10.1007/bf02167631
- Barton, D. H. R.; de Mayo, P. J. *J. Chem. Soc.* **1957**, 150–158. doi:10.1039/jr9570000150
- Hendrickson, J. B. *Tetrahedron* **1959**, *7*, 82–89. doi:10.1016/0040-4020(59)80055-7
- Xu, H.; Dickschat, J. S. *Chem. – Eur. J.* **2020**, *26*, 17318–17341. doi:10.1002/chem.202002163
- Xu, H.; Dickschat, J. S. *Chem. – Eur. J.* **2022**, *28*, e202200405. doi:10.1002/chem.202200405
- Tsankova, E.; Enev, V. *Tetrahedron* **1987**, *43*, 4425–4432. doi:10.1016/s0040-4020(01)90318-7
- Piet, D. P.; Schrijvers, R.; Franssen, M. C. R.; de Groot, A. *Tetrahedron* **1995**, *51*, 6303–6314. doi:10.1016/0040-4020(95)00272-a
- Pérez Morales, M. C.; Catalán, J. V.; Domingo, V.; Jaraíz, M.; Herrador, M. M.; Quílez del Moral, J. F.; López-Pérez, J.-L.; Barrero, A. F. *Chem. – Eur. J.* **2013**, *19*, 6598–6612. doi:10.1002/chem.201300662
- Barrero, A. F.; Herrador, M. M.; López-Pérez, J.-L.; Arteaga, J. F.; Catalán, J. *Org. Lett.* **2009**, *11*, 4782–4785. doi:10.1021/ol901066x
- Sakui, N.; Kuroyanagi, M.; Ishitobi, Y.; Sato, M.; Ueno, A. *Phytochemistry* **1992**, *31*, 143–147. doi:10.1016/0031-9422(91)83023-e
- Ognjanov, I.; Ivanov, D.; Herout, V.; Horák, M.; Plíva, J.; Šorm, F. *Collect. Czech. Chem. Commun.* **1958**, *23*, 2033–2045. doi:10.1135/cccc19582033
- Brown, E. D.; Solomon, M. D.; Sutherland, J. K.; Torre, A. *Chem. Commun. (London)* **1967**, 111–112. doi:10.1039/c19670000111
- Allen, F. H.; Rogers, D. *Chem. Commun. (London)* **1967**, 588–590. doi:10.1039/c19670000588
- Hartley, R. D.; Fawcett, C. H. *Phytochemistry* **1969**, *8*, 1793–1796. doi:10.1016/s0031-9422(00)85970-1
- Nishimura, K.; Shinoda, N.; Hirose, Y. *Tetrahedron Lett.* **1969**, *10*, 3097–3100. doi:10.1016/s0040-4039(01)88358-1
- Rücker, G.; Silva, G. A. A. B. E.; Bauer, L.; Schikarski, M. *Planta Med.* **1977**, *31*, 322–327. doi:10.1055/s-0028-1097540
- Clark, B. C., Jr.; Chamblee, T. S.; Iacobucci, G. A. *J. Agric. Food Chem.* **1987**, *35*, 514–518. doi:10.1021/jf00076a019
- Adio, A. M.; Paul, C.; Tesso, H.; Kloth, P.; König, W. A. *Tetrahedron: Asymmetry* **2004**, *15*, 1631–1635. doi:10.1016/j.tetasy.2004.03.030
- Badalamenti, N.; Bruno, M.; Formisano, C.; Rigano, D. *Nat. Prod. Commun.* **2022**, *17*, 1934578X221096963. doi:10.1177/1934578x221096963
- Zafra-Polo, M. C.; Blazquez, M. A.; Villar, A. *Fitoterapia* **1989**, *60*, 469–473.
- Gundidza, M.; Chinyanganya, F.; Chagonda, L.; De Pooter, H. L.; Mavi, S. *Flavour Fragrance J.* **1994**, *9*, 299–303. doi:10.1002/ffj.2730090604
- Bleeker, P. M.; Spyropoulou, E. A.; Diergaarde, P. J.; Volpin, H.; De Both, M. T. J.; Zerbe, P.; Bohlmann, J.; Falara, V.; Matsuba, Y.; Pichersky, E.; Haring, M. A.; Schuurink, R. C. *Plant Mol. Biol.* **2011**, *77*, 323–336. doi:10.1007/s11103-011-9813-x
- Zager, J. J.; Lange, I.; Srividya, N.; Smith, A.; Lange, B. M. *Plant Physiol.* **2019**, *180*, 1877–1897. doi:10.1104/pp.18.01506
- Colby, S. M.; Crock, J.; Dowdle-Rizzo, B.; Lemaux, P. G.; Croteau, R. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 2216–2221. doi:10.1073/pnas.95.5.2216
- Picaud, S.; Olsson, M. E.; Brodelius, M.; Brodelius, P. E. *Arch. Biochem. Biophys.* **2006**, *452*, 17–28. doi:10.1016/j.abb.2006.06.007
- Chou, W. K. W.; Fanizza, I.; Uchiyama, T.; Komatsu, M.; Ikeda, H.; Cane, D. E. *J. Am. Chem. Soc.* **2010**, *132*, 8850–8851. doi:10.1021/ja103087w
- Pyle, B. W.; Tran, H. T.; Pickel, B.; Haslam, T. M.; Gao, Z.; MacNevin, G.; Vederas, J. C.; Kim, S.-U.; Ro, D.-K. *FEBS J.* **2012**, *279*, 3136–3146. doi:10.1111/j.1742-4658.2012.08692.x
- Baer, P.; Rabe, P.; Fischer, K.; Citron, C. A.; Klapschinski, T. A.; Groll, M.; Dickschat, J. S. *Angew. Chem., Int. Ed.* **2014**, *53*, 7652–7656. doi:10.1002/anie.201403648
- Haruhisa, S.; Eiji, Ö.; Takeshi, M. *Tetrahedron Lett.* **1979**, *20*, 2245–2246. doi:10.1016/s0040-4039(01)93687-1

38. Nishino, C.; Bowers, W. S.; Montgomery, M. E.; Nault, L. R.; Nielson, M. W. *J. Chem. Ecol.* **1977**, *3*, 349–357. doi:10.1007/bf00988450
39. Faraldos, J. A.; Wu, S.; Chappell, J.; Coates, R. M. *Tetrahedron* **2007**, *63*, 7733–7742. doi:10.1016/j.tet.2007.04.037
40. Rinkel, J.; Dickschat, J. S. *Org. Lett.* **2019**, *21*, 2426–2429. doi:10.1021/acs.orglett.9b00725
41. Xu, H.; Lackus, N. D.; Köllner, T. G.; Dickschat, J. S. *Org. Lett.* **2022**, *24*, 587–591. doi:10.1021/acs.orglett.1c04021
42. Jones, R. V. H.; Sutherland, M. D. *Chem. Commun. (London)* **1968**, 1229–1230. doi:10.1039/c19680001229
43. Brown, E. D.; Sam, T. W.; Sutherland, J. K.; Torre, A. *J. Chem. Soc., Perkin Trans. 1* **1975**, 2326–2332. doi:10.1039/p19750002326
44. Gough, J. H.; Sutherland, M. D. *Aust. J. Chem.* **1964**, *17*, 1270–1281. doi:10.1071/ch9641270
45. Thomas, A. F. *Helv. Chim. Acta* **1972**, *55*, 2429–2432. doi:10.1002/hlca.19720550718
46. Okude, T.; Hayashi, S. *Bull. Chem. Soc. Jpn.* **1970**, *43*, 2984–2985. doi:10.1246/bcsj.43.2984
47. Agarwal, S. G.; Vashist, V. N.; Atal, C. K. *Phytochemistry* **1974**, *13*, 2024–2025. doi:10.1016/0031-9422(74)85160-5
48. Craveiro, A. A.; Andrade, C. H. S.; Matos, F. J. A.; Alencar, J. W. *J. Nat. Prod.* **1979**, *42*, 669–671. doi:10.1021/np50006a014
49. Rucker, G.; Langmann, B.; de Siqueira, N. S. *Planta Med.* **1981**, *41*, 143–149. doi:10.1055/s-2007-971691
50. Sam, T. W.; Sutherland, J. K. *J. Chem. Soc. D* **1971**, 970–971. doi:10.1039/c29710000970
51. Peijnenburg, W. J. G. M.; Dormans, G. J. M.; Buck, H. M. *Tetrahedron* **1988**, *44*, 2339–2350. doi:10.1016/s0040-4020(01)81743-9
52. Minnaard, A. J.; Wijnberg, J. B. P. A.; de Groot, A. *J. Org. Chem.* **1997**, *62*, 7346–7350. doi:10.1021/jo970902r
53. Šorm, F.; Suchý, M.; Vonašek, F.; Plíva, J.; Herout, V. *Collect. Czech. Chem. Commun.* **1951**, *16*, 268–277. doi:10.1135/cccc19510268
54. Hartley, R. D.; Fawcett, C. H. *Phytochemistry* **1969**, *8*, 637–643. doi:10.1016/s0031-9422(00)85414-x
55. Hendriks, H.; Malingré, T. M.; Batterman, S.; Bos, R. *Phytochemistry* **1975**, *14*, 814–815. doi:10.1016/0031-9422(75)83045-7
56. Ganter, C.; Keller-Wojtkiewicz, B. *Helv. Chim. Acta* **1971**, *54*, 183–206. doi:10.1002/hlca.19710540117
57. Endo, J.; Nagasawa, M. *Yakugaku Zasshi* **1974**, *94*, 1574–1579. doi:10.1248/yakushi1947.94.12\_1574
58. Könen, P. P.; Wüst, M. *J. Agric. Food Chem.* **2020**, *68*, 8936–8941. doi:10.1021/acs.jafc.0c03273
59. Wang, C.-C.; Kuoh, C.-S.; Wu, T.-S. *J. Nat. Prod.* **1996**, *59*, 409–411. doi:10.1021/np9601824
60. Křepinský, J.; Herout, V. *Collect. Czech. Chem. Commun.* **1962**, *27*, 2459–2462. doi:10.1135/cccc19622459
61. Kugler, E.; Kováts, E. s. *Helv. Chim. Acta* **1963**, *46*, 1480–1513. doi:10.1002/hlca.19630460506
62. Nigam, M. C.; Nigam, I. C.; Levi, L.; Handa, K. L. *Can. J. Chem.* **1964**, *42*, 2610–2615. doi:10.1139/v64-382
63. Halim, A. F.; Collins, R. P. *Phytochemistry* **1973**, *12*, 1077–1083. doi:10.1016/0031-9422(73)85019-8
64. Bohlmann, F.; Zdero, C.; Pickard, J.; Robinson, H.; King, R. M. *Phytochemistry* **1981**, *20*, 1323–1333. doi:10.1016/0031-9422(81)80032-5
65. Lorimer, S. D.; Weavers, R. T. *Phytochemistry* **1987**, *26*, 3207–3215. doi:10.1016/s0031-9422(00)82471-1
66. Steele, C. L.; Crock, J.; Bohlmann, J.; Croteau, R. *J. Biol. Chem.* **1998**, *273*, 2078–2089. doi:10.1074/jbc.273.4.2078
67. Booth, J. K.; Yuen, M. M. S.; Jancsik, S.; Madilao, L. L.; Page, J. E.; Bohlmann, J. *Plant Physiol.* **2020**, *184*, 130–147. doi:10.1104/pp.20.00593
68. Rabe, P.; Dickschat, J. S. *Angew. Chem., Int. Ed.* **2013**, *52*, 1810–1812. doi:10.1002/anie.201209103
69. Wang, Y.-H.; Xu, H.; Zou, J.; Chen, X.-B.; Zhuang, Y.-Q.; Liu, W.-L.; Celik, E.; Chen, G.-D.; Hu, D.; Gao, H.; Wu, R.; Sun, P.-H.; Dickschat, J. S. *Nat. Catal.* **2022**, *5*, 128–135. doi:10.1038/s41929-022-00735-0
70. Posner, G. H.; Loomis, G. L.; Sawaya, H. S. *Tetrahedron Lett.* **1975**, *16*, 1373–1376. doi:10.1016/s0040-4039(00)72146-0
71. van Beek, T. A.; Kleis, R.; Posthumus, M. A.; van Veldhuizen, A. *Phytochemistry* **1989**, *28*, 1909–1911. doi:10.1016/s0031-9422(00)97885-3
72. Vanderah, D. J.; Rutledge, N.; Schmitz, F. J.; Ciereszko, L. S. *J. Org. Chem.* **1978**, *43*, 1614–1616. doi:10.1021/jo00402a040
73. Suzuki, T.; Kikuchi, H.; Kurosawa, E. *Chem. Lett.* **1980**, *9*, 1267–1270. doi:10.1246/cl.1980.1267
74. Feger, W.; Brandauer, H.; Ziegler, M. *J. Essent. Oil Res.* **1999**, *11*, 556–562. doi:10.1080/10412905.1999.9701213
75. Motl, O.; Herout, V.; Šorm, F. *Collect. Czech. Chem. Commun.* **1958**, *23*, 1293–1296. doi:10.1135/cccc19581293
76. Varma, K. R.; Jain, T. C.; Bhattacharyya, S. C. *Tetrahedron* **1962**, *18*, 979–984. doi:10.1016/s0040-4020(01)92752-8
77. Chetty, G. L.; Zalkow, V. B.; Zalkow, L. H. *Tetrahedron Lett.* **1968**, *9*, 3223–3225. doi:10.1016/s0040-4039(00)89530-1
78. Lassak, E. V. *Phytochemistry* **1969**, *8*, 2097. doi:10.1016/s0031-9422(00)88102-9
79. Bohlmann, F.; Zdero, C.; King, R. M.; Robinson, H. *Phytochemistry* **1982**, *21*, 147–150. doi:10.1016/0031-9422(82)80032-0
80. Barbosa-Filho, J. M.; Yoshida, M.; Gottlieb, O. R.; Barbosa, R. d. C. S. B. C.; Giesbrecht, A. M.; Young, M. C. M. *Phytochemistry* **1987**, *26*, 2615–2617. doi:10.1016/s0031-9422(00)83890-x
81. Barrero, A. F.; Sánchez, J. F.; Oltra, J. E.; Altarejos, J.; Ferrol, N.; Barragán, A. *Phytochemistry* **1991**, *30*, 1551–1554. doi:10.1016/0031-9422(91)84207-9
82. Raharivelomanana, P.; Bianchini, J.-P.; Ramanoelina, A. R. P.; Rasoarahona, J. R. E.; Faure, R.; Cambon, A. *Phytochemistry* **1998**, *47*, 1085–1088. doi:10.1016/s0031-9422(98)80077-0
83. Weyerstahl, P.; Schneider, S.; Marschall, H. *Flavour Fragrance J.* **1996**, *11*, 81–94. doi:10.1002/(sici)1099-1026(199603)11:2<81::aid-ffj553>3.0.co;2-8
84. Bastos, J. K.; Gottlieb, O. R.; Sarti, S. J.; Filho, D. S. *Nat. Prod. Lett.* **1996**, *9*, 65–70. doi:10.1080/10575639608043580
85. Toyota, M.; Saito, T.; Asakawa, Y. *Phytochemistry* **1999**, *51*, 913–920. doi:10.1016/s0031-9422(99)00159-4
86. Hayashi, S.; Yano, K.; Hayashi, N.; Matsuura, T. *Bull. Chem. Soc. Jpn.* **1968**, *41*, 1465–1466. doi:10.1246/bcsj.41.1465
87. Wang, S.-Y.; Ding, L.-F.; Su, J.; Peng, L.-Y.; Song, L.-D.; Wu, X.-D. *Phytochem. Lett.* **2018**, *23*, 127–131. doi:10.1016/j.phytol.2017.11.021
88. Ren, F.; Mao, H.; Liang, J.; Liu, J.; Shu, K.; Wang, Q. *Planta* **2016**, *244*, 1065–1074. doi:10.1007/s00425-016-2570-y
89. Dutta, M.; Nezam, M.; Chowdhury, S.; Rakib, A.; Paul, A.; Sami, S. A.; Uddin, M. Z.; Rana, M. S.; Hossain, S.; Effendi, Y.; Idroes, R.; Tallei, T.; Alqahtani, A. M.; Emran, T. B. *Front. Mol. Biosci.* **2021**, *8*, 625391. doi:10.3389/fmolb.2021.625391

90. Maupetit, P. *Perfum. Flavor* **1985**, *9*, 19–37.
91. Adio, A. M.; König, W. A. *Phytochemistry* **2005**, *66*, 599–609. doi:10.1016/j.phytochem.2005.01.015
92. Rieck, A.; Bülow, N.; Jung, S.; Saritas, Y.; König, W. A. *Phytochemistry* **1997**, *44*, 453–457. doi:10.1016/s0031-9422(96)00482-7
93. Zhao, Y.; Yue, J.; Lin, Z.; Ding, J.; Sun, H. *Phytochemistry* **1997**, *44*, 459–464. doi:10.1016/s0031-9422(96)00521-3
94. Adio, A. M.; Paul, C.; König, W. A.; Muhle, H. *Phytochemistry* **2003**, *64*, 637–644. doi:10.1016/s0031-9422(03)00298-x
95. Jommi, G.; Křepinský, J.; Herout, V.; Šorm, F. *Tetrahedron Lett.* **1967**, *8*, 677–681. doi:10.1016/s0040-4039(00)90572-0
96. Takahashi, S. *Chem. Pharm. Bull.* **1968**, *16*, 2447–2450. doi:10.1248/cpb.16.2447
97. Maurer, B.; Fracheboud, M.; Grieder, A.; Ohloff, G. *Helv. Chim. Acta* **1972**, *55*, 2371–2382. doi:10.1002/hlca.19720550712
98. Hanayama, N.; Kido, F.; Sakuma, R.; Uda, H.; Yoshikoshi, A. *Tetrahedron Lett.* **1968**, *9*, 6099–6102. doi:10.1016/s0040-4039(00)70804-5
99. Woof, C. R.; Durand, D. J.; Fey, N.; Richards, E.; Webster, R. L. *Chem. – Eur. J.* **2021**, *27*, 5972–5977. doi:10.1002/chem.202004980
100. Lago, J. H. G.; Brochini, C. B.; Roque, N. F. *J. Essent. Oil Res.* **2005**, *17*, 271–273. doi:10.1080/10412905.2005.9698899
101. Crockett, S. L.; Demirci, B.; Başer, K. H. C.; Khan, I. A. *J. Essent. Oil Res.* **2007**, *19*, 302–306. doi:10.1080/10412905.2007.9699287
102. Ludwiczuk, A.; Nagashima, F.; Gradstein, R. S.; Asakawa, Y. *Nat. Prod. Commun.* **2008**, *3*, 133–140. doi:10.1177/1934578x0800300205
103. Wajs-Bonikowska, A.; Sienkiewicz, M.; Stobiecka, A.; Maciag, A.; Szoka, L.; Karna, E. *Chem. Biodiversity* **2015**, *12*, 407–418. doi:10.1002/cbdv.201400167
104. Ha, C. T. T.; Diep, L. N.; Thuy, D. T. T.; Bon, T. N.; Cham, L. T. T.; Setzer, W. N. *Rec. Nat. Prod.* **2022**, *16*, 503–508. doi:10.25135/rnp.305.2112.2299
105. Ludwiczuk, A.; Komala, I.; Pham, A.; Bianchini, J.-P.; Raharivelomanana, P.; Asakawa, Y. *Nat. Prod. Commun.* **2009**, *4*, 1387–1392. doi:10.1177/1934578x0900401015
106. Zhao, Y.; Schenk, D. J.; Takahashi, S.; Chappell, J.; Coates, R. M. *J. Org. Chem.* **2004**, *69*, 7428–7435. doi:10.1021/jo049058c
107. Hwu, J. R.; Wetzol, J. M. *J. Org. Chem.* **1992**, *57*, 922–928. doi:10.1021/jo00029a025
108. Bates, R. B.; Slagel, R. C. *Chem. Ind. (London)* **1962**, 1715–1716.
109. Akhila, A.; Sharma, P. K.; Thakur, R. S. *Phytochemistry* **1988**, *27*, 2105–2108. doi:10.1016/0031-9422(88)80105-5
110. Dolejš, L.; Mironov, A.; Šorm, F. *Tetrahedron Lett.* **1960**, *1* (32), 18–21. doi:10.1016/s0040-4039(01)84073-9
111. Bates, R. B.; Slagel, R. C. *J. Am. Chem. Soc.* **1962**, *84*, 1307–1308. doi:10.1021/ja00866a051
112. Heathcock, C. H.; Ratcliffe, R. *J. Am. Chem. Soc.* **1971**, *93*, 1746–1757. doi:10.1021/ja00736a029
113. Oppolzer, W.; Wylie, R. D. *Helv. Chim. Acta* **1980**, *63*, 1198–1203. doi:10.1002/hlca.19800630511
114. Sammes, P. G.; Street, L. J. *J. Chem. Soc., Chem. Commun.* **1983**, 666–668. doi:10.1039/c39830000666
115. Bromidge, S. M.; Sammes, P. G.; Street, L. J. *J. Chem. Soc., Perkin Trans. 1* **1985**, 1725–1730. doi:10.1039/p19850001725
116. Negishi, E.-i.; Ma, S.; Sugihara, T.; Noda, Y. *J. Org. Chem.* **1997**, *62*, 1922–1923. doi:10.1021/jo970009s
117. Abreu, P. M.; Noronha, R. G. *Flavour Fragrance J.* **1997**, *12*, 79–83. doi:10.1002/(sici)1099-1026(199703)12:2<79::aid-ffj617>3.0.co;2-7
118. Acree, W. E.; Chickos, J. S. *Phase Transition Enthalpy Measurements of Organic and Organometallic Compounds*. In *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*; Linstrom, P. J.; Mallard, W. G., Eds.; National Institute of Standards and Technology: Gaithersburg, MD, USA.
119. de Moraes, S. M.; Facundo, V. A.; Braz Filho, R. *J. Essent. Oil Res.* **2002**, *14*, 274–275. doi:10.1080/10412905.2002.9699851
120. Xu, H.; Goldfuss, B.; Dickschat, J. S. *Chem. – Eur. J.* **2021**, *27*, 9758–9762. doi:10.1002/chem.202101371
121. Tuttle, F. E. *Justus Liebigs Ann. Chem.* **1894**, *279*, 391–397. doi:10.1002/jlac.18942790315
122. Gandurin, A. *Ber. Dtsch. Chem. Ges.* **1908**, *41*, 4359–4363. doi:10.1002/cber.190804103163
123. Šorm, F.; Holub, M.; Sýkora, V.; Mleziva, J.; Streibl, M.; Plíva, J.; Schneider, B.; Herout, V. *Collect. Czech. Chem. Commun.* **1953**, *18*, 512–526. doi:10.1135/cccc19530512
124. Schroeter, G.; Lichtenstadt, L.; Irineu, D. *Ber. Dtsch. Chem. Ges.* **1918**, *51*, 1587–1613. doi:10.1002/cber.19180510247
125. Ruzicka, L.; Rudolph, E. A. *Helv. Chim. Acta* **1926**, *9*, 118–140. doi:10.1002/hlca.19260090114
126. Ruzicka, L.; Haagen-Smit, A. J. *Helv. Chim. Acta* **1931**, *14*, 1104–1122. doi:10.1002/hlca.19310140520
127. Pfau, A. S.; Plattner, P. *Helv. Chim. Acta* **1936**, *19*, 858–879. doi:10.1002/hlca.193601901117
128. Plíva, J.; Šorm, F. *Collect. Czech. Chem. Commun.* **1949**, *14*, 274–286. doi:10.1135/cccc19490274
129. Minato, H. *Tetrahedron* **1962**, *18*, 365–371. doi:10.1016/s0040-4020(01)93252-1
130. Guz, L.; Wawrzykowski, J.; Adaszek, L. *Pol. J. Vet. Sci.* **2021**, *24*, 79–84. doi:10.24425/pjvs.2021.136795
131. Rücker, G.; Hefendehl, F. W. *Phytochemistry* **1978**, *17*, 809–810. doi:10.1016/s0031-9422(00)94241-9

## License and Terms

This is an open access article licensed under the terms of the Beilstein-Institut Open Access License Agreement (<https://www.beilstein-journals.org/bjoc/terms>), which is identical to the Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0>). The reuse of material under this license requires that the author(s), source and license are credited. Third-party material in this article could be subject to other licenses (typically indicated in the credit line), and in this case, users are required to obtain permission from the license holder to reuse the material.

The definitive version of this article is the electronic one which can be found at: <https://doi.org/10.3762/bjoc.19.18>

## Appendix D

### Isotopic Labeling Experiments Solve the Hedycaryol Problem

*Org. Lett.* **2022**, *24*, 587

DOI: [10.1021/acs.orglett.1c04021](https://doi.org/10.1021/acs.orglett.1c04021)

# Isotopic Labeling Experiments Solve the Hedycaryol Problem

Houchao Xu, Nathalie D. Lackus, Tobias G. Köllner, and Jeroen S. Dickschat\*



Cite This: *Org. Lett.* 2022, 24, 587–591



Read Online

ACCESS |



Metrics & More

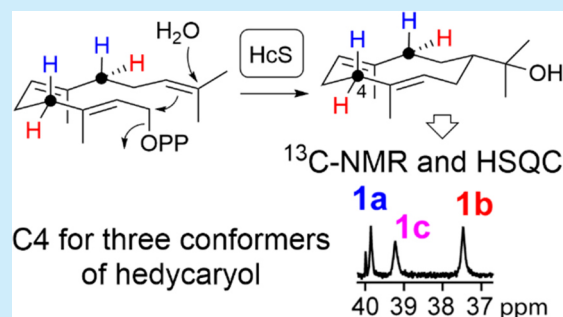


Article Recommendations



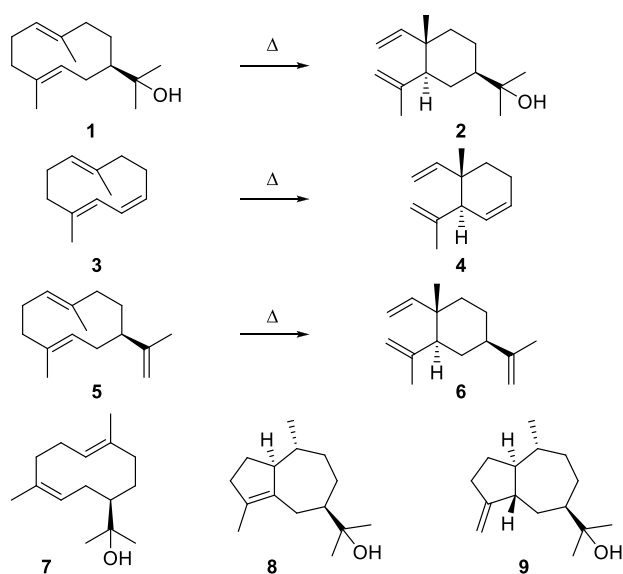
Supporting Information

**ABSTRACT:** Hedycaryol is a widespread sesquiterpene alcohol and important biosynthetic intermediate toward eudesmols and guaiols. A full NMR assignment for this compound has been hampered because of the unique molecular mechanics of its conformers in complex mixtures. This problem was solved through the enzymatic synthesis of isotopically labeled materials using a mutated plant and a bacterial enzyme for access to both enantiomers of hedycaryol, which also allowed us to follow the stereochemical course of its Cope rearrangement.



In 1916, Semmler reported on the monocyclic sesquiterpene alcohol elemol (**2**, Scheme 1) from the tree *Canarium*

## Scheme 1. Cope Rearrangements of 10-Membered Sesquiterpenes and Structures of 7–9



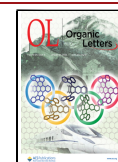
*luzonicum* (elemi) that is native to the Philippines.<sup>1</sup> The compound was also found in large quantities (60%) in the essential oil from *Hedycarya angustifolia*.<sup>2</sup> After the discovery that geijerene (**4**) from *Geijera parviflora*<sup>3</sup> is the product of a thermal (Cope) rearrangement of pregeijerene (**3**),<sup>4</sup> **2** was also shown to stand in a similar relation to hedycaryol (**1**).<sup>5</sup> Cope rearrangements also explain the conversion of germacrene A (**5**) to elemene (**6**)<sup>6</sup> and are known for several other

germacrane sesquiterpenes.<sup>7,8</sup> Most germacrane exhibit a set of fairly rigid, slowly interconverting conformers, causing line broadening in the NMR spectra and often multiple sets of signals depending on their abundance up to one for each conformer.<sup>7</sup> Despite the fact that **5** is a widespread natural product and an important biosynthetic intermediate toward many eudesmanes, guaianes, and related sesquiterpenes,<sup>9</sup> several attempts to assign the NMR data for **5** from the mixture of conformers gave only partial data sets, albeit of better and better quality with the development of NMR techniques.<sup>6,10–12</sup> However, even with the high-field NMR technology available today, the problem still remained largely open but was recently solved by us for **5** using extensive <sup>13</sup>C and stereoselective <sup>2</sup>H labeling experiments.<sup>13</sup>

Not only the sesquiterpene alcohol **1** but also its three geometrical double-bond isomers show similar transannular reactions, as observed for **5**, including thermal Cope rearrangements, and are observed as a mixture of strained conformers.<sup>14,15</sup> Although **1**<sup>16</sup> and its double-bond stereoisomers<sup>17,18</sup> have been synthesized, very limited <sup>1</sup>H NMR data only for Me groups and olefinic hydrogen atoms of the isolated **1** have been reported.<sup>19</sup> Plant sesquiterpene synthases (STPSs) converting farnesyl diphosphate (FPP) into **1** as the main product are known from *Populus trichocarpa* (PtTPS7),<sup>20</sup> *Camellia brevistyla* (CbTPS1),<sup>21</sup> and *Liquidambar formosana* (LfTPS01),<sup>22</sup> whereas the 2*Z*,6*E* stereoisomer **7** was tentatively assigned by GC/MS as the product of a bacterial

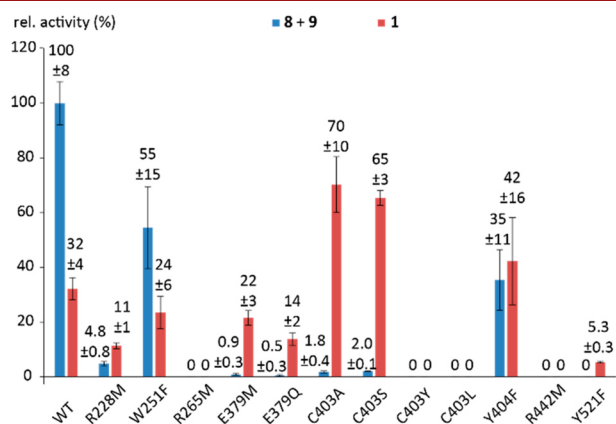
Received: November 26, 2021

Published: January 5, 2022



STPS from *Kitasatospora setae* (HcS), in which case the crystal structure has also been obtained.<sup>23</sup> We have recently identified another STPS from *P. trichocarpa* (PtTPSS) that produces a mixture of (1*S*,5*S*,7*R*,10*R*)-guaia-4(15)-en-11-ol (**8**) and (1*S*,7*R*,10*R*)-guaia-4-en-11-ol (**9**)<sup>24</sup> and studied the cyclization mechanism that proceeds through (*R*)-**1**.<sup>25</sup> Now we report on site-directed mutagenesis experiments on PtTPSS yielding *inter alia* an enzyme variant (PtTPSS-C403A) that selectively produces (*R*)-**1**. We further report on the NMR-based structure assignment for the HcS product as (*S*)-**1**. Isotopic labeling experiments were performed to assign full NMR data for **1** and to study its unique chemistry.

We aimed at the usage of two hedycaryol synthases, one from a plant and one from a bacterium (Figure S1), as plant and bacterial enzymes usually make different enantiomers.<sup>9,26,27</sup> PtTPSS from *P. trichocarpa* was considered to be a good starting point because for tobacco *5-epi-aristolochene* synthase (TEAS), the exchange of the active site residue Y520 with Phe (Y520F) has been demonstrated to interrupt the cyclization cascade at the intermediate **5**, suggesting its involvement in the reprotonation for the second cyclization to *5-epi-aristolochene*.<sup>28</sup> Also, for other 1,10-cyclizing plant terpene synthases that produce eudesmane or guaiane sesquiterpene hydrocarbons or alcohols, the accumulation of intermediate **1** or **5**, respectively, was observed upon the exchange of the corresponding Tyr residue.<sup>29–31</sup> With FPP as the substrate, the enzyme variant Y521F of PtTPSS also resulted in an abolished production of **8** and **9** with accumulation of **1**, but the production was strongly reduced ( $5.3 \pm 0.3\%$  of wild-type level of **8** + **9**, Figure 1); therefore,



**Figure 1.** Relative activity of enzyme variants of PtTPSS. The production of **8** + **9** by the wild type (WT) was set to 100%. The bars show the mean and standard deviation in triplicate.

this mutant was not suitable for labeling experiments. As previously observed for pseudolaratriene synthase (PxaTPS8) from *Pseudolarix amabilis*,<sup>32</sup> mutation of the first Arg in the conserved RxR motif<sup>33</sup> (R265M) gave an inactive enzyme. Also, mutation of the third active-site Arg<sup>33</sup> three positions before the highly conserved DTE (Asp-Thr-Glu) triad (R442M) resulted in an inactive protein. Further highly conserved active-site residues include W251, R288, and E379 (equal to W250, R287, and E379 in TEAS). The exchange W251F showed decreased activity ( $55 \pm 15\%$  for **8** + **9**) with a nearly unchanged production ratio for intermediate **1**, whereas the activity for R288M was nearly lost with slightly increased relative amounts of **1** ( $4.8 \pm 0.8\%$  **8** + **9**,  $11 \pm 1\%$  **1**). The

enzyme variants E379M and E379Q lost the ability to produce **8** and **9** almost completely, with some retained production of **1** ( $22 \pm 3$  and  $14 \pm 2\%$ , respectively), establishing the importance of this residue that shows a long-range interaction with  $Mg^{2+}$  in the TEAS structure<sup>33</sup> for catalysis. The observed amino acid residues near the helix G kink in plant type I terpene synthases are diverse, and it has been shown in several mutational studies that exchanges in this region can dramatically influence the product profile.<sup>34–37</sup> Specifically, the exchanges T409G in *Zea mays* eudesmanediol synthase (ZmEDS) and H415A in *Senecio scandens* liguloxide synthase (SsLOS) caused a product shift toward the intermediate **1**.<sup>30,31</sup> The substitutions C403Y and C403L (equal to T409 in ZmEDS) yielded inactive proteins, whereas C403A and C403S lost the ability to form **8** and **9**, with a shift toward **1** in good yields ( $70 \pm 10\%$  and  $65 \pm 3\%$ , respectively). The exchange Y404F (equal to H415 in SsLOS) showed a moderately reduced activity, with **1** and **8** + **9** produced in nearly equal amounts. Thus the PtTPSS-C403A variant was selected for further studies on **1**.

Notably, the enzyme product obtained from FPP with HcS showed no difference from the product from PtTPSS-C403A in a comparison of mass spectra and GC retention times (Figure S2). The thermal Cope reaction also explains the unique behavior of **1** in gas chromatographic analyses with partial rearrangement to **2** in the injector, forming a sharp peak with a short retention time, and partial rearrangement on the GC column, forming a second broad peak of later retention time. Together with the results from the NMR analyses discussed as follows, the previously tentatively identified product **7** of HcS<sup>23</sup> must be reassigned as **1**. On the basis of the known absolute configurations of **8** and **9**<sup>24</sup> the product of PtTPSS-C403A can be assigned as (*R*)-**1**. The optical rotation of the HcS product,  $[\alpha]_D^{25} = -21.3$  (*c* 0.29,  $C_6H_6$ ), in comparison with that for (*R*)-**1** from *H. angustifolia*,  $[\alpha]_D^{25} = +30.8$  ( $CHCl_3$ ),<sup>5</sup> pointed to the absolute configuration of (*S*)-**1** for the bacterial compound. Its Cope rearrangement yielded (+)-**2** (*c* 0.21,  $C_6H_6$ ), lit. for (–)-**2** from *H. angustifolia*:  $[\alpha]_D^{25} = -4.5$  ( $CHCl_3$ );<sup>5</sup> this first isolation paper shows the wrong absolute configurations for **1** and **2**, despite the fact that the absolute configuration of **2** had been clarified 4 years before<sup>38</sup>). These data confirmed the structural reassignment for the HcS product because the previously assigned **7** is known to form *cis*-isoelemol through Cope rearrangement,<sup>14</sup> a stereoisomer of **2**, but not **2**. Cope rearrangement of the PtTPSS-C403A product gave (–)-**2**,  $[\alpha]_D^{25} = -1.4$  (*c* 0.14,  $C_6H_6$ ), with identical NMR data to those for the rearranged material from HcS. GC analysis using a chiral stationary phase (Figure S11) confirmed the enantiomeric relationship and purity of both Cope rearrangement products.

Hedycaryol (**1**) is difficult to purify, as it can undergo acid-induced transannular cyclization reactions<sup>5,14</sup> even under only very mildly acidic conditions of column chromatographic purifications on silica gel. Nevertheless, the conversion of FPP by HcS gives high yields of (*S*)-**1** (43%) of sufficient purity for direct NMR analysis, but the spectra show line broadening (Figures S12 and S13), and the data assignment is only possible to a very limited extent. Therefore, the signals of the individual carbon atoms of **1** were enhanced using <sup>13</sup>C labelings that were enzymatically introduced through the conversion of all 15 isotopomers of (<sup>13</sup>C)FPP<sup>39</sup> with HcS. The <sup>13</sup>C NMR spectra of the obtained products show one

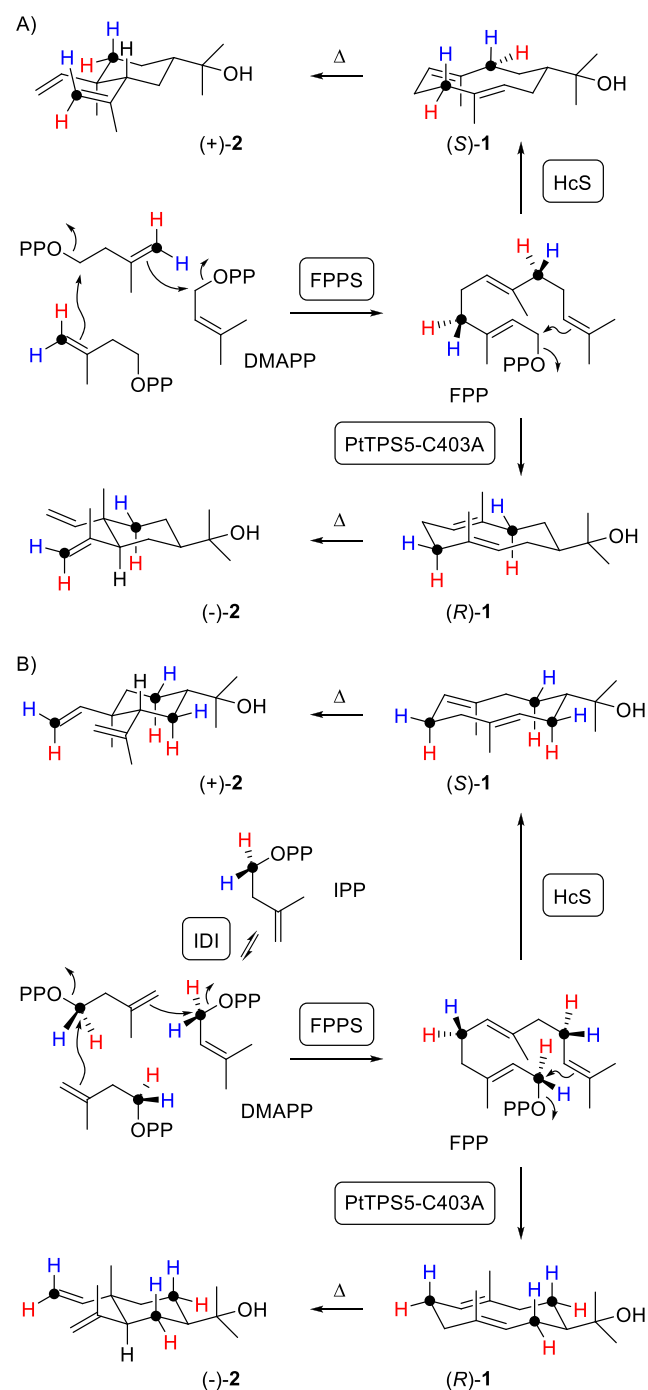
comparably sharp and two broad signals for the three conformers of **1** in each case (Figures S14–S16), but in this stage, it was not fully clear which group of 15 signals belonged to one conformer of **1**. Therefore, completely labeled ( $^{13}\text{C}_{15}$ )FPP<sup>39</sup> was converted with HcS to ( $^{13}\text{C}_{15}$ )-**1**, which gave three contiguous spin systems in the  $^{13}\text{C}$ , $^{13}\text{C}$ -COSY analysis (Figures S17–S19) and thus allowed us to fully assign the  $^{13}\text{C}$  NMR data to the three conformers of **1** (Figures S20–S22, Tables S4–S6). Furthermore, the heteronuclear single quantum coherence (HSQC) spectra of the 15 singly labeled ( $^{13}\text{C}$ )-**1** isotopomers showed strongly enhanced cross-peaks for the attached hydrogen atoms at the labeled carbon atoms (Figures S23–S31). The enzymatic conversion of dimethylallyl pyrophosphate (DMAPP) with (*E*)-(4- $^{13}\text{C}$ ,4- $^2\text{H}$ )IPP<sup>40</sup> (blue H =  $^2\text{H}$ ) and (*Z*)-(4- $^{13}\text{C}$ ,4- $^2\text{H}$ )IPP<sup>40</sup> (red H =  $^2\text{H}$ ) using FPP synthase (FPPS) from *Streptomyces coelicolor*<sup>41</sup> and HcS then resulted in the vanishing of one of the diastereotopic hydrogen atoms within the methylene groups C4 and C8 of (*S*)-**1** (Scheme 2A, Figures S25 and S28). Similar experiments with (*R*)-(1- $^{13}\text{C}$ ,1- $^2\text{H}$ )IPP<sup>13</sup> (blue H =  $^2\text{H}$ ) and (*S*)-(1- $^{13}\text{C}$ ,1- $^2\text{H}$ )IPP<sup>13</sup> (red H =  $^2\text{H}$ ) using the enzyme mix of *Escherichia coli* isopentenyl diphosphate isomerase (IDI),<sup>42</sup> FPPS, and HcS gave corresponding results for the methylene groups C1, C5, and C9, ultimately allowing a full assignment of all hydrogen signals for each of the three conformers of (*S*)-**1** (Scheme 2B, Figures S23, S26, and S29).

Corresponding experiments with PtTPS5-C403A showed incorporations into the same carbon atoms from the 15 isotopomers of ( $^{13}\text{C}$ )FPP (Figures S32–S34) and the same spin systems in the  $^{13}\text{C}$ , $^{13}\text{C}$ -COSY for the product from ( $^{13}\text{C}_{15}$ )FPP (Figures S35–S37). Because the product (*R*)-**1** is the enantiomer of the HcS product, the stereoselective labeling experiments with (*E*)- and (*Z*)-(4- $^{13}\text{C}$ ,4- $^2\text{H}$ )IPP and (*R*)- and (*S*)-(1- $^{13}\text{C}$ ,1- $^2\text{H}$ )IPP gave opposite outcomes for the incorporations into the diastereotopic methylene group hydrogen atoms (Scheme 2 and Figures S23, S25, S26, S28, and S29).

A comparison of the  $^{13}\text{C}$  NMR data of the conformers of **1** to those reported previously for the conformers of **5**, for which single  $^{13}\text{C}$  labelings have also been performed for each position,<sup>13</sup> indicates that for both compounds, the same set of three conformers exists, allowing for a structural assignment of the conformers of **1**. Specifically, the  $^{13}\text{C}$  NMR chemical shifts for each  $^{13}\text{C}$ -labeling experiment for **1** and **5** are very similar for most carbon atoms, appear in the same order for the three conformers, and always show one sharp and two broad peaks. The three conformers of **1**, designated as **1a–1c** (Scheme 3), are observed in a ratio of 20% **1a** (DD = both Me groups down, double bonds crossed), 45% **1b** (DU = Me groups down and up, double bonds parallel), and 35% **1c** (UD, parallel) based on a peak integration for the signals of olefinic hydrogen atoms in the  $^1\text{H}$  NMR, whereas **1d** (UU, crossed) was not observed. The Cope rearrangement to **2** must start from the minor conformer **1a** to establish the correct stereochemistry in the product. This has been pointed out before based on a computational study, suggesting that **1a** is not the most stable conformer, but the computed transition-state energy is lowest for the Cope rearrangement of **1a**.<sup>43</sup>

Further labeling experiments were carried out to investigate the Cope rearrangement of **1** to **2** in more detail. For this purpose, all 15 enzymatically prepared isotopomers of ( $^{13}\text{C}$ )-**1** were heated to 130 °C in  $\text{C}_6\text{D}_6$  in a sealed pressure tube. Each experiment gave a clean conversion to the corresponding isotopomer of ( $^{13}\text{C}$ )-**2**, for which sharp signals were observed

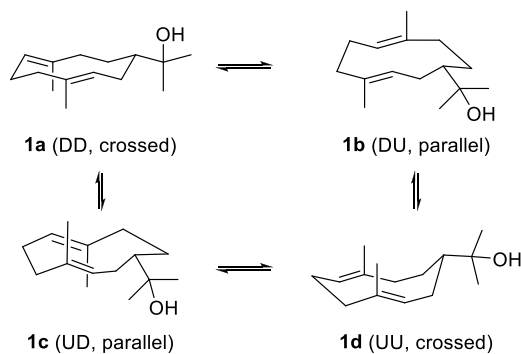
## Scheme 2. Stereoselective Deuteration Experiments for Enzymatic Reactions to **1** and Cope Rearrangement to **2**



in the  $^{13}\text{C}$  NMR (Figure S38). Cope rearrangement of the labeled isotopomers of **1** obtained from stereoselectively deuterated IPP isotopomers with HcS and PtTPS5-C403A gave a specific incorporation into the diastereotopic hydrogen positions of all methylene groups of **2**. Notably, for the HcS and the PtTPS5-C403A products, again, the opposite outcome was observed in all experiments, reflecting the enantiomeric relation of their products (*S*)-**1** and (*R*)-**1**. HSQC analysis of all eight Cope rearrangement products (from four stereoselectively deuterated IPP isotopomers with two enzymes, Figures S39–S43) not only confirmed the correctness of the nuclear Overhauser effect spectroscopy (NOESY)-based



## Scheme 3. Conformers of Hedycaryol (S)-1



assignments for the diastereotopic hydrogen atoms of **2** (Table S2) but also showed a clear stereochemical course for the C4 and C5 hydrogen atoms in the Cope rearrangement (Scheme 2) that is in agreement with the chair–chair transition state starting from conformer **1a**.

### ■ ASSOCIATED CONTENT

#### Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.orglett.1c04021>.

Tabulated NMR data and spectra of **1** and **2**, experimental procedures, and results from labeling experiments (PDF)

### ■ AUTHOR INFORMATION

#### Corresponding Author

Jeroen S. Dickschat – *Kekulé Institute for Organic Chemistry and Biochemistry, University of Bonn, 53121 Bonn, Germany*; [orcid.org/0000-0002-0102-0631](https://orcid.org/0000-0002-0102-0631);  
Email: [dickschat@uni-bonn.de](mailto:dickschat@uni-bonn.de)

#### Authors

Houchao Xu – *Kekulé Institute for Organic Chemistry and Biochemistry, University of Bonn, 53121 Bonn, Germany*  
Nathalie D. Lackus – *Max Planck Institute for Chemical Ecology, 07745 Jena, Germany*  
Tobias G. Köllner – *Max Planck Institute for Chemical Ecology, 07745 Jena, Germany*

Complete contact information is available at: <https://pubs.acs.org/10.1021/acs.orglett.1c04021>

#### Author Contributions

H.X. performed the labeling experiments and analysis of enzyme variants. N.D.L. and T.G.K. constructed the expression plasmids for enzyme variants. J.S.D. designed and coordinated the research. The manuscript was written by J.S.D. with contributions and approval by all authors.

#### Notes

The authors declare no competing financial interest.

### ■ ACKNOWLEDGMENTS

This work was funded by the DFG (DI1536/7-2).

### ■ REFERENCES

- Semmler, F. W.; Liao, F. *Chem. Ber.* **1916**, *49*, 794–798.
- Hellyer, R. O. *Aust. J. Chem.* **1962**, *15*, 157.

- Sutherland, M. D. *Chem. Ind.* **1959**, 1220–1221.
- Jones, R. V. H.; Sutherland, M. D. *Aust. J. Chem.* **1968**, *21*, 2255–2264.
- Jones, R. V. H.; Sutherland, M. D. *Chem. Commun.* **1968**, 1229–1230.
- Weinheimer, A. J.; Youngblood, W. W.; Washecheck, P. H.; Karns, T. K. B.; Ciereszko, L. S. *Tetrahedron* **1970**, *11*, 497–500.
- Takeda, K. *Tetrahedron* **1974**, *30*, 1525–1534.
- Adio, A. M. *Tetrahedron* **2009**, *65*, 1533–1552.
- Xu, H.; Dickschat, J. S. *Chem.—Eur. J.* **2020**, *26*, 17318–17341.
- Nishino, C.; Bowers, W. S.; Montgomery, M. E.; Nault, L. R.; Nielson, M. W. *J. Chem. Ecol.* **1977**, *3*, 349–357.
- Adio, A. M.; Paul, C.; Tesso, H.; Kloth, P.; König, W. A. *Tetrahedron: Asymmetr.* **2004**, *15*, 1631–1635.
- Faraldos, J. A.; Wu, S.; Chappell, J.; Coates, R. M. *Tetrahedron* **2007**, *63*, 7733–7742.
- Rinkel, J.; Dickschat, J. S. *Org. Lett.* **2019**, *21*, 2426–2429.
- Kodama, M.; Yokoo, S.; Matsuki, Y.; Ito, S. *Tetrahedron Lett.* **1979**, *20*, 1687–1690.
- Sutherland, J. K. *Tetrahedron* **1974**, *30*, 1651–1660.
- Minnaard, A. J.; Wijnberg, J. B. P. A.; de Groot, A. *Tetrahedron* **1994**, *50*, 4755–4764.
- Kodama, M.; Matsuki, Y.; Ito, S. *Tetrahedron Lett.* **1976**, *17*, 1121–1124.
- Kodama, M.; Yokoo, S.; Yamada, H.; Ito, S. *Tetrahedron Lett.* **1978**, *19*, 3121–3124.
- Stahl, E. *Planta Med.* **1984**, *50*, 157–160.
- Irmisch, S.; Jiang, Y.; Chen, F.; Gershenzon, J.; Köllner, T. G. *BMC Plant Biol.* **2014**, *14*, 270.
- Hattan, J.; Shindo, K.; Ito, T.; Shibuya, Y.; Watanabe, A.; Tagaki, C.; Ohno, F.; Sasaki, T.; Ishii, J.; Kondo, A.; Misawa, N. *Planta* **2016**, *243*, 959–972.
- Chuang, L.; Wen, C.-S.; Lee, Y.-R.; Lin, Y.-L.; Hsu, L.-R.; Wang, S.-Y.; Chu, F.-H. *J. Nat. Prod.* **2018**, *81*, 1162–1172.
- Baer, P.; Rabe, P.; Citron, C. A.; de Oliveira Mann, C. C.; Kaufmann, N.; Groll, M.; Dickschat, J. S. *ChemBioChem.* **2014**, *15*, 213–216.
- Lackus, N. D.; Morawetz, J.; Xu, H.; Gershenzon, J.; Dickschat, J. S.; Köllner, T. G. *Molecules* **2021**, *26*, 555.
- Xu, H.; Goldfuss, B.; Dickschat, J. S. *Chem.—Eur. J.* **2021**, *27*, 9758–9762.
- Rabe, P.; Schmitz, T.; Dickschat, J. S. *Beilstein J. Org. Chem.* **2016**, *12*, 1839–1850.
- Ding, L.; Goerls, H.; Dornblut, K.; Lin, W.; Maier, A.; Fiebig, H.-H.; Hertweck, C. *J. Nat. Prod.* **2015**, *78*, 2963–2967.
- Rising, K. A.; Starks, C. M.; Noel, J. P.; Chappell, J. *J. Am. Chem. Soc.* **2000**, *122*, 1861–1866.
- Little, D. B.; Croteau, R. B. *Arch. Biochem. Biophys.* **2002**, *402*, 120–135.
- Liang, J.; Liu, J.; Brown, R.; Jia, M.; Zhou, K.; Peters, R. J.; Wang, Q. *Plant J.* **2018**, *94*, 847–856.
- Wang, L.; Liang, J.; Xie, X.; Liu, J.; Shen, Q.; Li, L.; Wang, Q. *Plant Mol. Biol.* **2021**, *105*, 55–64.
- Mafu, S.; Karunanithi, P. S.; Palazzo, T. A.; Harrod, B. L.; Rodriguez, S. M.; Mollhoff, I. N.; O'Brien, T. E.; Tong, S.; Fiehn, O.; Tantillo, D. J.; Bohlmann, J.; Zerbe, P. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, 974–979.
- Starks, C. M.; Back, K.; Chappell, J.; Noel, J. P. *Science* **1997**, *277*, 1815–1820.
- Köllner, T. G.; Schnee, C.; Gershenzon, J.; Degenhardt, J. *Plant Cell* **2004**, *16*, 1115–1131.
- Wilderman, P. R.; Peters, R. J. *J. Am. Chem. Soc.* **2007**, *129*, 15736–15737.
- Liang, J.; Wang, L.; Liu, J.; Shen, Q.; Fu, J.; Peters, R. J.; Wang, Q. *Biochemistry* **2020**, *59*, 2660–2666.
- Xu, J.; Ai, Y.; Wang, J.; Xu, J.; Zhang, Y.; Yang, D. *Phytochemistry* **2017**, *137*, 34–41.
- Halsall, T. G.; Theobald, D. W.; Walshaw, K. B. *J. Chem. Soc.* **1964**, 1029–1037.

- (39) Rabe, P.; Barra, L.; Rinkel, J.; Riclea, R.; Citron, C. A.; Klapschinski, T. A.; Janusko, A.; Dickschat, J. S. *Angew. Chem., Int. Ed.* **2015**, *54*, 13448–13451.
- (40) Lauterbach, L.; Rinkel, J.; Dickschat, J. S. *Angew. Chem., Int. Ed.* **2018**, *57*, 8280–8283.
- (41) Rabe, P.; Rinkel, J.; Nubbemeyer, B.; Köllner, T. G.; Chen, F.; Dickschat, J. S. *Angew. Chem., Int. Ed.* **2016**, *55*, 15420–15423.
- (42) Hahn, F. M.; Hurlburt, A. P.; Poulter, C. D. *J. Bacteriol.* **1999**, *181*, 4499–4504.
- (43) Terada, Y.; Yamamura, S. *Tetrahedron Lett.* **1979**, *20*, 3303–3306.

## Recommended by ACS

### Quantitative High-Field NMR- and Mass Spectrometry-Based Fatty Acid Sequencing Reveals Internal Structure in Ru-Catalyzed Deuteration of Docosahexaenoic Acid

Dong Hao Wang, Mikhail S. Shchepinov, *et al.*

SEPTEMBER 13, 2022  
ANALYTICAL CHEMISTRY

READ 

### Structure of the Repurposed Fungal Terpene Cyclase FlvF Implicated in the C–N Bond-Forming Reaction of Flavunoidine Biosynthesis

Margarita A. Tararina, David W. Christianson, *et al.*

AUGUST 29, 2022  
BIOCHEMISTRY

READ 

### Insights into the Metabolomic Capacity of *Podaxis* and Isolation of Podaxisterols A–D, Ergosterol Derivatives Carrying Nitrosyl Cyanide-Derived Modifications

Huijuan Guo, Christine Beemelmanns, *et al.*

AUGUST 30, 2022  
JOURNAL OF NATURAL PRODUCTS

READ 

### Ansafurantrienins, Unprecedented Ansatrienin Derivatives Formed via Photocatalytic Intramolecular [3 + 2] Oxidative Cycloaddition

Hongji Li, Peng Sun, *et al.*

JANUARY 04, 2022  
ORGANIC LETTERS

READ 

Get More Suggestions >

## Appendix E

**Mechanistic Divergence between (4*S*,7*R*)-Germacra-(1(10)*E*,5*E*)-dien-11-ol Synthases from *Dictyostelium purpureum* and *Streptomyces coelicolor***

*Org. Biomol. Chem.* **2021**, *19*, 370

DOI: 10.1039/D0OB02361B



Cite this: DOI: 10.1039/d0ob02361b

Received 26th November 2020,  
Accepted 15th December 2020

DOI: 10.1039/d0ob02361b

rsc.li/obc

## Mechanistic divergence between (4*S*,7*R*)-germacra-(1(10)*E*,5*E*)-dien-11-ol synthases from *Dictyostelium purpureum* and *Streptomyces coelicolor*†

Houchao Xu,<sup>a</sup> Jan Rinkel,<sup>a</sup> Xinlu Chen,<sup>b</sup> Tobias G. Köllner,<sup>c</sup> Feng Chen<sup>b</sup> and Jeroen S. Dickschat<sup>\*,a</sup>

The main product of DpTPS9 from the social amoeba *Dictyostelium purpureum* was identified as (4*S*,7*R*)-germacra-(1(10)*E*,5*E*)-dien-11-ol that is also known as an intermediate of bacterial geosmin synthase, but the experimentally verified cyclisation mechanisms differ. Together with the low sequence identity this points to convergent evolution. The functionality of selected residues in DpTPS9 was investigated via site-directed mutagenesis experiments.

Terpene synthases are remarkable biocatalysts that convert structurally simple, acyclic and achiral oligoprenyl diphosphates into usually (poly)cyclic and chiral terpene hydrocarbons or alcohols. For the classical type I enzymes the substrates vary only in the chain length ranging from dimethylallyl diphosphate (DMAPP, C<sub>5</sub>, hemiterpenes) through geranyl-(GPP, C<sub>10</sub>, monoterpenes), farnesyl (FPP, C<sub>15</sub>, sesquiterpenes) and geranylgeranyl diphosphate (GGPP, C<sub>20</sub>, diterpenes) to geranylgeranyl diphosphate (GFPP, C<sub>25</sub>, sesterterpenes). Each of these substrates can be made from DMAPP by successive additions of isopentenyl diphosphate (IPP, C<sub>5</sub>) by a prenyltransferase, thus tracing back terpene biosynthesis to only two isomeric C<sub>5</sub> building blocks. Terpene biosynthesis proceeds then, in case of type I enzymes, by the abstraction of diphosphate, producing a reactive allyl cation that can cyclise through the attack of an olefinic double bond to the cationic centre. The number of possible cyclisation events depends on the number of double bonds in the substrate, while also hydride or proton shifts, Wagner-Meerwein rearrangements or sometimes ring-opening reactions may contribute to the cat-

ionic cascade that finally establishes the terpene skeleton. The cascade is terminated by deprotonation to yield a terpene hydrocarbon or by attack of water to result in a terpene alcohol. Such terpene cyclisation mechanisms can be studied by isotopic labelling experiments<sup>1</sup> or DFT calculations,<sup>2</sup> while the role of the enzyme can experimentally be addressed by site-directed mutagenesis,<sup>3–6</sup> with identification of functionally or structurally important residues based on a sequence alignment, a crystal structure or homology model, or theoretically by molecular mechanics simulations.<sup>7</sup> A large number of terpene synthases has been characterised from plants, bacteria, and fungi,<sup>8–13</sup> but knowledge about these enzymes from other organisms such as eukaryotic microorganisms is scarce. A few previous studies in our laboratories have resulted in the identification and functional characterisation of terpene synthases in the social amoebae *Dictyostelium discoideum* and *D. purpureum*.<sup>14–19</sup> Here we report on the functional characterisation of a sesquiterpene synthase from *D. purpureum* and enzyme mechanistic investigations by isotopic labelling experiments and site-directed mutagenesis.

The dictyostelid amoeba *D. purpureum* encodes twelve terpene synthases in its genome designated DpTPS1–12 that all exhibited sesquiterpene synthase activity *in vitro*.<sup>17</sup> Specifically, DpTPS9 produced a sesquiterpene alcohol **1** as the major product from FPP, besides smaller amounts of a sesquiterpene hydrocarbon **2** (Fig. S1†). For structure elucidation a preparative scale incubation of FPP with purified recombinant DpTPS9 was performed, followed by compound purification and extensive one- and two-dimensional NMR spectroscopy (Table 1, Fig. S2–S8†), resulting in the structure of (4*S*\*,7*R*\*)-germacra-(1(10)*E*,5*E*)-dien-11-ol for **1** (Scheme 1).

The absolute configurations of terpenes can be determined through a labelling strategy using enantioselectively deuterated oligoprenyl diphosphate precursors. Their enzymatic conversion by terpene synthases leads to stereoselectively deuterated products with known absolute configuration at the deuterated carbon. The absolute configuration of the investigated terpene can then be concluded from the relative orientation of the

<sup>a</sup>Kekulé-Institute of Organic Chemistry and Biochemistry, University of Bonn, Gerhard-Domagk-Strasse 1, 53121 Bonn, Germany. E-mail: dickschat@uni-bonn.de

<sup>b</sup>Department of Plant Sciences, University of Tennessee, 2431 Joe Johnson Drive, Knoxville, TN 37996-4561, USA

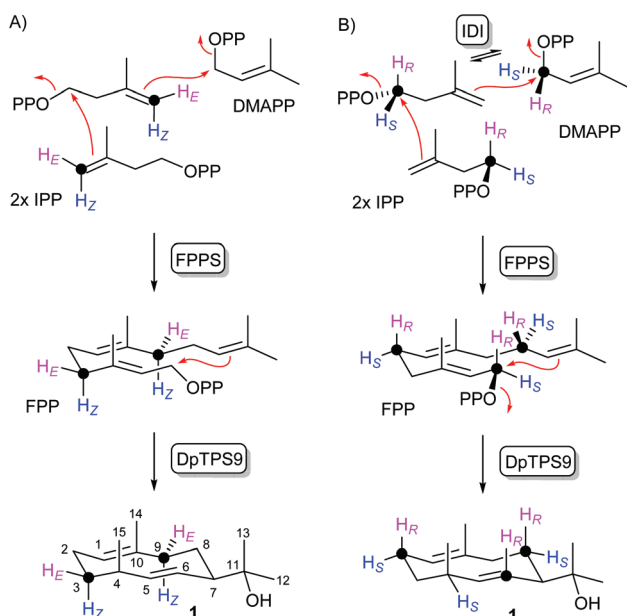
<sup>c</sup>Max Planck Institute for Chemical Ecology, Hans-Knöll-Straße 8, 07745 Jena, Germany

† Electronic supplementary information (ESI) available: Experimental procedures, EI mass spectra and NMR spectra of DpTPS9 enzyme products, and results from labelling experiments. See DOI: 10.1039/d0ob02361b

**Table 1** NMR data of (4*S*,7*R*)-germacra-(1(10)*E*,5*E*)-dien-11-ol (**1**)

C <sup>a</sup>	<sup>13</sup> C <sup>b</sup> (δ)	<sup>1</sup> H <sup>b</sup> (δ, m, <i>J</i> , int)	<sup>13</sup> C <sup>c</sup> (δ)
1	130.68 (CH)	4.98 (br d, <i>J</i> = 11.7)	130.66
2	24.15 (CH <sub>2</sub> )	2.34 (m, H <sub>β</sub> ) 1.85 (m, H <sub>α</sub> )	23.80
3	33.20 (CH <sub>2</sub> )	1.52 (m, H <sub>α</sub> ) 1.34 (m, H <sub>β</sub> )	32.87
4	34.19 (CH)	2.27 (m)	33.93
5	142.24 (CH)	5.44 (dd, <i>J</i> = 15.9, 3.5)	143.15
6	124.79 (CH)	4.92 (ddd, <i>J</i> = 15.9, 9.8, 2.0)	123.84
7	59.26 (CH)	2.12 (m)	58.99
8	21.99 (CH <sub>2</sub> )	1.43 (m, H <sub>α</sub> ) 1.12 (m, H <sub>β</sub> )	22.12
9	41.71 (CH <sub>2</sub> )	2.22 (m, H <sub>β</sub> ) 2.17 (m, H <sub>α</sub> )	41.37
10	131.16 (C)	—	131.18
11	71.44 (C)	—	71.81
12	26.80 (CH <sub>3</sub> )	1.12 (s)	26.37
13	27.64 (CH <sub>3</sub> )	1.05 (s)	26.91
14	16.83 (CH <sub>3</sub> )	1.46 (s)	16.77
15	14.90 (CH <sub>3</sub> )	1.00 (d, <i>J</i> = 6.9)	14.79

<sup>a</sup> Carbon numbering as in Scheme 1. <sup>b</sup> Recorded in (<sup>2</sup>H<sub>6</sub>)benzene at 700 MHz (<sup>1</sup>H) or 175 MHz (<sup>13</sup>C), chemical shifts δ in ppm, multiplicity m (s = singlet, d = doublet, m = multiplet, br = broad), coupling constants *J* are given in Hertz. <sup>c</sup> Recorded in (<sup>2</sup>H)chloroform at 75 MHz.<sup>24</sup>



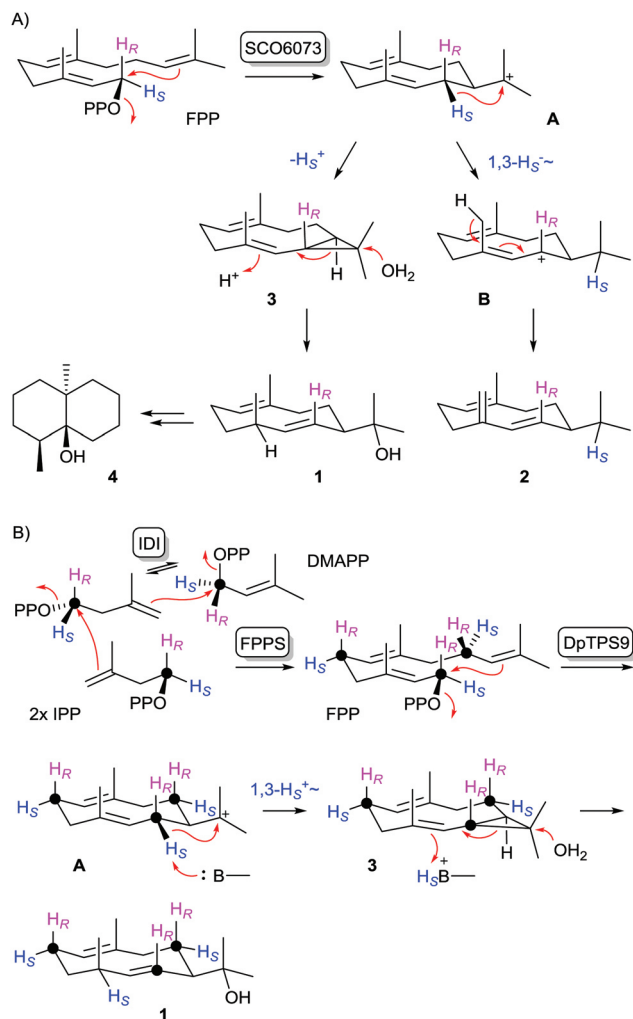
**Scheme 1** Determination of the absolute configuration of **1** using the stereoselectively deuterated substrates (A) (*E*- and (*Z*)-(<sup>4-<sup>13</sup>C,4-<sup>2</sup>H</sup>)IPP with DMAPP, FPPS and DpTPS9, and (B) (*R*- and (*S*)-(<sup>1-<sup>13</sup>C,1-<sup>2</sup>H</sup>)IPP with IDI, FPPS and DpTPS9. Black dots indicate <sup>13</sup>C-labelled carbons, coloured hydrogens were substituted with deuterium.

naturally present stereogenic centres to the additional stereogenic centre introduced by the deuteration. For this purpose, we have developed the stereoselectively deuterated probes (*E*- and (*Z*)-(<sup>4-<sup>13</sup>C,4-<sup>2</sup>H</sup>)IPP<sup>20</sup> that can be coupled with DMAPP using FPP synthase (FPPS) from *Streptomyces coelicolor*<sup>15</sup> to yield FPP with enantioselective deuterations at C4 and C8

(Scheme 1, Fig. S9†). Additional stereoselective labellings can be introduced at C1, C5 and C9 with (*R*- and (*S*)-(<sup>1-<sup>13</sup>C,1-<sup>2</sup>H</sup>)IPP<sup>21</sup> in conjunction with isopentenyl diphosphate isomerase (IDI) from *Escherichia coli*<sup>22</sup> and FPPS (Fig. S10†). After formation of all four FPP isotopomers through a known stereochemical course with inversion of configuration at C1 of the allyl diphosphates and attack of IPP from the *Si* face at C4,<sup>23</sup> their conversion with DpTPS9 gave stereoselectively labelled products that supported the absolute configuration of (4*S*,7*R*)-**1**, thus identifying this enzyme as a eukaryotic (4*S*,7*R*)-germacra-(1(10)*E*,5*E*)-dien-11-ol synthase. These data also confirmed the previously tentatively assigned absolute configuration of (4*S*,7*R*)-**1** from *Streptomyces citreus* by the same sign of the optical rotation (**1** from DpTPS9: [ $\alpha$ ]<sub>D</sub><sup>20</sup> = -297, *c* 0.03, CHCl<sub>3</sub>; **1** from *S. citreus*: [ $\alpha$ ]<sub>D</sub><sup>20</sup> = -82, *c* 0.23, CHCl<sub>3</sub>)<sup>24</sup> that was later secured by chemical degradation of (4*S*,7*R*)-**1** from the liverwort *Dumortiera hirsuta* ([ $\alpha$ ]<sub>D</sub> = -153.3, *c* 1.5, CHCl<sub>3</sub>)<sup>25</sup> The sesquiterpene hydrocarbon **2** was identified as (-)-germacrene D by comparison to both enantiomers present in the essential oil of *Solidago canadensis* through gas chromatography using a homochiral stationary phase (Fig. S11†).<sup>26</sup>

Compound **1** has been reported as a volatile emitted by several myxobacteria,<sup>27,28</sup> streptomycetes,<sup>29,30</sup> cyanobacteria,<sup>31</sup> and sponge-associated fungi.<sup>32</sup> It is also an intermediate in the biosynthesis of the earthy odorant geosmin (**4**).<sup>33</sup> As established by Cane and coworkers, **1** is produced from FPP by the N-terminal domain and further converted by the C-terminal domain of the bifunctional geosmin synthase SCO6073 from *Streptomyces coelicolor* into **4**, while **2** is known as a side product of this enzyme (Scheme 2A).<sup>34–36</sup> Isotopic labelling experiments with (1,1-<sup>2</sup>H<sub>2</sub>)FPP and both enantiomers of (*R*- and (*S*)-(<sup>1-<sup>2</sup>H</sup>)FPP revealed that after the cyclisation of FPP to the (*E,E*)-germacradienyl cation (**A**) the 1-*pro-S* hydrogen of FPP (H<sub>S</sub>, blue) shifts into the isopropyl group to form **2** via **B**. For the biosynthesis of **1** this hydrogen is lost by deprotonation of **A** to isolepidozene (**3**), followed by protonation induced ring opening with attack of water to **1**. This mechanism was also supported by the uptake of one deuterium during incubation in D<sub>2</sub>O.<sup>34,35</sup>

For DpTPS9, the incubations of (*R*- and (*S*)-(<sup>1-<sup>13</sup>C,1-<sup>2</sup>H</sup>)IPP with IDI, FPPS and DpTPS9 indicated some important mechanistic differences, as none of the deuterium atoms from either substrate was lost, *i.e.* besides the three <sup>13</sup>C atoms also three deuterium atoms were incorporated into **1** in both cases (Fig. S12†). The <sup>13</sup>C-NMR spectra of the obtained products (Fig. S13†) indicated that the 1-*pro-R* hydrogen of FPP (H<sub>R</sub>, purple) did not migrate, leading to an upfield shifted triplet for C6 of **1** as a result of <sup>13</sup>C-<sup>2</sup>H spin coupling ( $\Delta\delta$  = -0.33 ppm, <sup>1</sup>*J*<sub>C,D</sub> = 22.8 Hz), while the 1-*pro-S* hydrogen shifted away as indicated by a singlet for C6 that was slightly shifted downfield as a result of a deuterium isotope effect ( $\Delta\delta$  = +0.06 ppm). The enzymatic conversion of (3-<sup>13</sup>C,1,1-<sup>2</sup>H<sub>2</sub>)FPP<sup>15</sup> with DpTPS9 resulted in an upfield shifted triplet for C4 of **1** ( $\Delta\delta$  = -0.48 ppm, <sup>1</sup>*J*<sub>C,D</sub> = 19.6 Hz), demonstrating that the 1-*pro-S* hydrogen of FPP ends up at this carbon (Fig. S14†). Product analysis by GC/MS showed the same result for com-



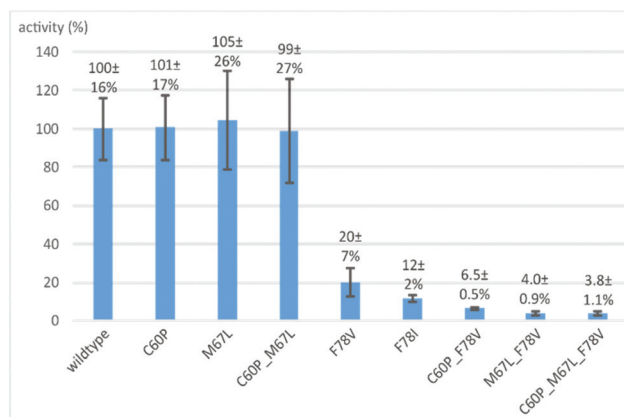
**Scheme 2** Divergent cyclisation mechanisms from FPP to 1. (A) Mechanism for the geosmin synthase SCO6073 from *S. coelicolor*, (B) mechanism for DpTPS9 from *D. purpureum*. Black dots indicate <sup>13</sup>C-labelled carbons, coloured hydrogens were substituted with deuterium.

pond 2 regarding the migration of the 1-*pro-S* hydrogen as observed for the geosmin synthase, which was evident from the base peak ion that arises by cleavage of the isopropyl group of 2 (Fig. S15†). Taken together, these experiments support a mechanism for DpTPS9 by cyclisation of FPP to A, a central intermediate in sesquiterpene biosynthesis,<sup>37</sup> and downstream steps to 2 as for the geosmin synthase (Scheme 2B). The further steps to 1 could also proceed through intermediate 3, formed from A by abstraction of the 1-*pro-S* proton e.g. by a basic active site residue, with strict reintroduction of the same proton at C4 of 3 in the subsequent reaction to 1. The different cyclisation mechanisms of DpTPS9 and bacterial geosmin synthase are also reflected by the low sequence identity of only 16% (comparison of DpTPS9 to the N-terminal domain, amino acid residues 1–345, of SCO6073), suggesting that these enzymes have independently evolved.

According to the amino acid sequence and the location of highly conserved motifs including the aspartate-rich motif

(<sup>81</sup>DDILD), the NSE triad (<sup>222</sup>NDMASYCKE), the RY pair (<sup>313</sup>RY) and the pyrophosphate sensor (<sup>175</sup>R), amoebal type I terpene synthases are similar to fungal and bacterial type I terpene synthases (bold residues are highly conserved, Fig. S16†).<sup>14,38</sup> Besides these motifs bacterial terpene synthases also exhibit conserved Pro and Leu residues 21 and 14 positions upstream of the Asp-rich motif for which a structural role has been assigned based on the crystal structure of selina-4(15),7(11)-diene synthase from *Streptomyces pristinaespiralis* (SdS)<sup>5</sup> and mutational studies with spiroalbatene synthase from *Allokutzneria albata* (SaS).<sup>6</sup> These positions are occupied by Cys60 and Met67 in DpTPS9, but the C60P, M67L and combined C60P-M67L enzyme variants exhibited a productivity that was very similar to the wildtype enzyme (set to 100 ± 16%) in all three cases (C60P: 101 ± 17%, M67L: 105 ± 26%, C60P-M67L: 99 ± 27%; Fig. 1), demonstrating that Cys60 and Met67 can efficiently substitute for the more regular Pro and Leu residues.

Bacterial enzymes also usually exhibit a Phe residue three positions upstream of the Asp-rich motif. The crystal structure of *epi*-isozizaene synthase from *S. coelicolor* showed that this residue is involved in the stabilisation of cationic intermediates through cation-π-interactions and its exchange (F96A) resulted in a loss of activity and formation of (*E*)-β-farnesene as the main product.<sup>39</sup> Similarly, the F97L, F97Y and F97W variants of selina-4(15),7(11)-diene synthase resulted in changed product profiles towards germacrene B and farnesenes.<sup>5</sup> In pentalenene synthase from *Streptomyces exfoliatus* (PS) for the F77Y variant a moderately reduced activity, but still production of pentalenene was observed.<sup>4</sup> Mutation of this residue in DpTPS9 showed a significantly reduced activity for both the F78V (20 ± 7%) and the F78I variant (12 ± 2%; Fig. 1), but no change in the product profile. Notably, the double mutations C60P-F78V (6.5 ± 0.5%), M67L-F78V (4.0 ± 0.9%) and the triple mutant C60P-M67L-F78V (3.8 ± 1.1%) showed a more strongly reduced activity, suggesting that the combined exchanges may have cooperative effects on the active site architecture of



**Fig. 1** Relative activities of wildtype DpTPS9 (set to 100%) and enzyme variants obtained by site-directed mutagenesis. The data are mean ± standard deviations obtained from triplicates.

DpTPS9. We have recently described a single highly conserved Asn residue eight or nine positions downstream of the NSE triad, whose mutation in the (-)- $\beta$ -araneosene synthase (DpTPS10, N249A, N249D) and in the (S)-(+)-nephthenol synthase from *D. purpureum* (DdTPS11, N243A, N243D) gave insoluble proteins, suggesting a structural function of this residue.<sup>38</sup> Similarly, for the (-)- $\beta$ -barbatene synthase from *D. discoideum* (DdTPS9) the N236D enzyme variant was insoluble, while only the N236A was obtained as a soluble protein, but exhibited strongly reduced catalytic activity.<sup>38</sup> Complementary alterations for DpTPS9 (N239A, N239D) also gave insoluble enzymes in both cases, confirming these earlier findings.

## Conclusions

In summary, we have characterised the sesquiterpene synthase DpTPS9 from *D. purpureum* as (-)-(4S,7R)-germacra-(1(10)E,5E)-dien-11-ol synthase and investigated its enzyme mechanism by isotopic labelling experiments and site-directed mutagenesis. The same product is known from the N-terminal domain of the bacterial geosmin synthase, but the amino acid sequences of the amoebal and the bacterial enzyme are different, suggesting that the same enzyme functions have evolved independently. Interestingly, while the products of DpTPS9 and the N-terminal domain of geosmin synthase are the same, their mechanisms of formation are different, demonstrating that for conclusive insights a mechanistic investigation of distantly related terpene synthases with the same product is required. The site-directed mutagenesis experiments reported here confirmed the importance of the Phe residue located three positions upstream of the Asp-rich motif and of the Asn residue located nine positions downstream of the NSE triad. Both residues are present in most bacterial, fungal and amoebal terpene synthases, and microbial type terpene synthases known as MTPSLs from plants,<sup>40</sup> but are not observed in typical plant terpene synthases.<sup>41</sup> Another highly conserved residue found in almost all characterised bacterial and fungal type terpene synthases is a Pro usually accompanied by a Leu residue located 21 and 14 residues upstream of the Asp-rich motif. Exemptions for which the conserved Pro is naturally substituted by another residue are the bacterial synthases for *epi*-isozizaene,<sup>42</sup> 1,8-cineol,<sup>43</sup> cembrene A<sup>44</sup> and hydroppyrene,<sup>45</sup> and the fungal synthases for trichodiene<sup>46</sup> and acoradiene.<sup>47</sup> As shown here, this Pro residue is not important for catalysis by DpTPS9 in which the corresponding position is occupied by Cys, but its exchange to Pro still yields a functional enzyme of similar catalytic activity. So far only a few terpene synthases from social amoebae have been characterised,<sup>14-19</sup> but more genetic information may soon become available and will allow us to continue our investigations on these enzymes in future experiments.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

This work was funded by the DFG (DI1536/7-2). We thank Andreas Schneider for the HPLC purification of 1.

## References

- 1 J. S. Dickschat, *Angew. Chem., Int. Ed.*, 2019, **58**, 15964.
- 2 D. J. Tantillo, *Nat. Prod. Rep.*, 2011, **28**, 1035; D. J. Tantillo, *Angew. Chem., Int. Ed.*, 2017, **56**, 10040.
- 3 M. Seemann, G. Zhai, K. Umezawa and D. Cane, *J. Am. Chem. Soc.*, 1999, **121**, 591.
- 4 M. Seemann, G. Z. Zhai, J. W. de Kraker, C. M. Paschall, D. W. Christianson and D. E. Cane, *J. Am. Chem. Soc.*, 2002, **124**, 7681.
- 5 P. Baer, P. Rabe, K. Fischer, C. A. Citron, T. A. Klapschinski, M. Groll and J. S. Dickschat, *Angew. Chem., Int. Ed.*, 2014, **53**, 7652.
- 6 J. Rinkel, L. Lauterbach, P. Rabe and J. S. Dickschat, *Angew. Chem., Int. Ed.*, 2018, **57**, 3238.
- 7 D. T. Major, *ACS Catal.*, 2017, **7**, 5461.
- 8 J. Degenhardt, T. G. Köllner and J. Gershenzon, *Phytochemistry*, 2009, **70**, 1621.
- 9 M. B. Quin, C. M. Flynn and C. Schmidt-Dannert, *Nat. Prod. Rep.*, 2014, **31**, 1449.
- 10 J. S. Dickschat, *Nat. Prod. Rep.*, 2016, **33**, 87.
- 11 D. W. Christianson, *Chem. Rev.*, 2017, **117**, 11570.
- 12 A. Minami, T. Ozaki, C. Liu and H. Oikawa, *Nat. Prod. Rep.*, 2018, **35**, 1330.
- 13 T. Mitsunashi and I. Abe, *ChemBioChem*, 2018, **19**, 1106.
- 14 X. Chen, T. G. Köllner, Q. Jia, A. Norris, B. Santhanam, P. Rabe, J. S. Dickschat, G. Shaulsky, J. Gershenzon and F. Chen, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 12132.
- 15 P. Rabe, J. Rinkel, B. Nubbemeyer, T. G. Köllner, F. Chen and J. S. Dickschat, *Angew. Chem., Int. Ed.*, 2016, **55**, 15420.
- 16 J. Rinkel, P. Rabe, X. Chen, T. G. Köllner, F. Chen and J. S. Dickschat, *Chem. – Eur. J.*, 2017, **23**, 10501.
- 17 X. Chen, T. G. Köllner, G. Shaulsky, Q. Jia, J. S. Dickschat, J. Gershenzon and F. Chen, *Sci. Rep.*, 2018, **8**, 14361.
- 18 X. Chen, K. Luck, P. Rabe, C. Q. D. Dinh, G. Shaulsky, D. R. Nelson, J. Gershenzon, J. S. Dickschat, T. G. Köllner and F. Chen, *eLife*, 2019, **8**, e44352.
- 19 J. Rinkel, T. G. Köllner, F. Chen and J. S. Dickschat, *ChemComm*, 2019, **55**, 13255.
- 20 L. Lauterbach, J. Rinkel and J. S. Dickschat, *Angew. Chem., Int. Ed.*, 2018, **57**, 8280.
- 21 J. Rinkel and J. S. Dickschat, *Org. Lett.*, 2019, **21**, 2426.
- 22 F. M. Hahn, A. P. Hurlburt and C. D. Poulter, *J. Bacteriol.*, 1999, **181**, 4499.
- 23 J. W. Cornforth, R. H. Cornforth, G. Popják and L. Yengoyan, *J. Biol. Chem.*, 1966, **241**, 3970.
- 24 D. Ganßer, F. C. Pollak and R. G. Berger, *J. Nat. Prod.*, 1995, **58**, 1790.
- 25 M. Toyota, T. Yoshida, J. Matsunami and Y. Asakawa, *Phytochemistry*, 1997, **44**, 293.

- 26 C. O. Schmidt, H. J. Bouwmeester, J. de Kraker and W. A. König, *Angew. Chem., Int. Ed.*, 1998, **37**, 1400.
- 27 S. Schulz, J. Fuhlendorff and H. Reichenbach, *Tetrahedron*, 2004, **60**, 3863.
- 28 J. S. Dickschat, S. C. Wenzel, H. B. Bode, R. Müller and S. Schulz, *ChemBioChem*, 2004, **5**, 778.
- 29 J. S. Dickschat, T. Martens, T. Brinkhoff, M. Simon and S. Schulz, *Chem. Biodivers.*, 2005, **2**, 837.
- 30 P. Rabe, C. A. Citron and J. S. Dickschat, *ChemBioChem*, 2013, **14**, 2345.
- 31 C. Höckelmann, P. G. Becher, S. H. von Reuss and F. Jüttner, *Z. Naturforsch.*, 2009, **64c**, 49.
- 32 L. Barra, P. Barac, G. M. König, M. Crüsemann and J. S. Dickschat, *Org. Biomol. Chem.*, 2017, **15**, 7411.
- 33 J. S. Dickschat, H. B. Bode, T. Mahmud, R. Müller and S. Schulz, *J. Org. Chem.*, 2005, **70**, 5174.
- 34 X. He and D. E. Cane, *J. Am. Chem. Soc.*, 2004, **126**, 2678.
- 35 J. Jiang, X. He and D. E. Cane, *J. Am. Chem. Soc.*, 2006, **128**, 8128.
- 36 J. Jiang, X. He and D. E. Cane, *Nat. Chem. Biol.*, 2007, **3**, 711.
- 37 H. Xu and J. S. Dickschat, *Chem. – Eur. J.*, 2020, **26**, DOI: 10.1002/chem.202002163.
- 38 J. Rinkel, T. G. Köllner, F. Chen and J. S. Dickschat, *ChemComm*, 2019, **55**, 13255.
- 39 J. A. Aaron, X. Lin, D. E. Cane and D. W. Christianson, *Biochemistry*, 2010, **49**, 1787.
- 40 Q. Jia, G. Li, T. G. Köllner, J. Fu, X. Chen, W. Xiong, B. J. Crandall-Stotler, J. L. Bowman, D. J. Weston, Y. Zhang, L. Chen, Y. Xie, F.-W. Li, C. J. Rothfels, A. Larsson, S. W. Graham, D. W. Stevenson, G. K.-S. Wong, J. Gershenson and F. Chen, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 12328.
- 41 F. Chen, D. Tholl, J. Bohlmann and E. Pichersky, *Plant J.*, 2011, **66**, 212.
- 42 X. Lin and D. E. Cane, *J. Am. Chem. Soc.*, 2009, **131**, 6332.
- 43 C. Nakano, H.-K. Kim and Y. Ohnishi, *ChemBioChem*, 2011, **12**, 1988.
- 44 J. Rinkel, L. Lauterbach, P. Rabe and J. S. Dickschat, *Angew. Chem., Int. Ed.*, 2018, **57**, 3238.
- 45 Y. Yamada, T. Kuzuyama, M. Komatsu, K. Shin-ya, S. Omura, D. E. Cane and H. Ikeda, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 857.
- 46 T. M. Hohn and P. D. Beremand, *Gene*, 1989, **79**, 131.
- 47 D. A. Yee, T. B. Kakule, W. Cheng, M. Chen, C. T. Y. Chong, Y. Hai, L. F. Hang, Y.-S. Hung, N. Liu, M. Ohashi, I. C. Okorafor, Y. Song, M. Tang, Z. Zhang and Yi Tang, *J. Am. Chem. Soc.*, 2020, **142**, 710.



## Appendix F







### Catalytic Role of Carbonyl Oxygens and Water in Selinadiene Synthase

*Nat. Catal.* **2022**, 5, 128

DOI: 10.1038/s41929-022-00735-0



# Catalytic role of carbonyl oxygens and water in selinadiene synthase

Yong-Heng Wang<sup>1,2,4</sup>, Houchao Xu<sup>3,4</sup>, Jian Zou<sup>1,4</sup>, Xian-Bo Chen<sup>1</sup>, Yu-Qing Zhuang<sup>1</sup>, Wei-Liang Liu<sup>1</sup>, Ersin Celik<sup>3</sup>, Guo-Dong Chen<sup>1</sup>, Dan Hu<sup>1</sup>, Hao Gao<sup>1</sup>, Ruibo Wu<sup>2</sup>  , Ping-Hua Sun<sup>1</sup>   and Jeroen S. Dickschat<sup>3</sup>  

**Terpene synthases (TSs) catalyse the most complex cyclization cascades in nature, with generation and taming of reactive carbocations. Although deprotonation–reprotonation sequences are frequently relevant for TS catalysis, little is known how the enzyme acts in these processes. Here we show, through quantum mechanics (density functional theory)/molecular mechanics molecular dynamics simulations that the main-chain carbonyl oxygen of Gly182 of selina-4(15),7(11)-diene synthase (SdS) has a dual role as a base and an acid and acts in synchrony with one water molecule. The computational model is supported by isotopic labelling experiments confirming the predicted stereochemical course associated with the deprotonation–reprotonation sequence. Gly182 is located within the G1/2 helix break of SdS, with all backbone carbonyl oxygens pointing into the active site having functions in recognizing substrate conformation, stabilizing carbocation intermediates and anchoring their poses. The strict conservation of the G1/2 helix break in type I TSs from bacteria, fungi and plants suggests that its functions as described here may be of general importance in TS catalysis.**


By first generating and then taming reactive carbocationic intermediates from a limited number of oligoprenyl diphosphates, terpene synthases (TSs) catalyse the most complex and diverse cyclization cascades found in nature. These transformations change on average more than half of the substrate's carbon atoms in terms of bonding situation and hybridization, with the introduction of stereochemical information into the achiral precursors in just a single enzymatic step. These remarkable biosynthetic processes provide diverse carbon skeletons that usually contain multiple stereogenic centres and can be polycyclic, giving access to the largest class of natural products, the terpenoids, with well over 80,000 known compounds<sup>1–4</sup>. The observed structural diversity of terpenoids is rooted in the pluripotent reactivity of the carbocationic intermediates, which can react through ring closures by intramolecular attack of an olefin to a cationic centre, hydride shifts, Wagner–Meerwein rearrangements and proton transfers, leading from one cationic intermediate to another and, by terminal deprotonation, to a neutral product that is ultimately released<sup>5–7</sup>. One key role for the TS is to steer this process by guiding the substrate through reactive conformations and buffering its reactivity—for example, by cation- $\pi$ -stabilizations<sup>8</sup> which often leads to the desired terpene product with astonishing selectivity. Herein TS catalysis appears to follow the principle of least effort: the most pronounced interventions of a TS lie in the initial carbocation formation, either by abstraction of diphosphate (type I enzymes) or by protonation of the substrate (type II), and the final cation quench, either by deprotonation or hydration, to yield the product. The intermediate events are believed to result from the inherent reactivity of the cationic intermediates, with their conformations being controlled by the hydrophobic active-site architecture<sup>9</sup>.

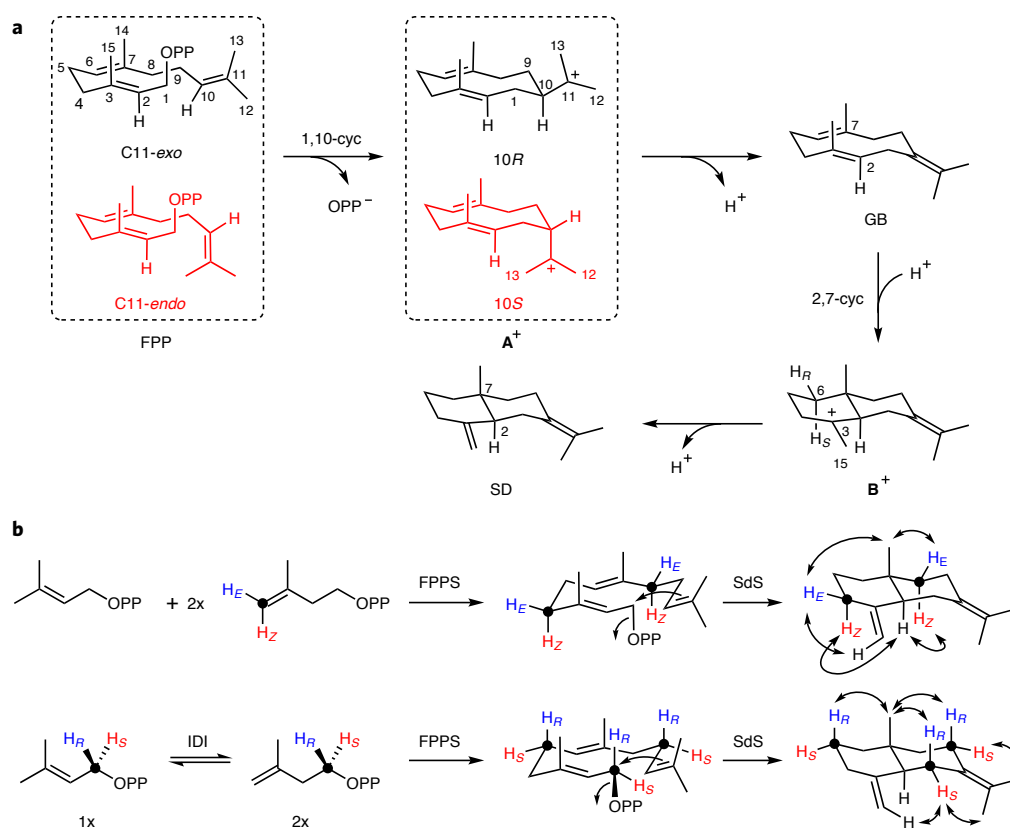
In many cases, deprotonation–reprotonation sequences through neutral intermediates<sup>10–17</sup> are proposed as an alternative to direct

intramolecular proton transfers<sup>18</sup> that often cannot be realized due to conformational and steric constraints. In contrast to intramolecular proton transfers, a deprotonation–reprotonation mechanism needs the direct intervention of the enzyme by providing a base for the deprotonation and an acid for the reprotonation. Little is known regarding how TSs intervene in such processes. Are two protein residues needed, one for deprotonation and another for reprotonation? Or is catalysis solved more economically with involvement of only one residue, which first abstracts the proton from one position of the cationic intermediate and then returns it to another position? The involvement of specific residues has been speculated on in previous work—for example, His309 was suggested to be relevant for deprotonation to  $\alpha$ -humulene and its reprotonation to initiate downstream cyclization to pentalene by bacterial pentalene synthase, but this hypothesis could not be confirmed by site-directed mutagenesis<sup>19</sup>.

Approaches to following deprotonation–reprotonation sequences are particularly difficult: labelling experiments with isotopic substitutions in the substrate can give insights into the fate of single atoms, but cannot divulge the exact sequence of events between the start and end of a reaction, so this approach deals with a black box problem and results from such experiments need to be interpreted with much care<sup>3</sup>. Site-directed mutagenesis can show the importance of active-site residues for catalytic activity<sup>19,20</sup>, but this approach alone does not give a detailed picture of how a residue identified as being critical actually acts. Together with an X-ray structure<sup>21–23</sup>, ideally of the wild type and the enzyme variant, or at least together with a structure homology model<sup>24,25</sup>, deeper insights may be obtained but enzyme structural data are usually obtained only with a co-crystallized substrate analogue and not with the relevant cationic intermediate<sup>1,2</sup>. Quantum chemical calculations have been extensively applied in investigations on the inherent reactivity of cationic

<sup>1</sup>College of Pharmacy, Jinan University, Guangzhou, P. R. China. <sup>2</sup>School of Pharmaceutical Sciences, Sun Yat-sen University, Guangzhou, P. R. China.

<sup>3</sup>Kekulé-Institute of Organic Chemistry and Biochemistry, University of Bonn, Bonn, Germany. <sup>4</sup>These authors contributed equally: Yong-Heng Wang, Houchao Xu, Jian Zou.  e-mail: [wurb3@mail.sysu.edu.cn](mailto:wurb3@mail.sysu.edu.cn); [pinghuasunny@163.com](mailto:pinghuasunny@163.com); [dickschat@uni-bonn.de](mailto:dickschat@uni-bonn.de)



**Fig. 1 | Cyclization mechanism of SdS. a**, Potential mechanisms for initiation of SdS catalysis for two different FPP conformations through 1,10-cyclization (1,10-cyc) and deprotonation to the neutral intermediate GB, followed by reprotonation-induced 2,7-cyclization (2,7-cyc) to SD. Carbon numbers follow FPP numbering (Supplementary Fig. 2 shows the difference between FPP and SD numbering). **b**, Stereoselective deuteration experiments ( $H_R$  or  $H_S$ ,  $^2H$ ; black dots,  $^{13}C$  labels) for the assignment of the absolute configuration of SD using labelled terpene monomers in conjunction with IDI, FPP synthase FPPS and SdS. Double-headed arrows indicate key NOESY correlations.  $H_Z$  and  $H_E$  are the hydrogens in the Z and E positions at C4 of IPP, respectively.

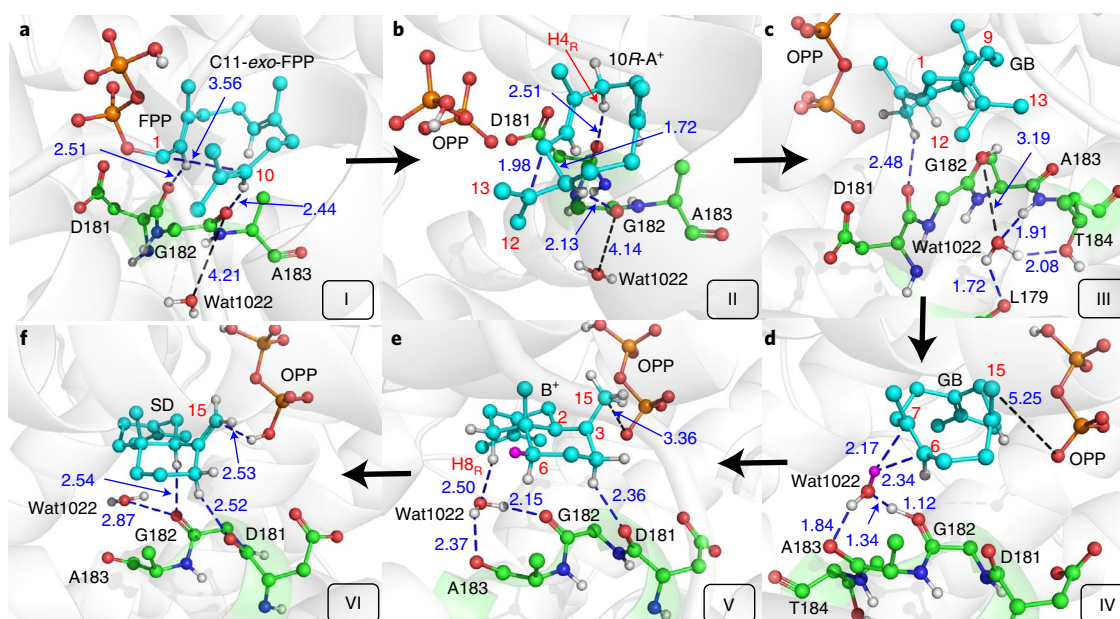
intermediates, but this approach is not capable of efficiently handling biomacromolecules and thus cannot take the enzyme environment into consideration<sup>4–8,26</sup>. In contrast, quantum mechanics and molecular mechanics (QM/MM), using QM to model the reactive part and MM to describe the enzyme environment, is ideally suited to studying the mechanisms of enzymatic reactions<sup>17,27–32</sup>.

Based on our previously reported crystal structure of selina-4(15),7(11)-diene synthase (SdS) from *Streptomyces pristinaespiralis* in complex with the non-reactive substrate surrogate 2,3-dihydrofarnesyl diphosphate (DHFPP)<sup>21</sup>, we now report on the deprotonation–reprotonation mechanism of SdS catalysis investigated by QM/MM molecular dynamics (MD) simulations in conjunction with isotopic labelling experiments. This work discloses many details of TS catalysis and shows how the enzyme recognizes the productive conformation of the substrate, how key carbocationic intermediates are stabilized and their poses are anchored and how the enzyme directly intervenes in proton transfers of the deprotonation–reprotonation cascade in selinadiene biosynthesis. All these functions are realized by the backbone carbonyl oxygens of the strictly conserved G1/2 helix motif of SdS and a nearby active-site water molecule.

## Results

**Hypothetical mechanisms for SdS catalysis.** Hypothetical mechanisms for SdS catalysis are shown in Fig. 1a. The substrate farnesyl diphosphate (FPP) can be cyclized starting from two different conformations, either with *exo*- or *endo*-orientation of the terminal alkene. Both conformations can be observed in the

homotetrameric crystal structure of SdS (PDB code: 4OKZ) in complex with four molecules of the substrate analogue DHFPP, in which two subunits accommodate DHFPP with C11-*exo*-conformation, while in one unit a C11-*endo*-conformation and in the fourth unit a C11-middle conformation is adopted (Supplementary Fig. 1; computed structures of free FPP and FPP in the enzyme are available as Supplementary Data 1–4<sup>21</sup>). Departure of the diphosphate group (OPP) initiates 1,10-cyclization to (10S)- or (10R)-A<sup>+</sup>; either enantiomer can explain selina-4(15),7(11)-diene (SD) biosynthesis, because the subsequent abstraction of hydrogen from C10 leads to the achiral minor product of SdS germacrene B (GB)<sup>21</sup>, with loss of the relevant stereochemical information for direct conclusions on the absolute configuration of A<sup>+</sup>. The observation of neutral GB favours a deprotonation–reprotonation mechanism towards SD that is also supported by DFT calculations<sup>33</sup>. In this mechanism the second C2–C7 cyclization needs triggering by reprotonation at C6 of GB to give B<sup>+</sup>, with incorporation of a proton from water into the *pro-R* ( $H_R$ ) or *pro-S* position ( $H_S$ ), and is followed by deprotonation from C15 to yield the final product SD. The C6 protonation in GB from the more readily accessible *re* side (rear view in the representation of Fig. 1a), leading to incorporation into the *pro-R* position, seems to be more likely than protonation from the *si* side (front view), requiring a stereochemically demanding C6–C7 *syn* addition with incorporation into the *pro-S* position. This second option would be possible if protonation and cyclization of GB proceed stepwise—that is, not concertedly to make the *syn* addition feasible. In a variation of this mechanistic model, the proton eliminated from C10 in A<sup>+</sup> could also be transferred directly to the



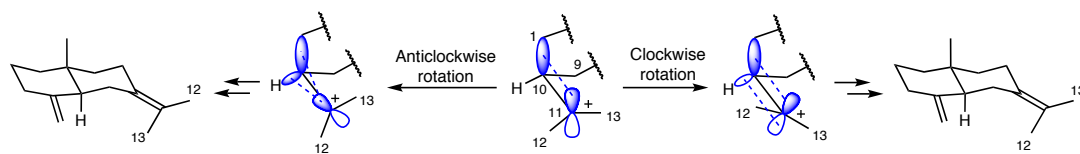
**Fig. 2 | Key structures for SdS catalysis obtained from QM(DFT)/MM simulations. a–f,** Illustrative structures from single MD trajectories with average distances from the final 10 ps, in Å. **a,** C11-*exo*-FPP at the active site of SdS with a hydrogen bridge from Asp181 to H2 and Gly182 to H10. **b,** 1,10-Cyclization results in the (*R*)-germacradienyl cation. Asp181 now has contact with H4<sub>R</sub>. **c,** Deprotonation by Gly182 as the catalytic base leads to neutral GB and protonated Gly with a hydrogen bond to an active-site water. **d,** The water is brought into proximity with C1 and C10 of GB (for further illustration of the movement of water from state II to IV, see Supplementary Fig. 12). **e,** This results in protonation at C1 of GB from the *re* face, with the water in contact with H8<sub>R</sub> and diphosphate in contact with H15. **f,** Diphosphate acts as the catalytic base in the final deprotonation to SD. All computed structures of intermediates and transition states are available as Supplementary Data 4–14.

*si* face of C6 in GB. However, previous DFT gas phase calculations disfavoured such a direct intramolecular proton transfer<sup>33</sup>, showing the requirement for enzyme guidance for such a hypothetical step.

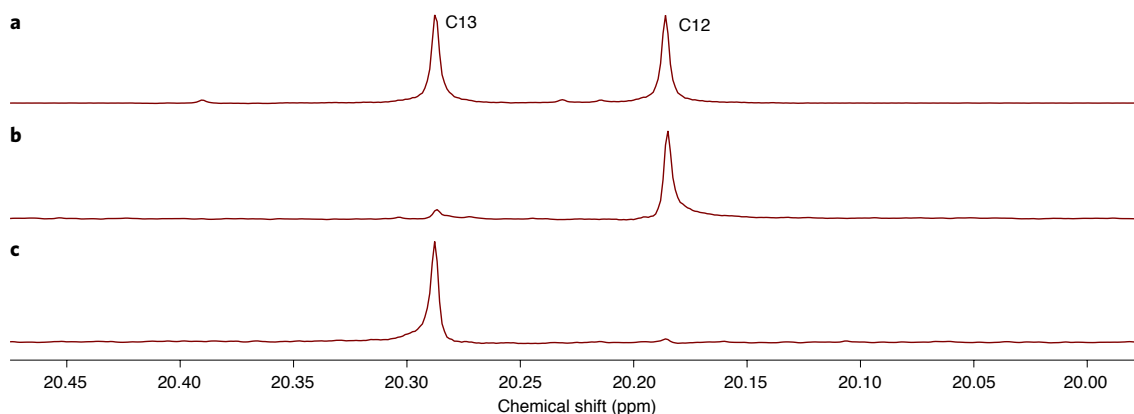
**Establishment of the absolute configuration of SD.** Knowledge of the absolute configuration of SD from SdS is critical for the computational work and interpretation of isotopic labelling experiments described in this Article, but this has not previously been established. For this purpose, a stereospecific deuteration was introduced from (*E*)- and (*Z*)-(4-<sup>13</sup>C,4-<sup>2</sup>H)isopentenyl pyrophosphate (IPP)<sup>25</sup> into SD by conversion with dimethylallyl diphosphate through farnesyl diphosphate synthase (FPPS) from *Streptomyces coelicolor*<sup>34</sup> and SdS, yielding stereoselectively labelled SD isotopomers (Fig. 1b). The additional <sup>13</sup>C labels allow for sensitive analysis by heteronuclear single quantum coherence (HSQC) spectroscopy without the need for compound isolation, allowing the labelling experiments to be conducted at small scale. Comparison of HSQC spectra from the labelling experiments with that of isolated SD, together with nuclear Overhauser effect spectroscopy (NOESY)-based hydrogen assignments, allows confirmation of the absolute configuration of (2*S*,7*R*)-SD (FPP numbering) from the known configurations at the deuterated carbons (Supplementary Fig. 3). Analogous experiments using (*R*)- and (*S*)-(1-<sup>13</sup>C,1-<sup>2</sup>H)-IPP<sup>35</sup> with conversion by isopentenyl diphosphate isomerase (IDI)<sup>36</sup>, FPPS and SdS confirmed the assigned absolute configuration of (2*S*,7*R*)-SD (Supplementary Fig. 4). Interpretation of the results from these stereoselective deuteration experiments relies on the known stereochemical course for the elongation of prenyl diphosphates with IPP by FPPS, with inversion of configuration at C1 of the elongated prenyl diphosphate and attack at C4 of IPP from the *si* face<sup>37</sup>.

**Formation of the neutral intermediate GB.** The relevant conformation of FPP, C11-*exo* or C11-*endo*, for the cyclization to SD was investigated by docking them to the active site of the SdS crystal

structure (Supplementary Fig. 5). During a 20 ps atomistic QM/MM MD simulation for FPP in the C11-*exo* conformation, a hydrogen bond interaction between H10 of FPP and the carbonyl oxygen of Gly182 with an average distance of 2.44 Å was observed (Fig. 2a; Supplementary Fig. 6 shows temporal evolution of this distance) that was absent in the equivalent simulation for the C11-*endo* conformation (Supplementary Fig. 7). In addition, in both conformations H2 forms a hydrogen bond with the carbonyl oxygen of Asp181 at an average distance of 2.51 Å and 2.52 Å, respectively (Fig. 2a and Supplementary Fig. 7). With the help of these hydrogen interactions, the average distance between C1 and C10 of FPP in the C11-*exo* conformation is only 3.56 Å, which fosters efficient 1,10-cyclization, while the average distance between these carbons in the C11-*endo* conformation is, at 4.13 Å, substantially longer, preventing an analogous cyclization reaction. This hypothesis is further supported by potential energy surface (PES) scans for diphosphate abstraction with 1,10-cyclization for both FPP conformers (Supplementary Fig. 8; Supplementary Fig. 9 shows the free energy profile of this step). With stabilization by the carbonyl group of Gly182, the C11-*exo* conformer can form the intermediate (10*R*)-A<sup>+</sup> (Supplementary Fig. 8, black curve) through an energy barrier of 21.7 kcal mol<sup>-1</sup>, which is comparable to the results from previous QM/MM calculations for other type I TSs<sup>17,30,38</sup>. Following diphosphate abstraction, C1 moves towards C10 causing a change in the hydrogen bond interaction of the carbonyl oxygen of Asp181 that switches from H2 to the 4-*pro-R* hydrogen (H4<sub>R</sub>), while the hydrogen bond between H10 and the carbonyl oxygen of Gly182 is maintained and stabilizes the cationic centre at C11 of (10*R*)-A<sup>+</sup> through electrostatic interaction. Removal of this electrostatic interaction by moving Gly182 from the QM to the MM region and setting its charge as zero in the QM/MM calculation resulted in an increased barrier by about 5.0 kcal mol<sup>-1</sup> for the 1,10-cyclization of C11-*exo*-FPP to (10*R*)-A<sup>+</sup> (Supplementary Fig. 8, blue curve), which is no longer stable and can persist only for <1 ps in the QM/MM MD simulation. Similarly,



**Fig. 3 | Molecular orbital interactions during C10–C11 bond rotation in (10R)-A<sup>+</sup>.** Blue dashed lines represent hyperconjugations, numbers at carbons follow FPP numbering. For clockwise rotation, the empty p-orbital interacts with both of its petals with the two  $\sigma$ -bonds and thus the two interactions are compatible. In contrast, for anticlockwise rotation the empty p-orbital uses only one of its petals in the interaction with both  $\sigma$ -bonds and thus these two interactions are competitive.



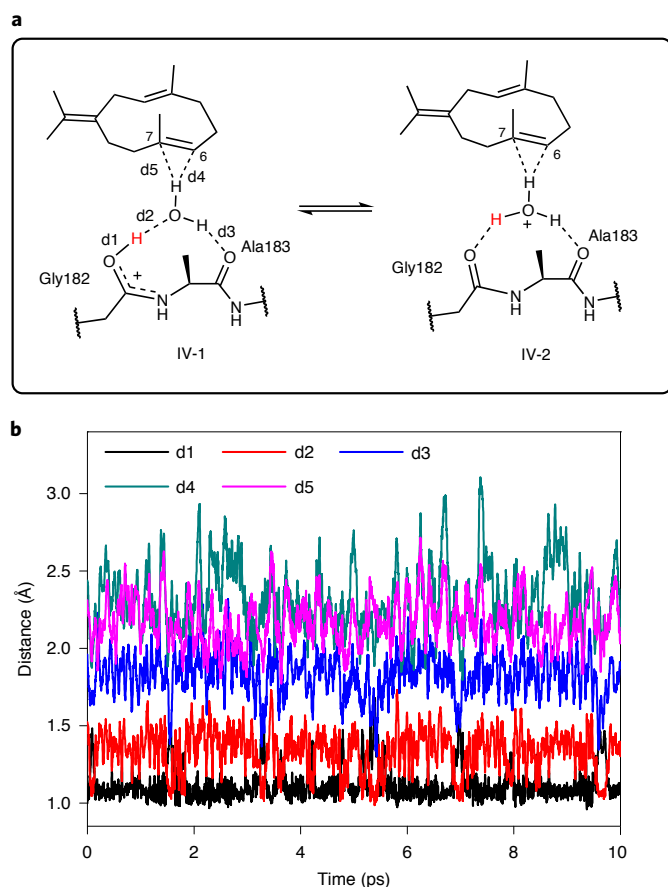
**Fig. 4 | Stereochemical fate of the geminal C12 and C13 methyl groups of FPP.** **a**, <sup>13</sup>C-NMR spectrum of SD (methyl region) showing signals for C12 and C13 (chemical shifts,  $\delta$  in ppm). **b**, <sup>13</sup>C-NMR spectrum of the extracted product obtained from the conversion of (12-<sup>13</sup>C)FPP with SdS. **c**, <sup>13</sup>C-NMR spectrum of the extracted product obtained from the conversion of (13-<sup>13</sup>C)FPP with SdS. The small peaks for C13 (**b**) and C12 (**c**) indicate very minor distribution of labelling by SdS. The signal for C13 (**b**) appears slightly larger than that for C12 (**c**), because the synthetic compound (12-<sup>13</sup>C)FPP has minor contamination with (13-<sup>13</sup>C)FPP.

(10S)-A<sup>+</sup> formed from C11-*endo*-FPP is, even with the involvement of Gly182, very unstable and reacts back to C11-*endo*-FPP in <2 ps (Supplementary Fig. 8, red curve). In summary, the carbonyl oxygen of Gly182 first recognizes C11-*exo*-FPP through hydrogen bonding and then has a critical function in stabilization of (10R)-A<sup>+</sup> through electrostatic interaction.

After cyclization of C11-*exo*-FPP (Fig. 2a), in the resulting (*E,E*)-germacradienyl cation (10R)-A<sup>+</sup> the atoms C9, C10, H10, C11, C12 and C13 remain approximately in one plane and the newly formed C1–C10 bond is perpendicular to this plane (Fig. 2b). Moreover, the average length of the C1–C10 bond (1.72 Å) is longer than that of a normal sp<sup>3</sup> C–C bond (~1.5 Å), while the average distance between C1 and C11 (1.98 Å) is much less than the sum of van der Waals radii of two carbon atoms (~3.4 Å). This indicates a strong hyperconjugation between the C1–C10  $\sigma$ -bond and the empty p-orbital at C11, which is also observed in the structure obtained from gas phase QM calculations<sup>33</sup> and which stabilizes the cationic centre located at C11 (Fig. 3, middle structure). The subsequent deprotonation from C10 requires that the C–H bond be perpendicular to the cationic plane so that the electron density located in the C–H  $\sigma$ -bonding orbital can fill the empty p-orbital of C11 with weakening of the C–H bond. Hence, either clockwise or anticlockwise rotation of C12–C11(+)-C13 around the C10–C11 bond is required. For both rotations, hyperconjugation between the empty C11 p-orbital and the C1–C10  $\sigma$ -orbital will decrease, with a simultaneous increase for the interaction of the C11 p-orbital with the C–H  $\sigma$ -orbital, yet their interaction modes are different (Fig. 3). For clockwise rotation, the empty p-orbital interacts with both of its petals with the two  $\sigma$ -bonds and thus the two interactions are compatible. In contrast,

for anticlockwise rotation the empty p-orbital uses only one of its petals in the interaction with both  $\sigma$ -bonds and these two interactions are thus competitive. Therefore, clockwise rotation may be more favourable in energy than anticlockwise, which is consistent with PES scans of the C11–C12 bond along the two rotational directions (Supplementary Fig. 10) with energy barriers for clockwise and anticlockwise rotation of 6.9 kcal mol<sup>-1</sup> and 9.6 kcal mol<sup>-1</sup>, respectively. The PES scans along the dihedral rotation cause a spontaneous deprotonation of C10 by Gly182, leading directly to GB (Supplementary Fig. 10). Therefore, the carbonyl oxygen of Gly182 serves as the base in the deprotonation of (10R)-A<sup>+</sup> (based on observations from single 20 ps QM/MM MD simulations without exploring alternatives, but the PES scans indicate that this mechanism is energetically feasible and no suitable conformations for alternative mechanisms were observed).

**Confirmation of the stereochemical course of GB formation.** The energetically preferred clockwise rotation of the C12–C11(+)-C13 group in the formation of GB must be associated with a specific stereochemical fate for C12 and C13, that should end up in the *Z* and *E* methyl groups, respectively, of the isopropylidene group of GB and consequently of SD, while the anticlockwise rotation predicted to be disfavoured based on QM/MM MD simulations should yield the opposite outcome (Fig. 3). These predictions can be tested experimentally using (12-<sup>13</sup>C)FPP<sup>39</sup> and (13-<sup>13</sup>C)FPP, enzymatically prepared from (9-<sup>13</sup>C)GPP<sup>15</sup> and IPP with FPPS, and their enzymatic conversion with SdS indeed gives a clear incorporation of <sup>13</sup>C-labelling from C12 into the *Z* methyl group and from C13 into the *E* methyl group, with only a very minor distribution of labelling between these positions (Fig. 4). Thus, the theoretically favoured



**Fig. 5 | Reprotonation of GB in the biosynthesis of SD. a**, Proton transfer between the carbonyl oxygen and the water molecule in state IV, with key distances d1, d2, d3, d4 and d5. **b**, Time evolution of key distances in state IV obtained by monitoring the final 10 ps QM/MM MD trajectory of state IV.

clockwise rotation of C12–C11(+)-C13 in the formation of GB is herewith confirmed experimentally.

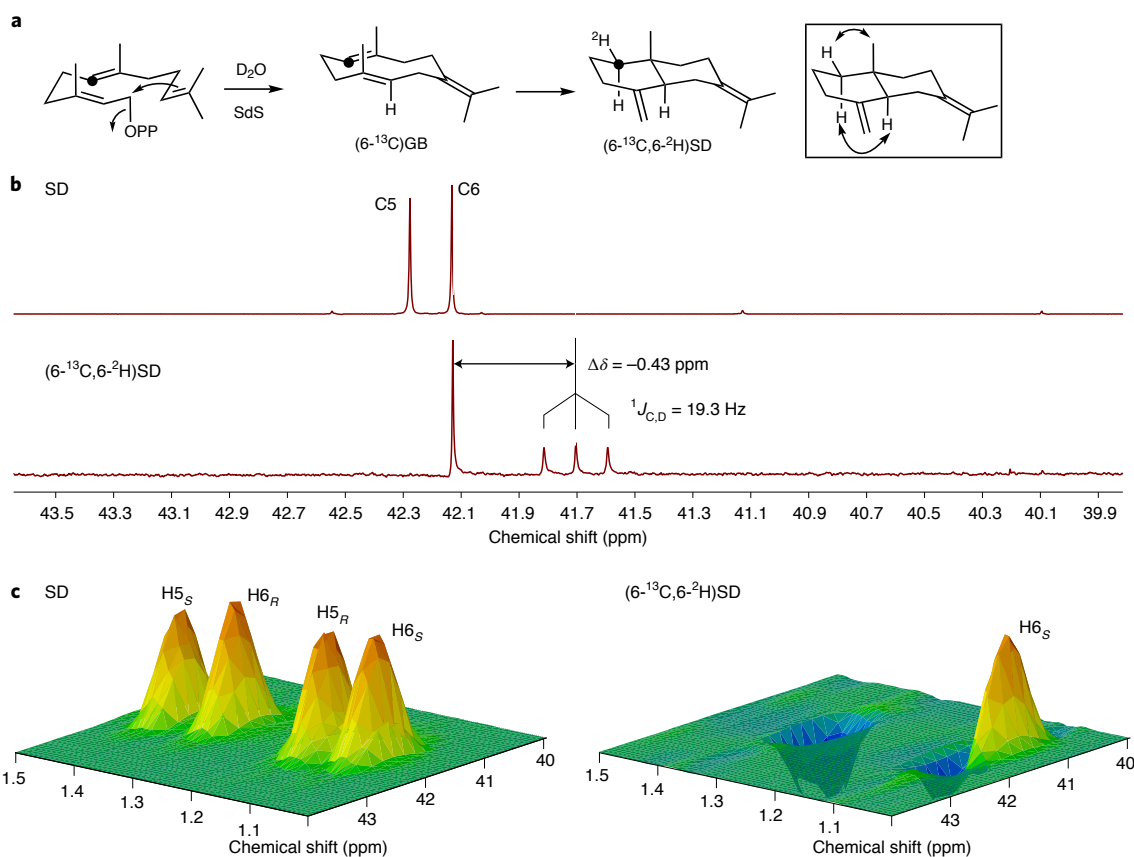
**Reactivation of the cyclization cascade by reprotonation.** After deprotonation of  $A^+$  to GB, the subsequent, second cyclization to SD requires reprotonation at C6, for which no suitable residue is observed in the SdS structure near this carbon. Therefore, the now protonated carbonyl of Gly182 was considered to be involved as a catalytic acid but, because of its fairly large distance to C6 and its unsuitable orientation, all attempts at direct proton transfer failed. Careful inspection of the QM/MM MD trajectories (single trajectories of the last 10 ps each) along the intermediate states of I, II and III (Fig. 2a–c) revealed the presence of a water molecule (Wat1022) that can be observed in all four subunits of the homotetrameric crystal structure of SdS. This water, at 5.82 Å, is too distant to H10 in state II to serve as the catalytic base in the deprotonation of  $A^+$  to GB. Notably, the average distance between the water oxygen and carbonyl oxygen of Gly182 decreases from 4.21 Å in state I to 4.14 Å in state II and then to 3.19 Å in state III, in which Gly182 is fully protonated (this distance in the crystal structure is, on average, 4.20 Å). In state III the water forms three hydrogen bonds with the NH and the side-chain oxygen of Thr184, and with the carbonyl group of Leu179, with the hydrogen bond to the side-chain oxygen of Thr184 being the least stable (Supplementary Fig. 11). After becoming fully protonated, the positively charged Gly182 may attract the water to the reactive site through electrostatic interaction

so that it could mediate proton transfer from Gly182 to C6 of GB<sup>40</sup>. This hypothesis is supported by a PES scan, leading to state IV (Fig. 2d; Supplementary Fig. 12 provides further illustration of the movement of water from state II to IV) in which water forms hydrogen bonds to the carbonyl oxygen of Ala183 and to protonated Gly182, with an additional O–H... $\pi$  interaction with the C6=C7 double bond of GB. The average distance between the water oxygen and the hydrogen of protonated Gly182 is only 1.34 Å, indicating a strong hydrogen bond<sup>41</sup> that may thus indeed be one of the main driving forces attracting water into the reactive site. The III-to-IV transition is exothermic (by 5.7 kcal mol<sup>-1</sup>), with a barrier of 13.7 kcal mol<sup>-1</sup> (Supplementary Fig. 13; alternatives were not explored, but no other water molecules suitable for migration to this site were observed in the single QM/MM MD simulation of state III).

Time-resolved monitoring of state IV (the single, final 10 ps QM/MM MD simulation) and, more specifically, of the key interactions between the active-site water, substrate and enzyme, revealed that the proton on Gly182 can be transferred to the water oxygen from time to time, as indicated by the evolution of distances d1 and d2 (Fig. 5b)—that is, state IV is best described as an equilibrium between states IV-1 and IV-2 (Fig. 5a). In the IV-2 state the positive charge of hydronium results in smaller distances, d3, d4 and d5 (note that the fluctuations of d3, d4 and d5 follow that of d2; Fig. 5b). The average value of d4 in IV is only 2.34 Å, suggesting that proton transfer to C6 of GB can be realized in state IV to trigger subsequent cyclization. The PES scan indicated that this protonation-induced 2,7-cyclization is a concerted but asynchronous, highly exothermic (21.3 kcal mol<sup>-1</sup>) process with a barrier of 12.7 kcal mol<sup>-1</sup> (Supplementary Figs. 14 and 15). During this process the carbonyl oxygen of Ala183 stays hydrogen bound to the water molecule and probably plays a major role in its active site fixation during the protonation-induced cyclization of GB (Fig. 2d,e). Hence Gly182 plays a crucial role in the deprotonation–reprotonation sequence, because it first serves as the base and then as the acid mediated through one water molecule.

Our previous mutation experiments indicated that the G182A enzyme variant is still active in generating GB, even in greater amounts than the wild type, while the production of SD is strongly reduced<sup>21</sup>, demonstrating a retained function in the deprotonation with simultaneous loss of function regarding reprotonation of GB. The homology modelled structure of state II for the G182A variant (Supplementary Fig. 16) reveals that the methyl group of Ala182 does not sterically interfere with intermediate  $A^+$ , and thus Ala182 can still function as a base in the formation of GB. On the other hand, G182A exchange may not only increase the hydrophobicity of the active site but may also sterically block the trajectory of the water whose introduction would be critical for mediating the reprotonation of GB. This hypothesis is supported by PES scans from state II to IV for the G182A mutant (Supplementary Figs. 17 and 18), which indicated that the mutant may make it less energetically feasible for water to move and to coordinate with Ala182. Therefore, the combined computational and previous experimental results<sup>21</sup> obtained for the G182A variant provide additional indirect evidence for the role of Gly182 and the active-site water in the deprotonation–reprotonation sequence.

**Confirmation of the reprotonation step.** According to the above QM/MM MD simulations, a stereochemical prediction for the protonation of GB can be made that should proceed with proton transfer to the *re* face at C6 (Fig. 2d,e). To confirm these computational results experimentally, (6-<sup>13</sup>C)FPP<sup>39</sup> was converted with SdS in a deuterium oxide (D<sub>2</sub>O) buffer, which leads through (6-<sup>13</sup>C)GB ultimately to (6-<sup>13</sup>C,6-<sup>2</sup>H)SD (FPP numbering; Fig. 6a). Incorporation of deuterium at C6 of SD in this double-labelling experiment can be concluded from the slightly upfield-shifted triplet for this carbon resulting from <sup>13</sup>C–<sup>2</sup>H spin coupling (Fig. 6b)<sup>42</sup>, while the



**Fig. 6 | Stereochemical course of the protonation-induced cyclization of GB to SD.** **a**, Conversion of ( $6\text{-}^{13}\text{C}$ )FPP in  $\text{D}_2\text{O}$  buffer with SdS into ( $6\text{-}^{13}\text{C}$ )GB and ( $6\text{-}^{13}\text{C}$ )SD. Double-headed arrows indicate key NOESY correlations observed for unlabelled SD for the assignment of hydrogens 1-*pro-R* and 1-*pro-S*. **b**, Partial  $^{13}\text{C}$ -NMR spectra showing signals for C6 (and C5) of unlabelled SD (top) and ( $6\text{-}^{13}\text{C}$ , $6\text{-}^2\text{H}$ )SD (bottom, upfield-shifted triplet signal). The additional singlet arises from non-deuterated ( $6\text{-}^{13}\text{C}$ )SD. **c**, Partial heteronuclear multiple quantum correlation spectrum (all signals for C–H correlations pointing upwards) showing signals for C6 (and C5) of unlabelled SD (left), and partial HSQC spectrum (C–H correlations for  $\text{CH}_3$  and CH groups pointing upwards and for  $\text{CH}_2$  groups pointing downwards) of ( $6\text{-}^{13}\text{C}$ , $6\text{-}^2\text{H}$ )SD (right, upfield-shifted signal pointing upwards). The signals pointing down arise from non-deuterated ( $6\text{-}^{13}\text{C}$ )SD. Signals for C5 are not visible, because this carbon is not labelled.  $\delta$ , chemical shifts in ppm.

stereochemical course of deuterium incorporation into the 6-*pro-R* position of SD, confirming *re* face attack, is evident from the HSQC spectrum of labelled SD (Fig. 6c)<sup>43</sup>.

**Terminal deprotonation step towards SD.** After 2,7-cyclization to  $\text{B}^+$ , the water molecule forms a hydrogen bond with the 8-*pro-R* hydrogen ( $\text{H8}_R$ ) and the cationic centre is located at C3 (Fig. 2e), showing electrostatic interaction with the negatively charged OPP group. Following cyclization to  $\text{B}^+$ , the average distance between C15 and the closest oxygen of the OPP group decreases from 5.25 Å in stage IV (Fig. 2d) to 3.36 Å in stage V (Fig. 2e), allowing for final deprotonation with transfer of a proton from C15 to OPP that is exothermic by 11.8 kcal mol<sup>-1</sup> and proceeds through a barrier of 5.3 kcal mol<sup>-1</sup> (Supplementary Fig. 19; alternatives were not explored, but no other suitable base was observed in the single QM/MM MD simulation of state IV). After completion of this deprotonation to SD, the water molecule loses all stable interactions with the product and all active-site residues, except for one hydrogen bond with Gly182 (Fig. 2f). This observed collapse of the active-site hydrogen-bonding network may be important for subsequent product release.

## Discussion

In summary, this work combining QM(DFT)/MM simulations with isotopic labelling experiments gives very detailed insights

into the mechanism of a terpene cyclase and substantially deepens our understanding of the complex reaction cascades performed by this intriguing class of enzyme. Specifically, our work highlights a catalytic role of the carbonyl oxygen of Gly182 that first recognizes the reactive conformation of the substrate through hydrogen bond interaction with the terminal alkenyl hydrogen of the substrate, and then stabilizes the cyclized (*E,E*)-germacradienyl cation  $\text{A}^+$  through electrostatic interaction. This intermediate is relevant to a large number of sesquiterpenes<sup>44</sup>, suggesting that this finding may also be of importance in the biosynthesis of other molecules from this class. The subsequent cooperation of Gly182 with an active-site water is important in the deprotonation–reprotonation mechanism of SdS catalysis, which gives a reasonable explanation for a long-standing problem of how neutral intermediates along terpene cyclization cascades can be reactivated. A similar involvement of a side-chain (amide) carbonyl oxygen acting together with water has been discussed for terminal deprotonation steps by type II diterpene synthases, but here also an additional histidine is involved whose basicity is highly important for this step<sup>45,46</sup>. In the present case, only a backbone carbonyl oxygen is involved in the deprotonation–reprotonation sequence and it is understandable that the reduced basicity of this group is needed to achieve this step, because not only proton uptake in the deprotonation, but also transfer back to the substrate in the reprotonation, must proceed smoothly. Other previously discussed bases, such as an active-site histidine residue or

pyrophosphate, may be less suitable for orchestration of deprotonation–reprotonation sequences, because they are sufficiently basic for terminal deprotonation steps but possibly too basic to allow triggering of reprotonations by their conjugate acids.

In addition, the main-chain carbonyl oxygens of Asp181 and Ala183 stabilize the pose of the substrate, reactive intermediates and the water in the active pocket of the enzyme through hydrogen bond interactions. All three functional residues (Asp181, Gly182 and Ala183) are located at the G1/2 helix break of SdS with their main-chain carbonyl oxygens pointing into the active site, which is a strictly conserved structural motif of type I TSs from bacteria, fungi and plants<sup>21</sup>. This may indicate a more general relevance of the findings of this study for TS catalysis. Interestingly, another study that supports this view has recently been published in regard to bacterial germacadien-11-ol synthase, for which a mutation of Gly188 at the G1/2 helix break (G188A, corresponding to G182A of SdS reported here) also resulted in the accumulation of the neutral intermediate isolepidozene, which is further converted in the wild type by reprotonation and hydroxylation<sup>47</sup>. However, additional work on other TSs is required to further confirm the generalizability of the detailed mechanistic model obtained here for SdS catalysis.

## Methods

**Computational methods.** The homotetrameric crystal structure of SdS (PDB code: 4OKZ; resolution: 1.9 Å) in complex with (Mg<sup>2+</sup>)<sub>3</sub>, as well as the non-reactive substrate surrogate DHFPP, was selected to build up our simulation models. For each model, a 30 ns classical MD simulation with AMBER99SB force field was performed to relax the enzyme–substrate complex and provide the initial structure for QM/MM MD simulations. In QM/MM MD simulations, the system was divided into QM and MM parts. The QM subsystem was described with the M06-2X method in combination with the 6-31 G(d) basis set, while the MM subsystem was described with the AMBER99SB force field. The QM/MM boundary was treated by the improved pseudo-bond approach. The reaction coordinate driving method was used to map out PESs. Supplementary Methods provides computational details.

**Production of SD by preparative scale incubation.** The enzymatic reaction mixture contained FPP (90 mg) dissolved in aqueous NH<sub>4</sub>HCO<sub>3</sub> solution (20 ml, 25 mM), incubation buffer (200 ml)<sup>21</sup> and SdS elution fraction (80 ml)<sup>21</sup>. After incubation with shaking at 28 °C overnight, the reaction mixture was extracted with pentane (3 × 100 ml). The combined organic layers were dried over MgSO<sub>4</sub> and concentrated under reduced pressure. The residue (9.0 mg) was then purified by column chromatography on silica gel with pentane to render SD as a colourless oil (Supplementary Methods shows spectroscopic characterization).

**Isotopic labelling experiments.** Isotopic labelling experiments were performed with the substrates and enzymes listed in Supplementary Table 1. Reaction mixtures contained substrates (1 mg each) in aqueous NH<sub>4</sub>HCO<sub>3</sub> solution (1 ml, 25 mM), enzyme elution fractions (2 ml each) and incubation buffer (4–7 ml) to a total volume of 10 ml. After incubation with shaking at 28 °C overnight, reaction mixtures were extracted with C<sub>6</sub>D<sub>6</sub> (0.6 + 0.3 ml). Extracts were dried over MgSO<sub>4</sub> and directly analysed by nuclear magnetic resonance (NMR) spectroscopy.

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

The authors declare that data supporting the findings of this study are available within the article and its Supplementary Information file. FPP parameters and key structures are given as Supplementary Data files. Data that support the plots within the paper and other findings of this study are available from the corresponding authors on reasonable request.

Received: 30 March 2021; Accepted: 5 January 2022;

Published online: 17 February 2022

## References

- Christianson, D. W. Structural and chemical biology of terpenoid cyclases. *Chem. Rev.* **117**, 11570–11648 (2017).
- Christianson, D. W. Structural biology and chemistry of the terpenoid cyclases. *Chem. Rev.* **106**, 3412–3442 (2006).
- Dickschat, J. S. Bacterial diterpene biosynthesis. *Angew. Chem. Int. Ed.* **58**, 15964–15976 (2019).
- Hong, Y. J. & Tantillo, D. J. Biosynthetic consequences of multiple sequential post-transition-state bifurcations. *Nat. Chem.* **6**, 104–111 (2014).
- Tantillo, D. J. Biosynthesis via carbocations: theoretical studies on terpene formation. *Nat. Prod. Rep.* **28**, 1035–1053 (2011).
- Hong, Y. J. & Tantillo, D. J. Branching out from the bisabolyl cation. Unifying mechanistic pathways to barbatene, bazzanene, chamigrene, chamipinene, cumacrene, cuprenene, dunnine, isobazzanene, iso-gamma-bisabolene, isochamigrene, laurene, microbotene, sesquithujene, sesquisabinene, thujopsene, trichodiene, and widdradiene sesquiterpenes. *J. Am. Chem. Soc.* **136**, 2450–2463 (2014).
- Hong, Y. J. & Tantillo, D. J. Consequences of conformational preorganization in sesquiterpene biosynthesis: theoretical studies on the formation of the bisabolene, curcumene, acoradiene, zizaene, cedrene, duprezianene, and sesquithuriferol sesquiterpenes. *J. Am. Chem. Soc.* **131**, 7999–8015 (2009).
- Mahadevi, A. S. & Sastry, G. N. Cation- $\pi$  interaction: its role and relevance in chemistry, biology, and material science. *Chem. Rev.* **113**, 2100–2138 (2013).
- Tantillo, D. J. Importance of inherent substrate reactivity in enzyme-promoted carbocation cyclization/rearrangements. *Angew. Chem. Int. Ed.* **56**, 10040–10045 (2017).
- Chen, M. et al. Mechanistic insights from the binding of substrate and carbocation intermediate analogues to aristolochene synthase. *Biochemistry* **52**, 5441–5453 (2013).
- Noel, J. P. et al. Structural elucidation of cisoid and transoid cyclization pathways of a sesquiterpene synthase using 2-fluorofarnesyl diphosphates. *ACS Chem. Biol.* **5**, 377–392 (2010).
- Minami, A., Ozaki, T., Liu, C. & Oikawa, H. Cyclopentane-forming di/ sesterterpene synthases: widely distributed enzymes in bacteria, fungi, and plants. *Nat. Prod. Rep.* **35**, 1330–1346 (2018).
- Rinkel, J. et al. Mechanisms of the diterpene cyclases  $\beta$ -pinacene synthase from *Dictyostelium discoideum* and hydroxyrene synthase from *Streptomyces clavuligerus*. *Chem. Eur. J.* **23**, 10501–10505 (2017).
- Rinkel, J., Lauterbach, L. & Dickschat, J. S. Spata-13,17-diene synthase—an enzyme with sesqui-, di-, and sesterterpene synthase activity from *Streptomyces xinghaiensis*. *Angew. Chem. Int. Ed.* **56**, 16385–16389 (2017).
- Bian, G. K. et al. A clade II-D fungal chimeric diterpene synthase from *Colletotrichum gloeosporioides* produces dolasta-1(15),8-diene. *Angew. Chem. Int. Ed.* **57**, 15887–15890 (2018).
- Rinkel, J., Lauterbach, L. & Dickschat, J. S. A branched diterpene cascade: the mechanism of spinodiene synthase from *Saccharopolyspora spinosa*. *Angew. Chem. Int. Ed.* **58**, 452–455 (2019).
- Zhang, F., Chen, N. H., Zhou, J. W. & Wu, R. B. Protonation-dependent diphosphate cleavage in FPP cyclases and synthases. *ACS Catal.* **6**, 6918–6929 (2016).
- Hong, Y. J. & Tantillo, D. J. Feasibility of intramolecular proton transfers in terpene biosynthesis-guiding principles. *J. Am. Chem. Soc.* **137**, 4134–4140 (2015).
- Seemann, M. et al. Pentalene synthase. Analysis of active site residues by site-directed mutagenesis. *J. Am. Chem. Soc.* **124**, 7681–7689 (2002).
- Jiang, J., He, X. & Cane, D. E. Biosynthesis of the earthy odorant geosmin by a bifunctional *Streptomyces coelicolor* enzyme. *Nat. Chem. Biol.* **3**, 711–715 (2007).
- Baer, P. et al. Induced-fit mechanism in class I terpene cyclases. *Angew. Chem. Int. Ed.* **53**, 7652–7656 (2014).
- Köksal, M., Chou, W. K. W., Cane, D. E. & Christianson, D. W. Structure of 2-methylisoborneol synthase from *Streptomyces coelicolor* and implications for the cyclization of a noncanonical C-methylated monoterpenoid substrate. *Biochemistry* **51**, 3011–3020 (2012).
- Aaron, J. A., Lin, X., Cane, D. E. & Christianson, D. W. Structure of epi-isozizaene synthase from *Streptomyces coelicolor* A3(2), a platform for new terpenoid cyclization templates. *Biochemistry* **49**, 1787–1797 (2010).
- Xu, H., Rinkel, J. R. & Dickschat, J. S. Isoishwarane synthase from *Streptomyces lincolnensis*. *Org. Chem. Front.* **8**, 1177–1184 (2021).
- Lauterbach, L., Rinkel, J. & Dickschat, J. S. Two bacterial diterpene synthases from *Allokutzneria albata* for bonnadiene and for phomopsene and allokutznerene. *Angew. Chem. Int. Ed.* **57**, 8280–8283 (2018).
- Tantillo, D. J. Walking in the woods with quantum chemistry—applications of quantum chemical calculations in natural products research. *Nat. Prod. Rep.* **30**, 1079–1086 (2013).
- Ganguly, A., Thaplyal, P., Rosta, E., Bevilacqua, P. C. & Hammes-Schiffer, S. Quantum mechanical/molecular free energy simulations of the self-cleavage reaction in the hepatitis Delta virus ribozyme. *J. Am. Chem. Soc.* **136**, 1483–1496 (2014).
- Rooklin, D. W., Lu, M. & Zhang, Y. Revelation of a catalytic calcium-binding site elucidates unusual metal dependence of a human apyrase. *J. Am. Chem. Soc.* **134**, 15595–15603 (2012).
- Ke, Z., Smith, G. K., Zhang, Y. & Guo, H. Molecular mechanism for eliminylation, a newly discovered post-translational modification. *J. Am. Chem. Soc.* **133**, 11103–11105 (2011).



30. Wang, Y.-H., Xie, H., Zhou, J., Zhang, F. & Wu, R. Substrate folding modes in trichodiene synthase: a determinant of chemo- and stereoselectivity. *ACS Catal.* **7**, 5841–5846 (2017).
31. Diao, H. et al. Biosynthetic mechanism of lanosterol: a completed story. *ACS Catal.* **10**, 2157–2168 (2020).
32. Chen, N., Wang, S., Smentek, L., Hess, B. A. & Wu, R. Biosynthetic mechanism of lanosterol: cyclization. *Angew. Chem. Int. Ed.* **54**, 8693–8696 (2015).
33. Das, S., Dixit, M. & Major, D. T. First principles model calculations of the biosynthetic pathway in selinadiene synthase. *Bioorg. Med. Chem.* **24**, 4867–4870 (2016).
34. Rabe, P. et al. Terpene cyclases from social *Amoebae*. *Angew. Chem. Int. Ed.* **55**, 15420–15423 (2016).
35. Rinkel, J. & Dickschat, J. S. Addressing the chemistry of germacrene A by isotope labeling experiments. *Org. Lett.* **21**, 2426–2429 (2019).
36. Hahn, F. M., Hurlburt, A. P. & Poulter, C. D. *Escherichia coli* open reading frame 696 is *idi*, a nonessential gene encoding isopentenyl diphosphate isomerase. *J. Bacteriol.* **181**, 4499–4504 (1999).
37. Cornforth, J. W., Cornforth, R. H., Popjak, G. & Yengoyan, L. Studies on the biosynthesis of cholesterol XX. Steric course of decarboxylation of 5-pyrophosphomevalonate and of the carbon to carbon bond formation in the biosynthesis of farnesyl pyrophosphate. *J. Biol. Chem.* **241**, 3970–3987 (1966).
38. Zhou, J. et al. Protonation-triggered carbon-chain elongation in geranyl pyrophosphate synthase (GPPS). *ACS Catal.* **5**, 4466–4478 (2015).
39. Rabe, P. et al. Conformational analysis, thermal rearrangement and EI-MS-fragmentation mechanism of (1(10)E,4E,6S,7R)-germacradien-6-ol by <sup>13</sup>C-labeling experiments. *Angew. Chem. Int. Ed.* **54**, 13448–13451 (2015).
40. Lin, F.-L. et al. Mechanistic characterization of the fusicoccane-type diterpene synthase for myrothec-15(17)-en-7-ol. *ACS Catal.* **10**, 4306–4312 (2020).
41. Dereka, B. et al. Crossover from hydrogen to chemical bonding. *Science* **371**, 160–164 (2021).
42. Rabe, P., Pahirulzaman, K. A. K. & Dickschat, J. S. Structures and biosynthesis of corvol ethers—sesquiterpenes from the actinomycete *Kitasatospora setae*. *Angew. Chem. Int. Ed.* **54**, 6041–6045 (2015).
43. Rabe, P., Rinkel, J., Klapschinski, T. A., Barra, L. & Dickschat, J. S. A method for investigating the stereochemical course of terpene cyclisations. *Org. Biomol. Chem.* **14**, 158–164 (2016).
44. Xu, H. & Dickschat, J. S. Germacrene A—a central intermediate in sesquiterpene biosynthesis. *Chem. Eur. J.* **26**, 17318–17341 (2020).
45. Potter, K., Criswell, J., Zi, J., Stubbs, A. & Peters, R. J. Novel product chemistry from mechanistic analysis of *ent*-copalyl diphosphate synthases from plant hormone biosynthesis. *Angew. Chem. Int. Ed.* **53**, 7198–7202 (2014).
46. Potter, K. C. et al. Blocking deprotonation with retention of aromaticity in a plant *ent*-copalyl diphosphate synthase leads to product rearrangement. *Angew. Chem. Int. Ed.* **55**, 634–638 (2016).
47. Srivastava, P. L. et al. Redesigning the molecular choreography to prevent hydroxylation in germacradien-11-ol synthase catalysis. *ACS Catal.* **11**, 1033–1041 (2021).

## Acknowledgements

This work was supported by the German Research Foundation DFG (no. DI1536/7-2 to J.S.D.); the National Key Research and Development Program of China (nos. 2018YFA0903200 and 2018YFA0903201 to H.G.); the National Natural Science Foundation of China (nos. 21803080 and 32070042 to Y.-H.W., 81925037 to H.G., 81872759 to P.-H.S. and 21773313 to R.W.); the National High-level Personnel of Special Support Program (no. 2017RA2259 to H.G.); the Chang Jiang Scholars Program (Young Scholar) from the Ministry of Education of China (to H.G.); and the K. C. Wong Education Foundation (to H.G.). We thank X.-S. Yao (Guangzhou) for his support of this study through the 111 Project of the Ministry of Education of the People's Republic of China (no. B13038), and the Guangzhou and Shenzhen Supercomputer Center for providing the computational source.

## Author contributions

J.S.D., R.W. and Y.-H.W. designed the research. J.S.D. supervised the experimental procedures. R.W. and P.-H.S. supervised the computational work. R.W. provided the Qchem-Tinker software package. Y.-H.W. and J.Z. performed the QM/MM MD simulations. H.X. and E.C. carried out the experimental work. Y.-H.W., J.Z., X.-B.C., Y.-Q.Z., W.-L.L., G.-D.C., D.H., H.G. and P.-H.S. analysed computational data. Y.-H.W., J.S.D., J.Z. and H.X. wrote the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41929-022-00735-0>.

**Correspondence and requests for materials** should be addressed to Ruibo Wu, Ping-Hua Sun or Jeroen S. Dickschat.

**Peer review information** *Nature Catalysis* thanks Hideaki Oikawa, Per-Olof Syren and Marc van der Kamp for their contribution to the peer review of this work.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2022

## Appendix G

### Isoishwarane synthase from *Streptomyces lincolnensis*

*Org. Chem. Front.* **2021**, *8*, 1177

DOI: 10.1039/D0QO01583K

## RESEARCH ARTICLE

View Article Online  
View Journal | View Issue

Check for updates

Cite this: *Org. Chem. Front.*, 2021, **8**, 1177Received 16th December 2020,  
Accepted 18th January 2021

DOI: 10.1039/d0qo01583k

rsc.li/frontiers-organic

Isoishwarane synthase from *Streptomyces lincolnensis*†

Houchao Xu, Jan Rinkel and Jeroen S. Dickschat \*

A new sesquiterpene synthase from *Streptomyces lincolnensis* was identified as isoishwarane synthase. The absolute configuration of its product was determined by an enantioselective deuteration strategy. The cyclisation mechanism from farnesyl diphosphate to isoishwarane was studied in incubation experiments with various isotopically labelled substrates and by site-directed mutagenesis. GC/MS analysis of all 15 isotopomers of (<sup>13</sup>C)isoishwarane also gave insights into the EI-MS fragmentation mechanism.

## Introduction

Terpene synthases (TSs) are remarkable enzymes that convert acyclic and achiral oligoprenyl diphosphates such as geranyl diphosphate (GPP, monoterpenes), farnesyl diphosphate (FPP, sesquiterpenes), geranylgeranyl diphosphate (GGPP, diterpenes) and geranylgeranyl diphosphate (GFPP, sesterterpenes) into usually (poly)cyclic terpene hydrocarbons or alcohols containing several contiguous stereogenic centres. For type I enzymes, these complex transformations are initiated by the abstraction of diphosphate with formation of an allyl cation that can undergo a cascade reaction involving typical carbocation chemistry with cyclisation steps through intramolecular attack of an olefinic double bond to the cation, hydride or proton migrations, and Wagner-Meerwein rearrangements. Some TSs show an astonishing selectivity and form only one product,<sup>1–3</sup> while in other cases multiple structurally related products are generated.<sup>4–6</sup> During the past years several bacterial type I TSs have been identified and characterised,<sup>7–10</sup> and a recent study on the distribution of TSs in bacteria with completely sequenced genomes from the genus *Streptomyces*<sup>11</sup> revealed that especially geosmin synthases,<sup>12</sup> 2-methylisoborneol synthases<sup>13,14</sup> and *epi*-isozizaene synthases<sup>2</sup> are widespread in this genus. One strain included in this study was *Streptomyces lincolnensis* NRRL 2936 that encodes a geosmin synthase and an *epi*-isozizaene synthase homolog in its genome. A third TS homolog did not show a close relationship to any characterised enzyme (Fig. S1†). Here we describe

the characterisation of this TS that exhibits a novel function as isoishwarane synthase, an enzyme mechanistic study by isotopic labelling experiments and site-directed mutagenesis, and labelling experiments that gave insights into the mechanisms of the mass spectrometric fragmentation reactions of **1**.

## Results and discussion

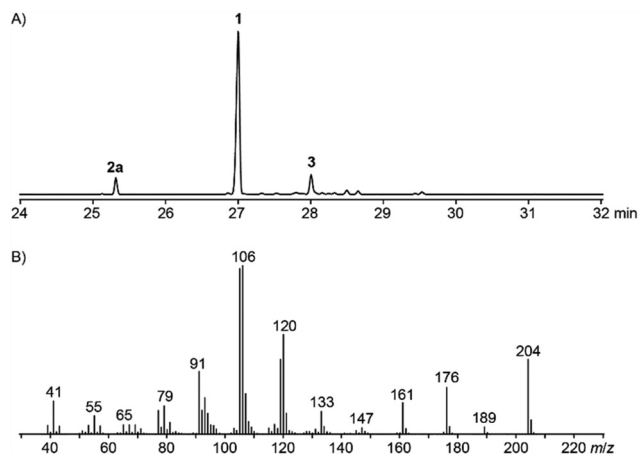
## Identification of the product of isoishwarane synthase

The gene of an uncharacterised TS from *Streptomyces lincolnensis* NRRL 2936 (accession number WP\_067429395, gene locus tag SLINC\_RS09705, Fig. S2†) was amplified by PCR and cloned into the *Escherichia coli* expression vector pYE-Express by homologous recombination in yeast.<sup>15</sup> After gene expression the protein was purified (Fig. S3†) and used in test incubations with GPP, FPP, GGPP and GFPP, followed by GC/MS analysis of the products. FPP was the only accepted substrate and was efficiently converted into a sesquiterpene hydrocarbon as main product whose mass spectrum was not included in our mass spectral libraries (Fig. 1). Furthermore, small amounts of  $\beta$ -elemene (**2a**), the Cope rearrangement product of germacrene A (**2**) that is formed under the thermal impact of the GC analysis,<sup>16–18</sup> and valencene<sup>19</sup> (**3**) were detected. The main product was isolated from a large scale incubation with FPP and its structure was elucidated by NMR spectroscopy (Table 1, Fig. S4–S10†), resulting in the structure of isoishwarane (**1**, Scheme 1) and thus in the identification of the TS as isoishwarane synthase (IWS).

The proposed cyclisation mechanism from FPP to **1** (Scheme 1) starts with the 1,10-cyclisation of FPP to the (*E,E*)-germacradienyl cation (**A**), followed by deprotonation to germacrene A (**2**), an important intermediate towards many sesquiterpenes,<sup>20</sup> explaining the observation of this compound as a side product of IWS. Its reprotonation at C6 induces a second cyclisation to **B** that is followed by a 1,2-hydride shift to **C** and

Kekulé-Institut für Organische Chemie und Biochemie, Rheinische Friedrich-Wilhelms-Universität Bonn, Gerhard-Domagk-Straße 1, 53121 Bonn, Germany.  
E-mail: dickschat@uni-bonn.de

†Electronic supplementary information (ESI) available: Experimental procedures, phylogenetic tree of bacterial terpene synthases, NMR spectra of isoishwarane, results from labelling experiments, and analysis of highly conserved residues in bacterial terpene synthases. See DOI: 10.1039/d0qo01583k



**Fig. 1** GC/MS analysis of the products obtained with IWS. (A) Total ion chromatogram (TIC) of products obtained from FPP (**2a**:  $\beta$ -elemene, retention index  $I = 1392$ , lit.  $I = 1389$ ,<sup>19</sup> **3**: valencene,  $I = 1497$ , lit.  $I = 1496$ <sup>19</sup>). (B) EI mass spectrum of **1** ( $I = 1457$ ).

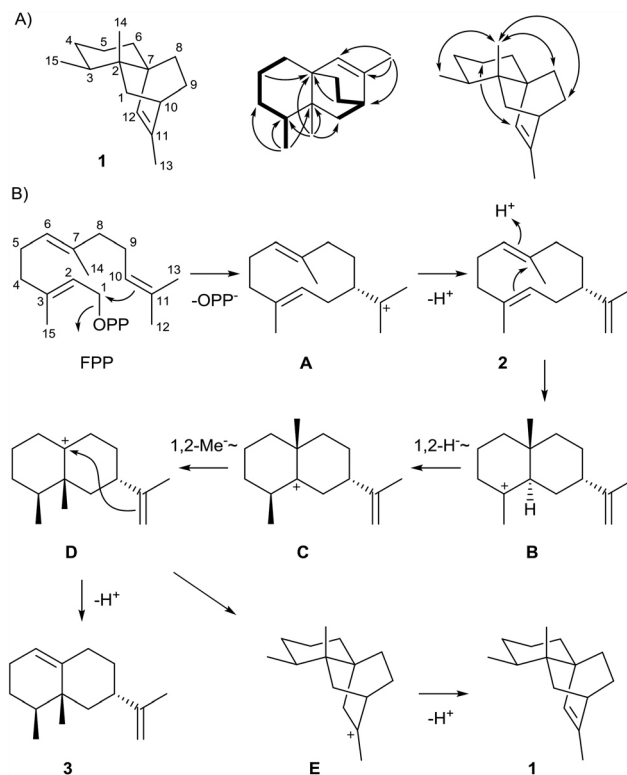
**Table 1** NMR data of isoishwarane (**1**)

$C^a$	$^{13}C$ ( $\delta$ ) <sup>b</sup>	type	$^1H$ ( $\delta$ , m, $J$ , int) <sup>b</sup>
1	42.96	CH <sub>2</sub>	1.21 (ddd, $^2J = 12.4$ , $^3J = 2.7$ , $^4J = 2.7$ , $^c H_\alpha$ ) 1.15 (dd, $^2J = 12.4$ , $^3J = 3.1$ , $H_\beta$ )
2	39.35	C <sub>q</sub>	—
3	40.53	CH	1.30 (m)
4	30.65	CH <sub>2</sub>	1.28 (m, $H_\alpha$ ) 1.10 (m, $H_\beta$ )
5	23.19	CH <sub>2</sub>	1.59 (m, $H_\alpha$ ) 1.55 (m, $H_\beta$ )
6	31.57	CH <sub>2</sub>	1.41 (m, $H_\beta$ ) 1.36 (m, $H_\alpha$ )
7	41.50	C <sub>q</sub>	—
8	31.84	CH <sub>2</sub>	1.63 (m, $H_\beta$ ) 0.92 (m, $H_\alpha$ )
9	24.38	CH <sub>2</sub>	1.44 (m, $H_\beta$ ) 1.34 (m, $H_\alpha$ )
10	37.30	CH	2.11 (br s)
11	140.04	C <sub>q</sub>	—
12	131.18	CH	5.69 (br s)
13	20.10	CH <sub>3</sub>	1.72 (d, $^3J = 1.4$ )
14	17.11	CH <sub>3</sub>	0.86 (s)
15	16.11	CH <sub>3</sub>	0.72 (d, $^3J = 6.7$ )

<sup>a</sup> Carbon numbering as in Scheme 1. <sup>b</sup> Recorded in C<sub>6</sub>D<sub>6</sub> recorded at 298 K. Chemical shifts  $\delta$  in ppm, multiplicity: s = singlet, d = doublet, m = multiplet, br = broad, coupling constants  $J$  are given in Hertz. <sup>c</sup>  $^4J$  coupling is explainable as W coupling with H<sub>9 $\alpha$</sub> .

a 1,2-methyl group migration to **D**. Deprotonation of this intermediate from C6 explains the formation of side product **3**. Along the main pathway, another ring closure to **E** and deprotonation result in **1**.

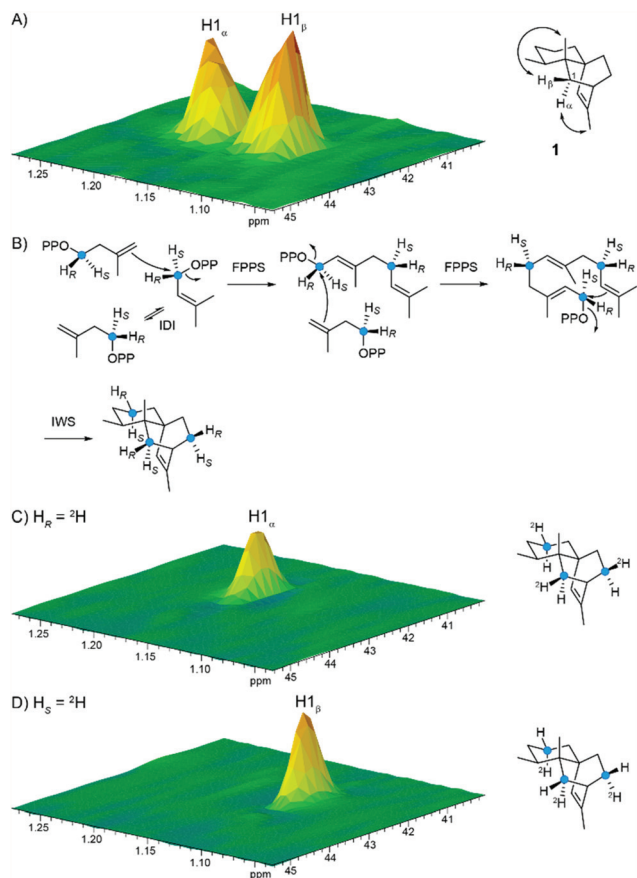
The absolute configurations of terpenes can be determined using enantioselectively deuterated oligoprenyl diphosphate precursors. Their conversion by a TS lead to stereoselectively deuterated compounds with one of the diastereotopic hydrogen atoms in a methylene group of a chiral terpene being exchanged against deuterium. In this way, a stereocentre of defined configuration is introduced into the terpene that can



**Scheme 1** Isoishwarane (**1**) from *S. lincolnensis*. (A) Structure elucidation by NMR (bold lines:  $^1H$ , $^1H$ -COSY, single headed arrows: HMBC, double headed arrows: NOESY correlations). Carbon numbering indicates the origin of each carbon from FPP by same number. (B) Cyclisation mechanism from FPP to **1**.

be used to determine the relative orientation of the naturally present stereogenic centres in the molecule, thus giving access to the absolute configuration of the terpene. We have recently developed an approach that makes use of stereoselectively deuterated isotopomers of IPP that are applied in the enzymatic synthesis of enantioselectively deuterated oligoprenyl diphosphate precursor through an oligoprenyl diphosphate synthase.<sup>18,21</sup> Additional  $^{13}C$  labellings were introduced at the deuterated carbons to allow for a highly sensitive HSQC analysis without the need of compound purification.

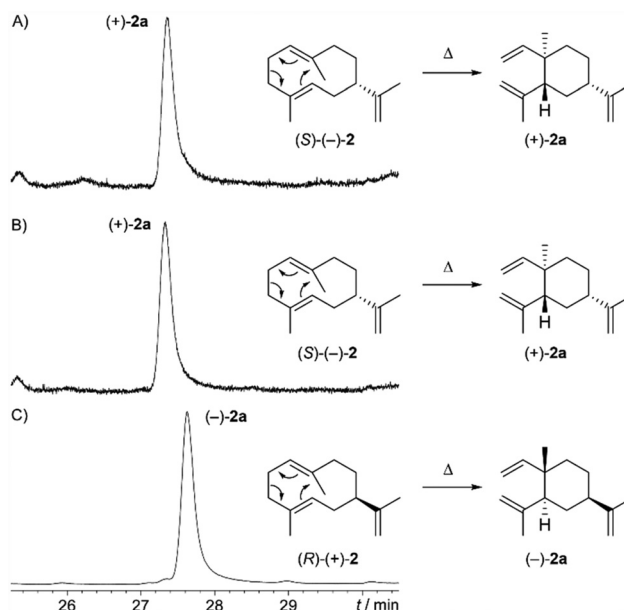
The absolute configuration of **1** was determined by conversion of (*R*)- and (*S*)-(1- $^{13}C$ ,1- $^2H$ )isopentenyl diphosphate (IPP)<sup>18</sup> with IPP isomerase (IDI) from *E. coli*,<sup>18,22</sup> FPP synthase (FPPS) from *Streptomyces coelicolor*<sup>23</sup> and IWS (Fig. 2, S12 and S13<sup>†</sup>). The incorporation of deuterium labellings together with the full assignments of the  $^1H$ -NMR data for all hydrogen atoms by NOESY (Scheme 1, Table 1) revealed the absolute configuration of (2*R*,3*S*,7*S*,10*S*)-**1**. Similar experiments with (*E*)- and (*Z*)-(4- $^{13}C$ ,4- $^2H$ )IPP<sup>21</sup> in conjunction with DMAPP, FPPS and IWS pointed to the same absolute configuration for **1** (Fig. S14<sup>†</sup>). The data interpretation in these labelling experiments relies on the known stereochemical course of oligoprenyl diphosphate synthases with inversion of configuration at C1 of allyl diphosphates and *Si* face attack at C4 of IPP.<sup>24,25</sup> Isoishwarane is a new natural product, but the synthesis of



**Fig. 2** Absolute configuration of isoishwarane (**1**). (A) Partial HSQC spectrum for C1 of unlabelled **1**, (B) enzymatic conversion of (*R*)- and (*S*)-(1-<sup>13</sup>C,1-<sup>2</sup>H)IPP by IDI, FPPS and IWS into labelled **1**, and partial HSQC spectra for C1 of labelled **1** obtained from (C) (*R*)-(1-<sup>13</sup>C,1-<sup>2</sup>H)IPP and (D) (*S*)-(1-<sup>13</sup>C,1-<sup>2</sup>H)IPP.

(*rac*)-**1** has been achieved along the lines of a total synthesis of ishwarane (structure shown in Fig. S15†) reported 50 years ago.<sup>26,27</sup> Furthermore, (*−*)-**1** has been synthesised from the related natural product ishwarone from *Aristolochia indica*<sup>28,29</sup> whose absolute configuration was chemically correlated to (+)-nootkatane.<sup>30</sup> Therefore, bacterial (+)-**1** is enantiomeric to (*−*)-**1** connected to a related family of plant-derived terpenoids.

The absolute configuration of **2** was deduced from the absolute configuration of its Cope rearrangement product **2a** by GC analysis on a chiral stationary phase (Fig. 3). Both enantiomers of **2** were available for comparison: (*S*)-(*−*)-**2** is the product of a recently characterised TS from the actinomycete *Micromonospora marina*,<sup>18</sup> while (*R*)-(+)-**2** is produced by the W335F variant of the intermedeol synthase from the basidiomycete *Termitomyces*.<sup>31</sup> Using these reference materials, the absolute configuration of (*S*)-(*−*)-**2** was established for the side product of IWS that is in agreement with the determined absolute configuration of **1** (Scheme 1). The absolute configuration of the side product **3** was also determined by GC using a chiral stationary phase, showing a different retention time to that of (+)-**3** from orange,<sup>32,33</sup> thus identifying the compound obtained with IWS as (*−*)-**3** (Fig. S16†). The enantiomer (+)-**3** is



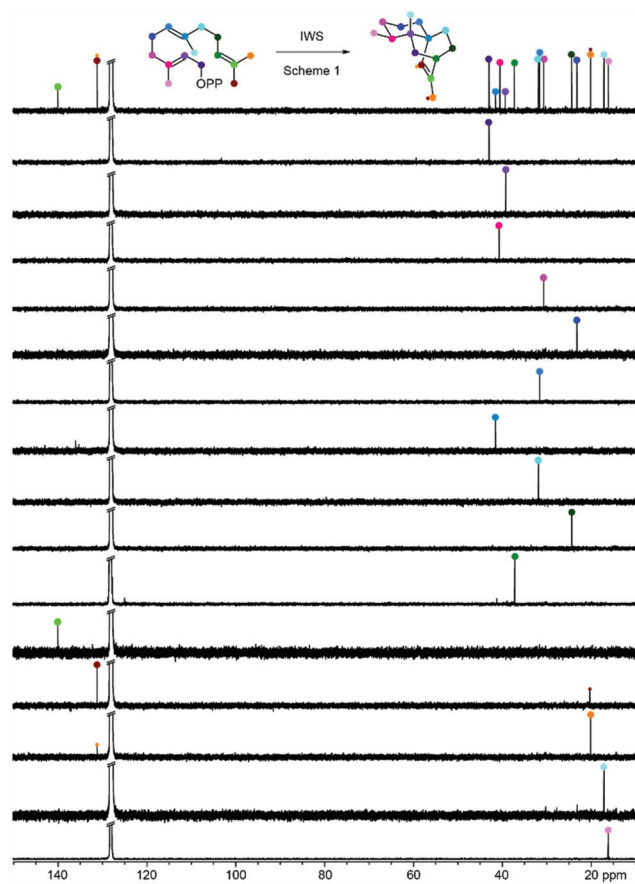
**Fig. 3** Absolute configuration of germacrene A (**2**) determined by GC on a chiral stationary phase. During GC analysis **2** rearranges to  $\beta$ -elemene (**2a**). The GC chromatograms show peaks for (A) (+)-**2a** obtained with IWS, (B) (+)-**2a** obtained with germacrene A synthase from *M. marina*, and (C) (*−*)-**2a** obtained with intermedeol synthase (W335F) from *Termitomyces*.

also produced by valencene synthase from *Vitis vinifera* (VvVal).<sup>34</sup> Thus, as frequently observed for bacterial sesquiterpenes, the absolute configurations of **1**, **2** and **3** are opposite to those of the same or related compounds in plants.<sup>20,35,36</sup>

### Cyclisation mechanism of isoishwarane synthase

The cyclisation mechanism from FPP to **1** (Scheme 1) was studied by isotopic labelling experiments. The enzymatic conversion of all 15 isotopomers of (<sup>13</sup>C)FPP<sup>37</sup> with IWS yielded labelled **1** with the incorporation of labelling into the expected positions for all substrates (Fig. 4), thus supporting the biosynthetic hypothesis including the methyl group migration from **C** to **D**. Notably, the <sup>13</sup>C labelling from C12 is mainly incorporated into the olefinic CH and the labelling from C13 is mainly introduced into the methyl group attached to the C=C double bond of **1**, but a minor exchange of these labellings is also observed. This is explainable by a relaxed regiochemistry for the deprotonation of **A** to **2** that proceeds mainly from C12, but also to a minor extent from C13. A comparable result was also obtained for germacrene A synthase from *M. marina*.<sup>18</sup>

The reprotonation of intermediate **2** at C6 for its further cyclisation to **1** was investigated by incubation of (6-<sup>13</sup>C)FPP in deuterium oxide buffer. The labelled product **1** obtained from this reaction showed a slightly upfield shifted triplet for C6 as a result of a <sup>13</sup>C-<sup>2</sup>H spin coupling, while the HSQC spectrum exhibited only one crosspeak for H6 $\alpha$ , and the crosspeak for H6 $\beta$  was abolished, indicating the specific incorporation of deuterium into the H6 $\beta$  position (Fig. 5A–D). Together with the stereochemistry at C7 of intermediate **B** that can be inferred

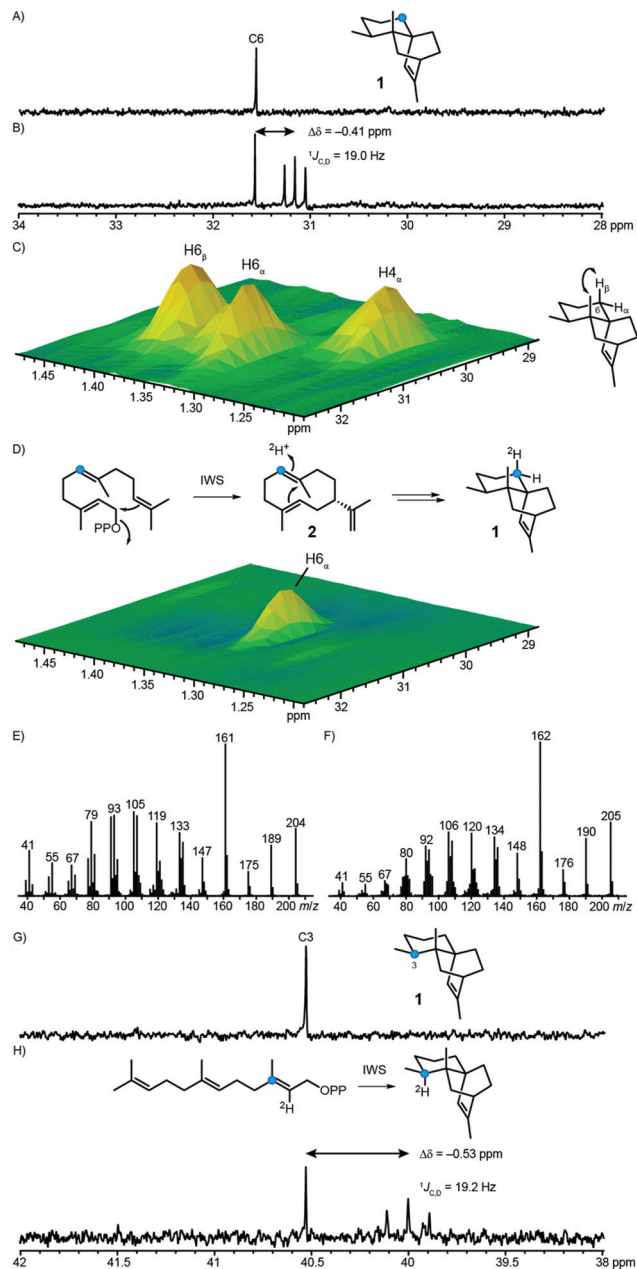


**Fig. 4** Cyclisation mechanism of IWS.  $^{13}\text{C}$ -NMR spectra of unlabelled **1** (top) and of labelled **1** obtained from all 15 isotopomers of ( $^{13}\text{C}$ )FPP with IWS. Coloured dots correlate the detected  $^{13}\text{C}$  signals to the carbons of **1**.

from the final structure of **1** this experiment demonstrates *anti* addition to the C6=C7 double bond in the cyclisation from **2** to **B**. For compound **3**, the incubation of ( $6\text{-}^{13}\text{C}$ )FPP in deuterium oxide buffer resulted in the incorporation of the  $^{13}\text{C}$ -labelling, but no additional deuterium uptake was detected by EIMS analysis (Fig. 5E and F), suggesting that the deprotonation from **D** to **3** proceeds with removal of the same proton as introduced in the reprotonation step from **2** to **B**. A complementary experiment with ( $2\text{-}^2\text{H}$ )FPP, enzymatically prepared from synthetic ( $2\text{-}^2\text{H}$ )GPP<sup>38,39</sup> and IPP with FPPS, showed full incorporation of deuterium into **3** (Fig. S17<sup>†</sup>). The 1,2-hydride shift from **B** to **C** was firmly established by enzymatic conversion of ( $3\text{-}^{13}\text{C}, 2\text{-}^2\text{H}$ )FPP<sup>40</sup> with IWS, resulting in an upfield shifted triplet for C3 of the obtained labelled **1** (Fig. 5G and H). Taken together, all labelling experiments support the mechanism for the FPP cyclisation to **1** as shown in Scheme 1.

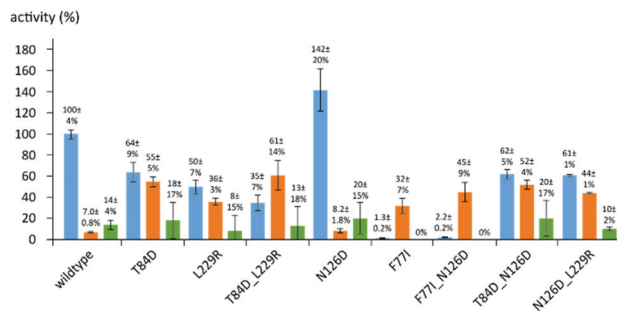
#### Site-directed mutagenesis of IWS

The unusual structure of **1** and cyclisation mechanism of IWS prompted us to investigate whether specific differences in the amino acid sequence in comparison to other bacterial TSs are relevant. Two obvious deviations can be observed in the aspar-



**Fig. 5** Cyclisation mechanism from FPP to **1**.  $^{13}\text{C}$ -NMR spectra of enzyme products obtained with IWS from (A) ( $6\text{-}^{13}\text{C}$ )FPP and (B) ( $6\text{-}^{13}\text{C}$ )FPP in  $\text{D}_2\text{O}$  buffer. (C) Partial HSQC spectrum of unlabelled **1**, (D) HSQC spectrum of ( $6\text{-}^{13}\text{C}, 6\text{-}^2\text{H}$ )-**1**. (E) EI mass spectrum of unlabelled **3**, (F) EI mass spectrum of labelled **3** obtained with IWS from ( $6\text{-}^{13}\text{C}$ )FPP in deuterium oxide buffer.  $^{13}\text{C}$ -NMR spectra of products obtained with IWS from (G) ( $3\text{-}^{13}\text{C}$ )FPP, and (H) ( $3\text{-}^{13}\text{C}, 2\text{-}^2\text{H}$ )FPP. Blue dots indicate  $^{13}\text{C}$ -labelled carbons.

tate-rich motif that is usually a variation of DDXX(X)D, but for IWS altered to  $^{80}\text{DDLHT}$  with an exchange of the third Asp against Thr, and in the NSE triad, usually represented by ND(L,I,V)XSXX(R,K)(E,D), for which in IWS the sequence  $^{222}\text{NDLHSIHLD}$  with a Leu instead of an Arg or Lys in the second last position is found. Changing these deviating to the usually observed residues in the T84D and L229R enzyme var-



**Fig. 6** Site-directed mutagenesis of IWS. Blue bars indicate production of **1**, orange bars production of **2**, and green bars production of **3**. Wildtype production of **1** is set to 100%, error bars indicate standard deviations determined from triplicates.

ants resulted in a decreased production of **1** ( $64 \pm 9\%$  and  $50 \pm 7\%$ , wildtype =  $100 \pm 4\%$ ) in favour of a nearly equivalent increase of **2** ( $55 \pm 5\%$  and  $36 \pm 3\%$ , wildtype =  $7.0 \pm 0.8\%$ , Fig. 6), showing that these residues are indeed important for the further cyclisation from **2** to **1**. The combined T84D-L229R variant resulted in a further decrease of **1** ( $35 \pm 7\%$ ) with corresponding increase of **2** ( $61 \pm 14\%$ ).

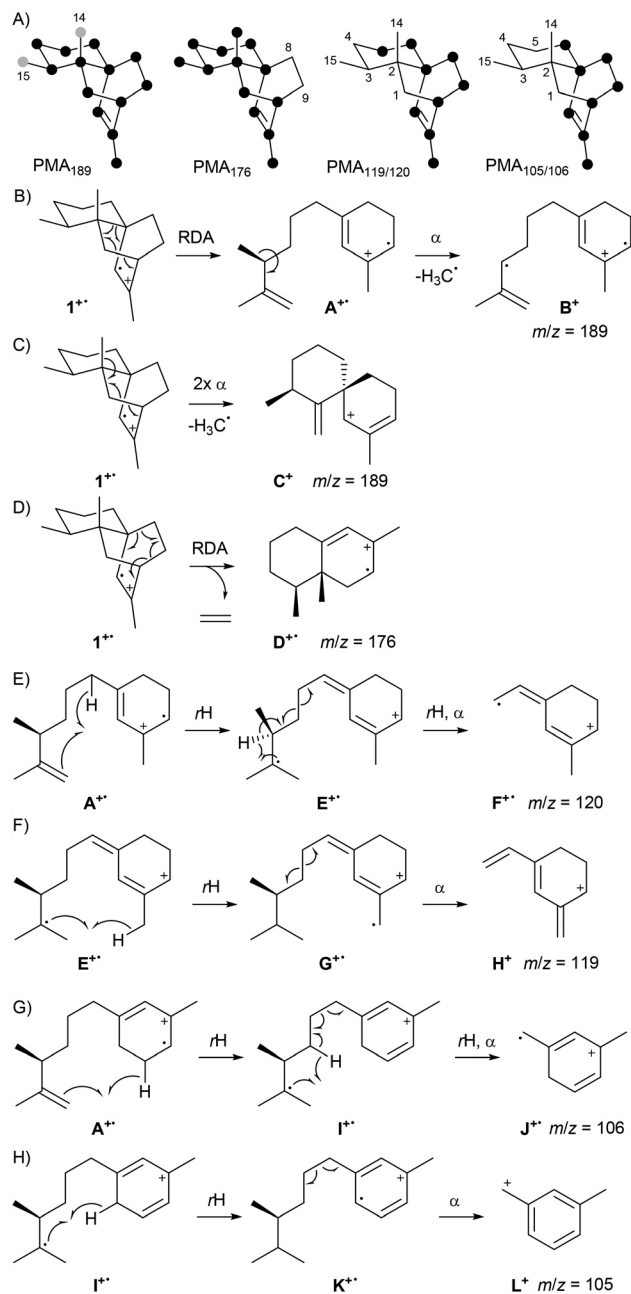
A detailed analysis of the amino acid sequences of characterised bacterial TSs and their homologs in genome sequenced bacteria shows that in most cases an Asp residue is present (Table S3<sup>†</sup>), representing a newly identified conserved residue. This Asp is located exactly 50 positions upstream of the pyrophosphate sensor, a highly conserved Arg residue with importance for catalytic activity that is involved in an active site hydrogen bond network with the substrate.<sup>41</sup> For IWS this position is occupied by Asn, and interestingly the N126D enzyme variant showed a moderately enhanced production level for **1** ( $142 \pm 20\%$ ), but almost no changes for **2** ( $8.2 \pm 1.8\%$ ). Bacterial TSs also exhibit a highly conserved Phe residue three positions upstream of the Asp-rich motif. Enzyme crystal structures revealed that this residue is involved in the stabilisation of cationic intermediates through cation- $\pi$ -interaction.<sup>41-43</sup> Accordingly, the F77Y variant of pentalenene synthase from *Streptomyces exfoliatus* showed a reduced activity,<sup>44</sup> while the F96A variant of *epi*-isozizaene synthase exhibited a lowered activity with acyclic  $\beta$ -farnesene as main product.<sup>42</sup> For the F79W variant of selina-4(15),7(11)-diene synthase from *Streptomyces pristinaespiralis* a reduced activity and product shift towards germacrene B, a neutral intermediate along the cyclisation cascade, has been observed.<sup>41</sup> Similar findings were obtained here for the F77I mutant that has a reduced activity with almost lost production of **1** ( $1.3 \pm 0.2\%$ ) and a product shift to the intermediate **2** ( $32 \pm 7\%$ ). The moderately increased activity of the N126D variant in comparison to the wildtype was paralleled for the F77I-N126D (**1**:  $2.2 \pm 0.2\%$ , **2**:  $45 \pm 9\%$ ) and N126D-L229R variants (**1**:  $61 \pm 1\%$ , **2**:  $44 \pm 1\%$ ) when compared to the F77I and L229R single exchanges, while almost no effect was observed for the T84D-N126D mutant (**1**:  $62 \pm 5\%$ , **2**:  $52 \pm 4\%$ ) in comparison to the T84D variant.

The production of **3** was  $14 \pm 4\%$  for the wildtype and ranged from  $20 \pm 15\%$  for the N126D variant to  $8 \pm 15\%$  for the L229R mutant, showing that the exchange of N126 against the usually observed Asp in this position is also beneficial for the formation of this side product. The exchange of F77 (F77I and F77I-N126D) was critical and resulted in a completely abolished production of **3** for both enzyme variants, demonstrating the importance of this residue for both downstream cyclisation products **1** and **3**.

### EI-MS fragmentation mechanism of isoishwaranone

The <sup>13</sup>C-labelled products **1** obtained with IWS from the fifteen isotopomers of (<sup>13</sup>C)FPP were analysed by GC/MS (Fig. S19<sup>†</sup>) to study the EI-MS fragmentation mechanism of this sesquiterpene hydrocarbon. For each carbon that participates in the formation of a fragment ion  $m/z$  the introduction of a <sup>13</sup>C-label will lead to an increase by 1 Da. This strategy has historically been applied through the introduction of isotope labels by total or semisynthesis.<sup>45-47</sup> In the present study, the relevant information was extracted from the mass spectra of all fifteen enzymatically prepared (<sup>13</sup>C)-**1** isotopomers and is summarised in a position-specific mass shift analysis (PMA <sub>$m/z$</sub> ),<sup>48,49</sup> in which labelled carbons resulting in a mass shift of +1 Da for a fragment ion  $m/z$  are indicated by black circles (Scheme 2A). Clear results are obtained, if one specific fragmentation mechanism leads to a fragment ion, but if different fragmentation reactions contribute to the formation of fragment ions with same  $m/z$ , representing different portions of the molecule, data interpretation becomes less clear. In such cases some carbons may only partially contribute to a fragment ion (indicated by grey circles in Scheme 2A).

The PMA<sub>189</sub> indicates that the fragment ion  $m/z = 189$  is formed by loss of either Me15 or Me14. This is explainable by ionisation of **1** at the double bond to **1**<sup>+</sup>, which may initiate a retro-Diels-Alder reaction (RDA) to **A**<sup>+</sup> (Scheme 2B). Subsequent  $\alpha$ -fragmentation with loss of Me15 yields **B**<sup>+</sup>. Alternatively, two  $\alpha$ -fragmentations from **1**<sup>+</sup> lead to **C**<sup>+</sup> with loss of Me14 (Scheme 2C). The even fragment ion at  $m/z = 176$  is generated with loss of C8 and C9, explainable by a RDA with cleavage of ethylene to **D**<sup>+</sup> (Scheme 2D). The PMA<sub>119/120</sub> demonstrates the formation of the corresponding fragment ions by loss of the C1-2(14)-3(15)-4 portion of **1**. Starting from **A**<sup>+</sup>, a hydrogen rearrangement ( $rH$ ) to **E**<sup>+</sup> followed by another hydrogen rearrangement and  $\alpha$ -cleavage can lead to **F**<sup>+</sup>, explaining the fragment ion at  $m/z = 120$  (Scheme 2E). An alternative hydrogen rearrangement from **E**<sup>+</sup> to **G**<sup>+</sup> and  $\alpha$ -cleavage to **H**<sup>+</sup> explains  $m/z = 119$  (Scheme 2F). Formation of the base peak ion at  $m/z = 106$  proceeds with loss of the C1-2(14)-3(15)-4-5 moiety of **1** and is connected to the formation of a second intensive fragment ion at  $m/z = 105$ . This can mechanistically be explained by hydrogen rearrangement from **A**<sup>+</sup> to **I**<sup>+</sup>, followed by a second hydrogen rearrangement and  $\alpha$ -cleavage to **J**<sup>+</sup> for  $m/z = 106$  (Scheme 2G). Starting from **I**<sup>+</sup> an alternative hydrogen rearrangement and  $\alpha$ -fragmentation leads to **L**<sup>+</sup> with  $m/z = 105$  (Scheme 2F). It should be emphasised that the <sup>13</sup>C-labelling experiments on which these



Scheme 2 EI-MS fragmentation mechanisms of **1**.

mechanistic proposals are based do not allow to follow the hydrogen rearrangements suggested here and alternative hydrogen migrations may be involved, but the carbon portions of **1** and the hydrogen contents for all the discussed fragment ions are clearly resolved.

### *In vivo* production and oxidation of isoishwarane

The production of the sesquiterpene hydrocarbon **1** was investigated by the collection of volatiles emitted by agar plate cultures of *S. lincolnensis* on charcoal filter traps using a closed-loop stripping apparatus (CLSA).<sup>50</sup> GC/MS analysis of the filter extracts revealed that the main volatile compound released by

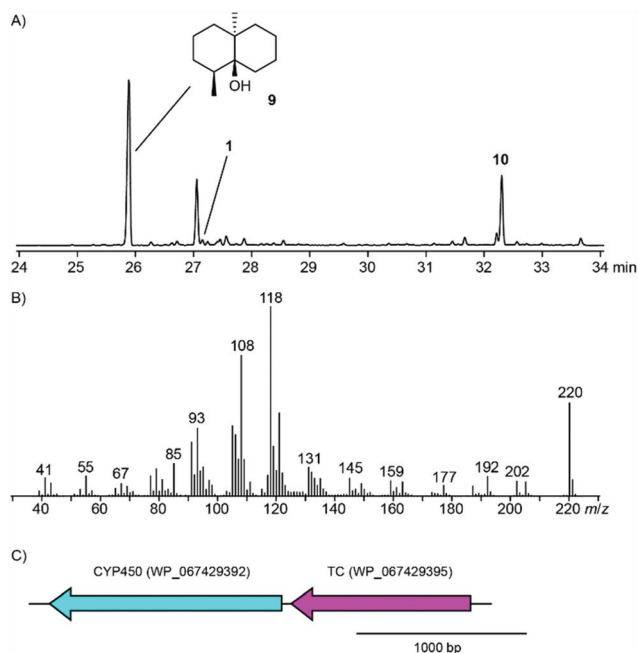


Fig. 7 Production of **1** and its oxidation product **10** by *S. lincolnensis*. (A) Total ion chromatogram of a CLSA headspace extract, (B) EI mass spectrum of **10**, (C) gene cluster of IWS (purple) and a CYP450 monooxygenase (turquoise) in *S. lincolnensis*.

*S. lincolnensis* is geosmin (**9**, Fig. 7), a degraded sesquiterpene that is produced by the geosmin synthase<sup>12</sup> (accession number WP\_067441821, gene locus tag SLINC\_RS34760) and is a typical volatile produced by many streptomycetes.<sup>51–53</sup> Compound **1** was only found in trace amounts, but clearly demonstrating production *in vivo*. The majority of **1** may be further converted into an unidentified oxidised sesquiterpene (**10**), whose EI mass spectrum also shows a neutral loss of ethylene by RDA reaction ( $m/z = 192$ , Fig. 7B). Notably, directly next to the gene for the IWS a gene for a cytochrome P450 (CYP450) is observed in the genome of *S. lincolnensis* (accession number WP\_067429392, gene locus tag SLINC\_RS09700), which further supports the hypothesis that **10** is derived from **1** by oxidation. Further research will be required to verify **10** as an oxidation product of **1** and the role of the CYP450 in its formation.

## Conclusions

In summary, we have characterised a sesquiterpene synthase from *Streptomyces lincolnensis* that makes the new natural product isoishwarane. The enzyme mechanism was deeply studied by isotopic labelling experiments giving evidence for all elementary steps along the cationic cyclisation cascade, and through site-directed mutagenesis showing that some of the observed sequence deviations in otherwise highly conserved motifs are relevant for the formation of the unusual skeleton of isoishwarane. Also, a newly identified highly conserved



residue, an Asp located 50 positions upstream of the pyrophosphate sensor, is altered to a  $^{126}\text{N}$  in IWS, and here a “sequence correction” leads to a moderately improved production of **1**. A BLAST search revealed the presence of closely related homologs of IWS in five other genome sequenced streptomycetes, and in all five homologs a  $^{126}\text{D}$  residue is naturally present (Fig. S20†). These enzymes likely also convert FPP into **1**, but further research is required to experimentally confirm their functions, as it is for the characterisation of the CYP450 that is genetically clustered with IWS.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft (DI1536/7-2). We thank Christina Braun for assistance in the experimental work.

## References

- 1 D. E. Cane, J.-K. Sohng, C. R. Lamberson, S. M. Rudnicki, Z. Wu, M. D. Lloyd, J. S. Oliver and R. B. Hubbard, Pentalenene Synthase. Purification, Molecular Cloning, Sequencing, and High-Level Expression in *Escherichia coli* of a Terpenoid Cyclase from *Streptomyces* UC5319, *Biochemistry*, 1994, **33**, 5846.
- 2 X. Lin, R. Hopson and D. E. Cane, Genome Mining in *Streptomyces coelicolor*: Molecular Cloning and Characterization of a New Sesquiterpene Synthase, *J. Am. Chem. Soc.*, 2006, **128**, 6022.
- 3 J. Rinkel and J. S. Dickschat, Stereochemical investigations on the biosynthesis of achiral (Z)- $\gamma$ -bisabolene in *Cryptosporangium arzum*, *Beilstein J. Org. Chem.*, 2019, **15**, 789.
- 4 S. Garms, T. G. Köllner and W. Boland, A Multiproduct Terpene Synthase from *Medicago truncatula* Generates Cadalane Sesquiterpenes via Two Different Mechanisms, *J. Org. Chem.*, 2010, **75**, 5590.
- 5 L. S. Vedula, Y. Zhao, R. M. Coates, T. Koyama, D. E. Cane and D. W. Christianson, Exploring biosynthetic diversity with trichodiene synthase, *Arch. Biochem. Biophys.*, 2007, **466**, 260.
- 6 J. Rinkel and J. S. Dickschat, Mechanistic investigations on multiproduct  $\beta$ -himachalene synthase from *Cryptosporangium arzum*, *Beilstein J. Org. Chem.*, 2019, **15**, 1008.
- 7 Y. Yamada, T. Kuzuyama, M. Komatsu, K. Shin-ya, S. Omura, D. E. Cane and H. Ikeda, Terpene synthases are widely distributed in bacteria, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 857.
- 8 J. S. Dickschat, Bacterial Terpene Cyclases, *Nat. Prod. Rep.*, 2016, **33**, 87.
- 9 J. S. Dickschat, Bacterial Diterpene Biosynthesis, *Angew. Chem., Int. Ed.*, 2019, **58**, 15964.
- 10 D. W. Christianson, Structural and Chemical Biology of Terpenoid Cyclases, *Chem. Rev.*, 2017, **117**, 11570.
- 11 L. Martin-Sanchez, K. S. Singh, M. Avalos, G. P. van Wezel, J. S. Dickschat and P. Garbeva, Phylogenomic analyses and distribution of terpene synthases among *Streptomyces*, *Beilstein J. Org. Chem.*, 2019, **15**, 1181.
- 12 J. Jiang, X. He and D. E. Cane, Biosynthesis of the earthy odorant geosmin by a bifunctional *Streptomyces coelicolor* enzyme, *Nat. Chem. Biol.*, 2007, **3**, 711.
- 13 M. Komatsu, M. Tsuda, S. Omura, H. Oikawa and H. Ikeda, Identification and functional analysis of genes controlling biosynthesis of 2-methylisoborneol, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 7422.
- 14 C.-M. Wang and D. E. Cane, Biochemistry and Molecular Genetics of the Biosynthesis of the Earthy Odorant Methylisoborneol in *Streptomyces coelicolor*, *J. Am. Chem. Soc.*, 2008, **130**, 8908.
- 15 J. S. Dickschat, K. A. K. Pahirulzaman, P. Rabe and T. Klapschinski, An Improved Technique for the Rapid Chemical Characterisation of Bacterial Terpene Cyclases, *ChemBioChem*, 2014, **15**, 810.
- 16 A. J. Weinheimer, W. W. Youngblood, P. H. Washecheck, T. K. B. Karns and L. S. Ciereszko, Isolation of the elusive (–)-germacrene A from the gorgonian *Eunicea mamosa*. Chemistry of coelenterates, *Tetrahedron Lett.*, 1970, **11**, 497.
- 17 J. A. Faraldos, S. Wu, J. Chappell and R. M. Coates, Conformational analysis of (+)-germacrene A by variable-temperature NMR and NOE spectroscopy, *Tetrahedron*, 2007, **63**, 7733.
- 18 J. Rinkel and J. S. Dickschat, Addressing the Chemistry of Germacrene A by Isotope Labeling Experiments, *Org. Lett.*, 2019, **21**, 2426.
- 19 R. P. Adams, *Identification of Essential Oil Components by Gas Chromatography/Mass Spectrometry*, Allured Business Media, Carol Stream, Illinois, USA, 4th edn, 2009.
- 20 H. Xu and J. S. Dickschat, Germacrene A - A Central Intermediate in Sesquiterpene Biosynthesis, *Chem. – Eur. J.*, 2020, **26**, 17318–17341.
- 21 L. Lauterbach, J. Rinkel and J. S. Dickschat, Two Bacterial Diterpene Synthases from *Allokutzneria albata* for Bonnadiene and for Phomopsene and Allokutznerene, *Angew. Chem., Int. Ed.*, 2018, **57**, 8280.
- 22 F. M. Hahn, A. P. Hurlburt and C. D. Poulter, *Escherichia coli* Open Reading Frame 696 Is idi, a Nonessential Gene Encoding Isopentenyl Diphosphate Isomerase, *J. Bacteriol.*, 1999, **181**, 4499.
- 23 P. Rabe, J. Rinkel, B. Nubbemeyer, T. G. Köllner, F. Chen and J. S. Dickschat, Terpene Cyclases from Social Amoebae, *Angew. Chem., Int. Ed.*, 2016, **55**, 15420.
- 24 J. W. Cornforth, R. H. Cornforth, C. Donninger and G. Popjak, Studies on the biosynthesis of cholesterol XIX. Steric course of hydrogen eliminations and of C-C bond

- formations in squalene biosynthesis, *Proc. R. Soc. London, Ser. B*, 1966, **163**, 492.
- 25 J. W. Cornforth, R. H. Cornforth, G. Popjak and L. Yengoyan, Studies on the Biosynthesis of Cholesterol, *J. Biol. Chem.*, 1966, **241**, 3970.
- 26 R. B. Kelly and J. Zamecnik, Total Synthesis of Isoishwarane: Structures of the Sesquiterpenoids Ishwarane and Ishwarone, *J. Chem. Soc., Chem. Commun.*, 1970, 1102.
- 27 R. B. Kelly, J. Zamecnik and B. A. Becket, Total Synthesis of ( $\pm$ )-Ishwarane, a Tetracyclic Sesquiterpenoid, *Can. J. Chem.*, 1972, **50**, 3455.
- 28 H. Fuhrer, A. K. Ganguly, K. W. Gopinath, T. R. Govindachari, K. Nagarajan, B. R. Pai and P. C. Parthasarathy, Ishwarone, *Tetrahedron*, 1970, **26**, 2371.
- 29 U. S. K. Rao, B. L. Manjunath and K. N. Menon, Chemical examination of the roots of *Aristolochia indica*, Linn. II. The essential oil, *J. Indian Chem. Soc.*, 1935, **12**, 494.
- 30 T. R. Govindachari, K. Nagarajan and P. C. Parthasarathy, Absolute stereochemistry of ishwarone, *J. Chem. Soc. D*, 1969, 823.
- 31 I. Burkhardt, N. Kreuzenbeck, C. Beemelmans and J. S. Dickschat, Mechanistic Characterization of Three Sesquiterpene Synthases from the Termite-Associated Fungus *Termitomyces*, *Org. Biomol. Chem.*, 2019, **17**, 3348.
- 32 G. L. K. Hunter and W. B. Brogden, Conversion of valencene to nootkatone, *J. Food Sci.*, 1965, **30**, 876.
- 33 W. D. MacLeod, The constitution of nootkatone, nootkatene and valencene, *Tetrahedron Lett.*, 1965, **6**, 4779.
- 34 J. Lückner, P. Bowen and J. Bohlmann, *Vitis vinifera* terpene cyclases: functional identification of two sesquiterpene synthase cDNAs encoding (+)-valencene synthase and (-)-germacrene D synthase and expression of mono- and sesquiterpene synthases in grapevine flowers and berries, *Phytochemistry*, 2004, **65**, 2649.
- 35 P. Rabe, T. Schmitz and J. S. Dickschat, Mechanistic Investigations on Six Bacterial Terpene Cyclases, *Beilstein J. Org. Chem.*, 2016, **12**, 1839.
- 36 L. Ding, H. Goerls, K. Dornblut, W. Lin, A. Maier, H.-H. Fiebig and C. Hertweck, Bacaryolanes A–C, Rare Bacterial Caryolanes from a Mangrove Endophyte, *J. Nat. Prod.*, 2015, **78**, 2963.
- 37 P. Rabe, L. Barra, J. Rinkel, R. Riclea, C. A. Citron, T. A. Klapschinski, A. Janusko and J. S. Dickschat, Conformational Analysis, Thermal Rearrangement and EI-MS-Fragmentation Mechanism of (1(10)E,4E,6S,7R)-Germacradien-6-ol by <sup>13</sup>C-Labeling Experiments, *Angew. Chem., Int. Ed.*, 2015, **54**, 13448.
- 38 J. A. Faraldos, S. Wu, J. Chappell and R. M. Coates, Doubly Deuterium-Labeled Patchouli Alcohol from Cyclization of Singly Labeled [2-<sup>2</sup>H<sub>1</sub>]Farnesyl Diphosphate Catalyzed by Recombinant Patchoulol Synthase, *J. Am. Chem. Soc.*, 2010, **132**, 2998.
- 39 G. Bian, J. Rinkel, Z. Wang, L. Lauterbach, A. Hou, Y. Yuan, Z. Deng, T. Liu and J. S. Dickschat, A Clade II-D Fungal Chimeric Diterpene Synthase from *Colletotrichum gloeosporioides* Making Dolasta-1(15),8-diene, *Angew. Chem., Int. Ed.*, 2018, **57**, 15887.
- 40 T. A. Klapschinski, P. Rabe and J. S. Dickschat, Pristinol, a Sesquiterpene Alcohol with Unprecedented Skeleton from *Streptomyces pristinaespiralis*, *Angew. Chem., Int. Ed.*, 2016, **55**, 10141.
- 41 P. Baer, P. Rabe, K. Fischer, C. A. Citron, T. A. Klapschinski, M. Groll and J. S. Dickschat, Induced Fit Mechanism in Class I Terpene Cyclases, *Angew. Chem., Int. Ed.*, 2014, **53**, 7652.
- 42 J. A. Aaron, X. Lin, D. E. Cane and D. W. Christianson, Structure of Epi-Isozizaene Synthase from *Streptomyces coelicolor*, A3(2), a Platform for New Terpenoid Cyclization Templates, *Biochemistry*, 2010, **49**, 1787.
- 43 C. A. Lesburg, G. Zhai, D. E. Cane and D. W. Christianson, Crystal Structure of Pentalenene Synthase: Mechanistic Insights on Terpenoid Cyclization Reactions in Biology, *Science*, 1997, **277**, 1820.
- 44 M. Seemann, G. Z. Zhai, J. W. de Kraker, C. M. Paschall, D. W. Christianson and D. E. Cane, Pentalenene Synthase. Analysis of Active Site Residues by Site-Directed Mutagenesis, *J. Am. Chem. Soc.*, 2002, **124**, 7681.
- 45 D. S. Weinberg and C. Djerassi, Mass Spectrometry in Structural and Stereochemical Problems. LXXXVIII. Rearrangements of Simple Terpenes on Electron Impact, *J. Org. Chem.*, 1966, **31**, 115.
- 46 J. Karliner and C. Djerassi, Terpenoids. LVII. Mass Spectral and Nuclear Magnetic Resonance Studies of Pentacyclic Triterpene Hydrocarbons, *J. Org. Chem.*, 1966, **31**, 1945.
- 47 R. R. Muccino and C. Djerassi, Mass Spectrometry in Structural and Stereochemical Problems. CCXL. The Effect of a 1401-Angular Methyl Group upon the Mass Spectral Fragmentation of 11- and 7-Keto Steroids, *J. Am. Chem. Soc.*, 1973, **95**, 8726.
- 48 P. Rabe, T. A. Klapschinski and J. S. Dickschat, Position-specific mass shift analysis: a systematic method to investigate the EI-MS fragmentation mechanism of epi-isozizaene, *ChemBioChem*, 2016, **17**, 1333.
- 49 A. Hou and J. S. Dickschat, On the mass spectrometric fragmentations of the bacterial sesterterpenes sesterterpenes A – C, *Beilstein J. Org. Chem.*, 2020, **16**, 2807.
- 50 K. Grob and F. Zürcher, Stripping of trace organic substances from water, Equipment and procedure, *J. Chromatogr.*, 1976, **117**, 285.
- 51 S. Schulz and J. S. Dickschat, Bacterial Volatiles: The Smell of Small Organisms, *Nat. Prod. Rep.*, 2007, **24**, 814.
- 52 K. Wilkins and C. Schöller, Volatile Organic Metabolites from Selected *Streptomyces* Strains, *Actinomycetologica*, 2009, **23**, 27.
- 53 P. Rabe, C. A. Citron and J. S. Dickschat, Volatile Terpenes from Actinomycetes - A Biosynthetic Study Correlating Chemical Analyses to Genome Data, *ChemBioChem*, 2013, **14**, 2345.

## Appendix H

**The Sesquiterpene Synthase PtTPS5 Produces (1*S*,5*S*,7*R*,10*R*)-Guaia-4(15)-en-11-ol and (1*S*,7*R*,10*R*)-Guaia-4-en-11-ol in Oomycete-Infected Poplar Roots**

*Molecules* **2021**, *26*, 555

DOI: [10.3390/molecules26030555](https://doi.org/10.3390/molecules26030555)

Article

# The Sesquiterpene Synthase PtTPS5 Produces (1S,5S,7R,10R)-Guaia-4(15)-en-11-ol and (1S,7R,10R)-Guaia-4-en-11-ol in Oomycete-Infected Poplar Roots

Nathalie D. Lackus <sup>1,†</sup>, Jennifer Morawetz <sup>1,†</sup>, Houchao Xu <sup>2</sup>, Jonathan Gershenzon <sup>1</sup> , Jeroen S. Dickschat <sup>2</sup> and Tobias G. Köllner <sup>1,\*</sup> 

<sup>1</sup> Max Planck Institute for Chemical Ecology, Hans-Knöll-Strasse 8, 07745 Jena, Germany; nlackus@ice.mpg.de (N.D.L.); jenny.morawetz@gmail.com (J.M.); gershenzon@ice.mpg.de (J.G.)

<sup>2</sup> Kekulé-Institute of Organic Chemistry and Biochemistry, University of Bonn, Gerhard-Domagk-Strasse 1, 53121 Bonn, Germany; houchao.xu@uni-bonn.de (H.X.); dickschat@uni-bonn.de (J.S.D.)

\* Correspondence: koellner@ice.mpg.de

† These authors contributed equally to this work.

**Abstract:** Pathogen infection often leads to the enhanced formation of specialized plant metabolites that act as defensive barriers against microbial attackers. In this study, we investigated the formation of potential defense compounds in roots of the Western balsam poplar (*Populus trichocarpa*) upon infection with the generalist root pathogen *Phytophthora cactorum* (Oomycetes). *P. cactorum* infection led to an induced accumulation of terpenes, aromatic compounds, and fatty acids in poplar roots. Transcriptome analysis of uninfected and *P. cactorum*-infected roots revealed a terpene synthase gene *PtTPS5* that was significantly induced upon pathogen infection. *PtTPS5* had been previously reported as a sesquiterpene synthase producing two unidentified sesquiterpene alcohols as major products and hedyacryol as a minor product. Using heterologous expression in *Escherichia coli*, enzyme assays with deuterium-labeled substrates, and NMR analysis of reaction products, we could identify the major *PtTPS5* products as (1S,5S,7R,10R)-guaia-4(15)-en-11-ol and (1S,7R,10R)-guaia-4-en-11-ol, with the former being a novel compound. The transcript accumulation of *PtTPS5* in uninfected and *P. cactorum*-infected poplar roots matched the accumulation of (1S,5S,7R,10R)-guaia-4(15)-en-11-ol, (1S,7R,10R)-guaia-4-en-11-ol, and hedyacryol in this tissue, suggesting that *PtTPS5* likely contributes to the pathogen-induced formation of these compounds in planta.

**Keywords:** sesquiterpene synthase; *Populus trichocarpa*; oomycete; *Phytophthora cactorum*; plant defense



**Citation:** Lackus, N.D.; Morawetz, J.; Xu, H.; Gershenzon, J.; Dickschat, J.S.; Köllner, T.G. The Sesquiterpene Synthase PtTPS5 Produces (1S,5S,7R,10R)-Guaia-4(15)-en-11-ol and (1S,7R,10R)-Guaia-4-en-11-ol in Oomycete-Infected Poplar Roots. *Molecules* **2021**, *26*, 555. <https://doi.org/10.3390/molecules26030555>

Academic Editor: Valeria

Patricia Sülsen

Received: 14 December 2020

Accepted: 19 January 2021

Published: 21 January 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Plants are constantly under attack from a multitude of pests, including pathogens and herbivores. Such biotic stresses often induce the formation of specialized plant metabolites that play major roles in plant defense. Terpenoids represent the largest class of natural compounds, and to date, more than 200,000 terpenoids are known, of which ~40,000 can be produced by plants [1]. Beside a few roles in primary metabolism and physiology, most plant terpenes function as specialized metabolites in processes such as plant signaling and defense. Volatile mono- and sesquiterpenes, for example, have been described as repellants for herbivores or attractants for beneficial insects and animals e.g., [2–4]. Non-volatile terpenoids, however, can act as phytoalexins and protect the plant against pathogen infection by inhibiting the growth and/or development of the attacking pathogen [5]. The sesquiterpene-derived zealexins and the diterpene-derived kauralexins in the grasses are well known examples for antimicrobial and locally accumulating plant terpenoids that are produced in response to pathogen attack [6,7].

The biosynthesis of terpenes starts with the formation of isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP), which represent the C<sub>5</sub> building blocks for all terpenes. IPP and DMAPP can be condensed by isopentenyl diphosphate synthases (IDS) to form a variety of prenyl diphosphates with various chain lengths, including geranyl diphosphate (GPP, C<sub>10</sub>), (*E,E*)-farnesyl diphosphate (FPP, C<sub>15</sub>), and (*E,E,E*)-geranylgeranyl diphosphate (GGPP, C<sub>20</sub>). The prenyl diphosphates are substrates for terpene synthases (TPS), which catalyze the formation of the basic mono-(C<sub>10</sub>), sesqui-(C<sub>15</sub>), and diterpene (C<sub>20</sub>) skeletons [8]. The terpenes formed can be stored in the plant tissue or released as volatiles. Additionally, terpenes can act as substrates for modifying enzymes such as cytochrome P450 monooxygenases, *O*-methyltransferases, and acyltransferases [8–10].

In recent years, we investigated the formation of defense terpenes in the model tree species Western balsam poplar (*Populus trichocarpa*). Nineteen out of the 38 TPS genes found in the *P. trichocarpa* genome and three IDS genes involved in GPP and FPP formation have been cloned and characterized so far [11–15]. Most of these genes are significantly upregulated upon leaf or root herbivory, indicating that their terpene products are involved in plant defense against insect herbivores. However, whether poplar terpenes can also be formed as potential phytoalexins in response to pathogen attack is unclear. The aim of this study was to investigate the formation of defense compounds including terpenes in *P. trichocarpa* roots upon infection with a plant pathogen. The root rot-causing hemibiotrophic generalist oomycete *Phytophthora cactorum* was selected as a model organism because of its broad host specificity and economic importance. It can infect more than 200 plant species, including important crops such as apple trees and strawberries or ornamentals such as orchids. Transcriptome sequencing and RT-qPCR analysis revealed a sesquiterpene synthase gene *PtTPS5*, which was highly expressed in *P. cactorum*-infected roots but not in non-infected control roots. Enzyme assays with recombinant PtTPS5 and (*E,E*)-FPP as substrate and subsequent NMR analysis of TPS reaction products allowed the identification of two sesquiterpene alcohols that also accumulated in infected poplar roots. We propose that the PtTPS5 sesquiterpenes or their potential conversion products function as a defensive barrier against pathogen infection in poplar roots.

## 2. Results

### 2.1. *P. cactorum* Infection Induces the Accumulation of Terpenes, Aromatic Compounds, and Fatty Acids in *P. trichocarpa* Roots

To investigate the formation of potential defense compounds upon pathogen infection in poplar roots, young *P. trichocarpa* trees were grown in liquid medium and inoculated with a zoospore suspension of the generalist oomycete *P. cactorum*. Roots were harvested five days after inoculation, extracted with hexane and the extracts were analyzed using gas chromatography-mass spectrometry (GC-MS). Beside traces of the monoterpenes limonene and 1,8-cineole, the monoterpene alcohol  $\alpha$ -terpineol, the sesquiterpene alcohol elemol and two so far unidentified sesquiterpene alcohols were detected. Elemol most likely represents a rearrangement product of hedycaryol formed during GC-MS analysis. In general, germacrene sesquiterpenoids such as hedycaryol or germacrene A are well known to undergo thermal Cope rearrangements to elemol or  $\beta$ -elemene, respectively [16]. Thus, the thermal formation of elemol from hedycaryol under the conditions of the GC analysis is more likely than a direct enzymatic formation, which has never been described and would be difficult to understand mechanistically. The two unidentified sesquiterpene alcohols were later identified in this study as (1*S*,5*S*,7*R*,10*R*)-guaia-4(15)-en-11-ol and (1*S*,7*R*,10*R*)-guaia-4-en-11-ol (see Section 2.3). While limonene and 1,8-cineole could not be quantified due to low amounts and partial overlap with other peaks,  $\alpha$ -terpineol, elemol, and the two unidentified sesquiterpene alcohols showed a significantly higher accumulation in *P. cactorum*-infected roots compared to uninfected control roots (Table 1, Supplemental Figure S1). *P. cactorum* mycelium grown in liquid poplar growth medium in the absence of poplar roots showed no terpene accumulation (Supplemental Figure S2), suggesting that the terpenes detected in *P. cactorum*-infected roots were produced by the plant and not the oomycete.

**Table 1.** Compounds detected in hexane extracts made from uninfected and *Phytophthora cactorum*-infected *Populus trichocarpa* roots. Means and SE in  $\mu\text{g/g}$  fresh weight are shown ( $n = 7-8$ ).

Compound	Uninfected Roots	Infected Roots	t-Value/T-Value	p-Value
<b>Aromatic compounds</b>				
Benzylalcohol #	0.22 $\pm$ 0.20	0.77 $\pm$ 0.20	48.00 (WR)	0.038 *
Salicylaldehyde #	4.02 $\pm$ 1.00	4.93 $\pm$ 1.79	0.02 (ST)	0.988
2-Phenylethanol #	0.11 $\pm$ 0.04	0.74 $\pm$ 0.23	39.00 (WR)	<0.001 ***
Benzyl salicylate	n.q.	n.q.	-	-
<b>Terpenes</b>				
Limonene #	n.q.	n.q.	-	-
1,8-Cineole #	n.q.	n.q.	-	-
$\alpha$ -Terpineol #	0.10 $\pm$ 0.03	0.44 $\pm$ 0.05	5.26 (ST)	<0.001 ***
Elemol	0.01 $\pm$ 0.01	0.27 $\pm$ 0.09	40.50 (WR)	0.002 **
Guaia-4(15)-en-11-ol # + Guaia-4-en-11-ol #	0.04 $\pm$ 0.04	4.31 $\pm$ 1.35	36.00 (WR)	<0.001 ***
<b>Fatty acids/aldehydes</b>				
(E)-4-Nonenal	n.q.	n.q.	-	-
Myristaldehyde	3.16 $\pm$ 0.34	8.53 $\pm$ 2.17	3.13 (ST)	0.007 **
Myristic acid #	0.40 $\pm$ 0.08	6.13 $\pm$ 1.59	8.19 (ST)	<0.001 ***
Pentadecanoic acid #	1.51 $\pm$ 0.30	4.49 $\pm$ 0.51	38.00 (WR)	<0.001 ***
Palmitic acid #	24.53 $\pm$ 3.57	72.04 $\pm$ 6.13	36.00 (WR)	<0.001 ***
Oleic acid #	5.04 $\pm$ 0.92	7.34 $\pm$ 1.09	1.51 (ST)	0.154
Stearic acid #	1.65 $\pm$ 0.19	5.29 $\pm$ 0.71	6.39 (ST)	<0.001 ***
<b>Others</b>				
1-Hexanol	0.14 $\pm$ 0.01	0.17 $\pm$ 0.01	12.00 (WR)	0.038 *
Unidentified compound	traces	1.64 $\pm$ 0.88	40.00 (WR)	0.002 **

Asterisks indicate statistical significance between uninfected roots and infected roots as assessed by Student's *t*-test (ST) or Wilcoxon rank sum test (WR) (\*,  $p \leq 0.05$ ; \*\*,  $p \leq 0.01$ ; \*\*\*,  $p \leq 0.001$ ). n.q., not quantified due to trace amounts or incomplete separation. Compounds marked with # were identified using authentic standards.

In addition to the terpenes, a number of aromatic compounds including benzylalcohol, salicylaldehyde, 2-phenylethanol, and benzyl salicylate, some fatty acids, and the fatty acid aldehyde myristaldehyde could be detected in the root hexane extracts (Table 1). Two of the aromatic compounds namely benzylalcohol and 2-phenylethanol, almost all fatty acids, and myristaldehyde were significantly upregulated upon oomycete infection. With the exception of myristic acid, all fatty acids also occurred in hexane extracts made from *P. cactorum* mycelium grown in the absence of poplar roots (Supplemental Figure S2).

Salicinoids, a group of salicylalcohol-derived glucosides, are major defense compounds in the Salicaceae (reviewed in Böckler et al. [17]). To test whether salicinoid levels were influenced by the *P. cactorum* treatment, root material was extracted with methanol and the extracts were analyzed using high performance liquid chromatography (HPLC)-UV and liquid chromatography-tandem mass spectrometry (LC-MS/MS). While the accumulation of most of the measured salicinoids including salicin, salirepin, salirepin-7-sulfate, salicortin, tremulacin, and homaloside D was not influenced by the oomycete treatment, salicin-7-sulfate showed a small but significant induction upon pathogen infection (Table 2).

**Table 2.** Salicinoids detected in methanol extracts made from uninfected and *Phytophthora cactorum*-infected *Populus trichocarpa* roots. Means and SE in  $\mu\text{g/g}$  fresh weight are shown ( $n = 7\text{--}8$ ).

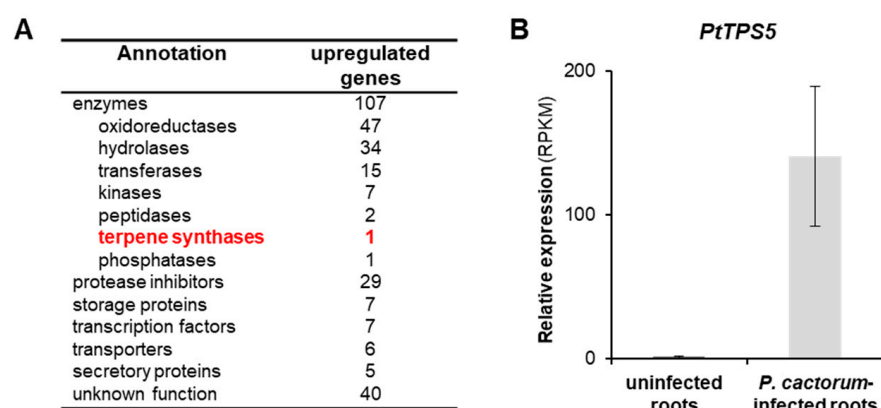
Compound	Uninfected Roots	Infected Roots	t-Value/T-Value	p-Value
Salicin	41.87 $\pm$ 16.03	77.22 $\pm$ 34.06	19.00 (WR)	0.336
Salicin-7-sulfate	2.08 $\pm$ 0.26	3.39 $\pm$ 0.39	2.52 (ST)	0.026 *
Salirepin	14.50 $\pm$ 2.39	20.62 $\pm$ 2.79	1.53 (ST)	0.151
Salirepin-7-sulfate	0.30 $\pm$ 0.03	0.42 $\pm$ 0.05	1.77 (ST)	0.1
Salicortin	368.31 $\pm$ 172.48	204.06 $\pm$ 70.32	2.25 (ST)	0.056
Tremulacin	3.16 $\pm$ 1.27	2.49 $\pm$ 1.98	0.28 (ST)	0.785
Homaloside D	16.77 $\pm$ 9.28	7.78 $\pm$ 3.04	33.00 (WR)	0.596

Asterisks indicate statistical significance between uninfected roots and infected roots as assessed by Student's *t*-test (ST) or Wilcoxon rank sum test (WR) (\*,  $p \leq 0.05$ ).

Infection of *P. trichocarpa* roots by *P. cactorum* was verified by measuring the transcript accumulation of the *Phytophthora*-specific Ras-related protein *Ypt1* [18] in the root material using RT-qPCR. *Ypt1* transcripts could be detected in *P. cactorum*-infected roots but not in uninfected control roots (Supplemental Figure S1), indicating a successful infection of the plant.

## 2.2. Transcriptome Analysis of Infected and Non-Infected Poplar Roots Revealed a Sesquiterpene Synthase Gene *PtTPS5* that Is highly Induced upon *P. cactorum* Infection

In order to identify genes involved in the *P. cactorum*-induced plant defense response, especially in terpene formation, we sequenced and analyzed the transcriptomes of infected and non-infected *P. trichocarpa* roots. Mapping the sequence reads onto the *P. trichocarpa* gene set revealed 201 genes that were significantly upregulated (fold change > 5) upon *P. cactorum* infection (Figure 1A, Supplemental Table S1). Among these genes, 107 encoded enzymes, including a highly upregulated terpene synthase (*PtTPS5*, Potri.005g095500). *PtTPS5* has recently been reported as sesquiterpene synthase producing unidentified sesquiterpene alcohols as major products and hedycaryol as a minor product [12]. Notably, the relatively high RPKM values (average  $\sim 140$ ) for *PtTPS5* in *P. cactorum*-infected roots were comparable to those of a variety of protease inhibitor genes known to be involved in plant defense (Figure 1B, Supplemental Table S1).



**Figure 1.** Transcript accumulation of the sesquiterpene synthase gene *PtTPS5* is upregulated after *Phytophthora cactorum* infection in *Populus trichocarpa* roots. (A) RNAseq, subsequent read mapping, and EDGE (estimated degree of gene expression) analysis was performed to identify genes significantly higher expressed in *P. cactorum*-infected roots compared to uninfected control roots. Genes with a fold change > 5 (false discovery rate < 0.01%;  $n = 4$ ) were considered as upregulated. (B) Relative gene expression of *PtTPS5*. Means and SE of RPKM values are shown ( $n = 4$ ). EDGE test ( $p = 4.9 \times 10^{-19}$ , weighted difference = 0.000254923).

### 2.3. PtTPS5 Produces (1*S*,5*S*,7*R*,10*R*)-guaia-4(15)-en-11-ol and (1*S*,7*R*,10*R*)-guaia-4-en-11-ol as Major Products

To elucidate the structure of the unidentified PtTPS5 products, the gene was synthesized, cloned into the bacterial expression vector pET100/D-TOPO, and heterologously expressed in *Escherichia coli*. Purified recombinant protein was incubated with (*E,E*)-FPP as substrate. GC-MS analysis of the products revealed the formation of hedyaryol, detected as its Cope rearrangement product elemol, and two coeluting sesquiterpene alcohols that were purified by column chromatography, followed by structure elucidation through one- and two-dimensional NMR spectroscopy (Table 3), resulting in the structures of the new compound guaia-4(15)-en-11-ol (**1**) and the known guaia-4-en-11-ol (**2**) that was previously reported from *Bulnesia sarmientoi* (Figure 2A) [19]. A biosynthetic hypothesis for these sesquiterpene alcohols suggests formation proceeds by 1,10-cyclisation and capture with water to form hedyaryol (Figure 2B). Its reprotonation at C10 can initiate a second cyclisation, as frequently observed for germacrane-type sesquiterpenes [20], to give a guaiane skeleton, followed by deprotonations from C15 or C5 to yield guaia-4(15)-en-11-ol and guaia-4-en-11-ol, respectively. The absolute configuration of guaia-4(15)-en-11-ol was determined through chemical correlation using stereoselectively deuterated precursors. The enzymatic conversion of (*R*)- or (*S*)-(1-<sup>13</sup>C,1-<sup>2</sup>H)IPP [21] with isopentenyl diphosphate isomerase from *E. coli* [21], FPP synthase from *Streptomyces coelicolor* [22], and PtTPS5 resulted in an enantioselective deuteration at C2, C6, and C8 of guaia-4(15)-en-11-ol with known configuration (Figure 2C), because it is well known that prenyl diphosphates are elongated by IPP with inversion of configuration at C1 [23]. The additional <sup>13</sup>C-label allowed for a highly sensitive detection of HSQC signals for the bound hydrogens, while signals for the hydrogens substituted by deuterium vanished (Supplemental Figure S3A–C). The labelled carbons together with a full assignment of hydrogen signals by NOESY helped determine the relative orientation of the naturally present stereogenic centres in **1** and thus its absolute configuration as (1*S*,5*S*,7*R*,10*R*)-guaia-4(15)-en-11-ol. A second set of experiments was performed with DMAPP and (*E*)- or (*Z*)-(4-<sup>13</sup>C,4-<sup>2</sup>H)IPP [24] (Figure 2D), known to react with attack at C4 of IPP from the *Si* face under FPPS catalysis [23]. Further conversion by PtTPS5 resulted in the introduction of additional stereogenic probes at C3 and C9, and HSQC analysis gave consistent results regarding the absolute configuration of (1*S*,5*S*,7*R*,10*R*)-guaia-4(15)-en-11-ol (Supplemental Figure S3D–F). The absolute configurations of hedyaryol and (1*S*,7*R*,10*R*)-guaia-4-en-11-ol were assigned based on biosynthetic considerations (Figure 2B).

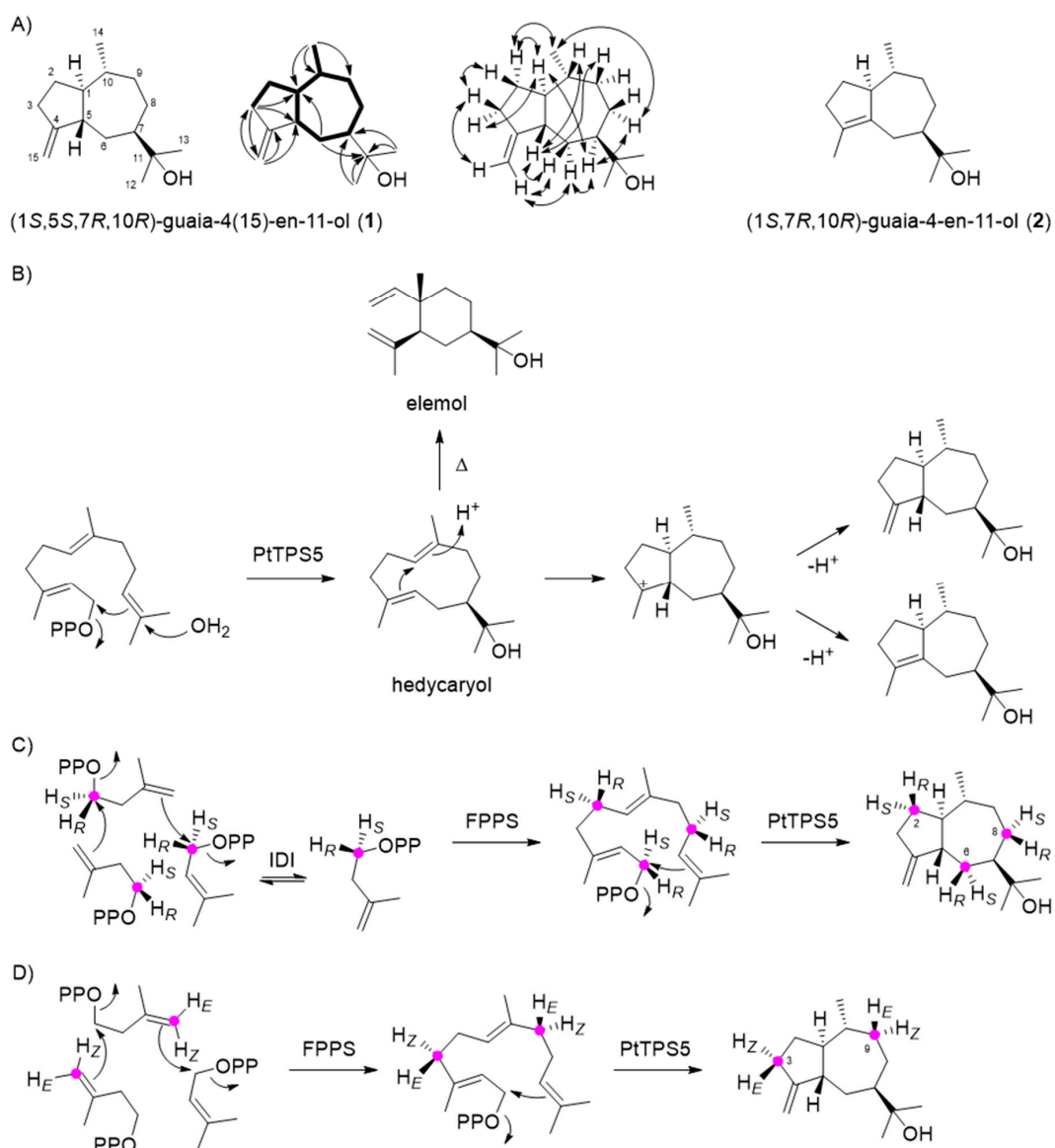
**Table 3.** NMR data of guaia-4(15)-en-11-ol and guaia-4-en-11-ol (isoguaiaol B). NMR data were recorded on a 700 MHz spectrometer in C<sub>6</sub>D<sub>6</sub> at 298 K. Coupling constants *J* are given in Hz and multiplicities are indicated by s = singlet, d = doublet, m = multiplet, br = broad.

C	Guaia-4(15)-en-11-ol ( <b>1</b> )		Guaia-4-en-11-ol ( <b>2</b> )	
	<sup>13</sup> C	<sup>1</sup> H	<sup>13</sup> C	<sup>1</sup> H
1	52.34 (CH)	1.18 (m)	55.60 (CH)	2.28 (m)
2	32.42 (CH <sub>2</sub> )	1.76 (m, H <sub>α</sub> ) 1.00 (m, H <sub>β</sub> )	30.42 (CH <sub>2</sub> )	1.92 (m) 1.51 (m)
3	32.96 (CH <sub>2</sub> )	2.32 (ddm, <i>J</i> = 15.9, 7.8, H <sub>β</sub> ) 2.19 (m, H <sub>α</sub> )	36.60 (CH <sub>2</sub> )	2.23 (m) 2.13 (m)
4	159.08 (C <sub>q</sub> )	–	131.74 (C <sub>q</sub> )	–
5	46.33 (CH)	2.15 (m)	138.62 (C <sub>q</sub> )	–
6	34.56 (CH <sub>2</sub> )	1.88 (ddd, <i>J</i> = 13.9, 9.2, 6.1, H <sub>β</sub> ) 1.54 (ddd, <i>J</i> = 13.9, 11.3, 8.0, H <sub>α</sub> )	30.37 (CH <sub>2</sub> )	2.67 (d, <i>J</i> = 15.2) 2.00 (m)
7	49.56 (CH)	1.41 (m)	49.30 (CH)	1.36 (m)
8	26.71 (CH <sub>2</sub> )	1.74 (m, H <sub>α</sub> ) 1.10 (m, H <sub>β</sub> )	31.51 (CH <sub>2</sub> )	1.90 (m) 1.00 (m)
9	40.29 (CH <sub>2</sub> )	1.77 (m, H <sub>β</sub> ) 0.96 (m, H <sub>α</sub> )	40.26 (CH <sub>2</sub> )	1.78 (m) 1.10 (m)
10	42.32 (CH)	1.06 (m)	39.37 (CH)	1.33 (m)



Table 3. Cont.

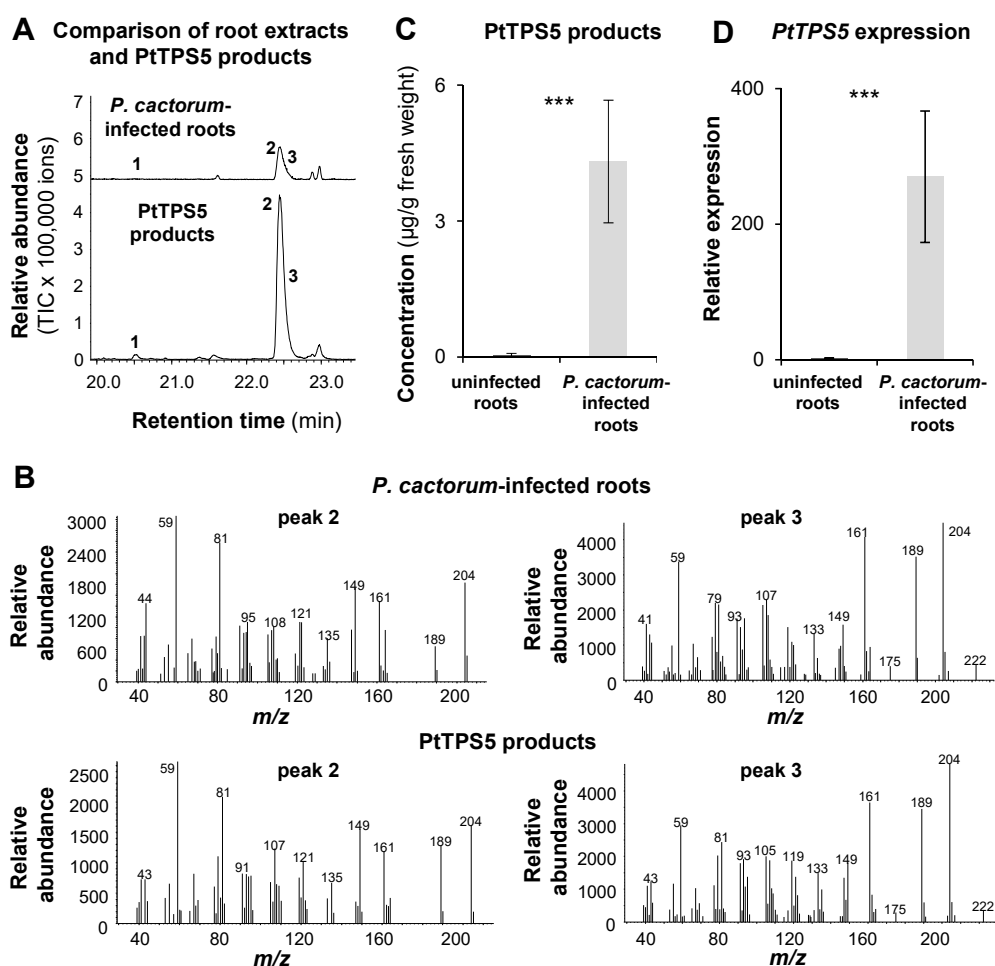
C	Guaia-4(15)-en-11-ol (1)		Guaia-4-en-11-ol (2)	
	$^{13}\text{C}$	$^1\text{H}$	$^{13}\text{C}$	$^1\text{H}$
11	73.26 (C <sub>q</sub> )	–	72.82 (C <sub>q</sub> )	–
12	27.75 (CH <sub>3</sub> )	1.03 (s)	26.89 (CH <sub>3</sub> )	1.03 (s)
13	25.78 (CH <sub>3</sub> )	1.01 (s)	26.66 (CH <sub>3</sub> )	1.02 (s)
14	21.58 (CH <sub>3</sub> )	0.87 (d, <i>J</i> = 6.5)	22.01 (CH <sub>3</sub> )	0.93 (d, <i>J</i> = 6.6)
15	104.24 (CH <sub>2</sub> )	5.00 (m, H <sub>E</sub> ) 4.91 (br, H <sub>Z</sub> )	14.60 (CH <sub>3</sub> )	1.63 (br s)



**Figure 2.** PtTSP5 produces (1*S*,5*S*,7*R*,10*R*)-guaia-4(15)-en-11-ol (1), (1*S*,7*R*,10*R*)-guaia-4-en-11-ol (2), and minor amounts of hedycaryol *in vitro*. (A) Structure elucidation (bold lines: H,H-COSY correlations, single-headed arrows: HMBC correlations, double-headed arrows: NOESY correlations). Carbon numbering is not systematic, but follows the FPP numbering. (B) Biosynthetic model for the cyclisation from FPP to the sesquiterpene alcohols and Cope rearrangement of hedycaryol to elemol under the thermal conditions of GC-MS analysis. Determination of the absolute configuration of (1*S*,5*S*,7*R*,10*R*)-guaia-4(15)-en-11-ol by enantioselective deuteration using (C) (*R*)- or (*S*)-(1- $^{13}\text{C}$ ,1- $^2\text{H}$ )IPP with IDI, FPPS and PtTSP5, and (D) (*E*)- or (*Z*)-(4- $^{13}\text{C}$ ,4- $^2\text{H}$ )IPP with FPPS and PtTSP5.

#### 2.4. The Accumulation of (1*S*,5*S*,7*R*,10*R*)-guaia-4(15)-en-11-ol, (1*S*,7*R*,10*R*)-guaia-4-en-11-ol, and Hedycaryol in *P. cactorum*-Infected and Non-Infected Roots Matches the Expression of PtTPS5

To figure out whether the two unidentified sesquiterpene alcohols detected in *P. cactorum*-infected poplar roots were identical to the PtTPS5 products (1*S*,5*S*,7*R*,10*R*)-guaia-4(15)-en-11-ol and (1*S*,7*R*,10*R*)-guaia-4-en-11-ol, we analyzed and compared hexane extracts prepared from a PtTPS5 enzyme assay and oomycete-infected root material using GC-MS. Although the two sesquiterpene alcohols could not be separated completely under the GC conditions we used in this experiment, the peaks of the PtTPS5 products and the two unidentified alcohols in the root extract had identical retention times and highly similar mass spectra (Figure 3A,B). Notably, the minor PtTPS5 product hedycaryol could also be detected as trace compound in the root extract. *PtTPS5* gene expression in uninfected and *P. cactorum*-infected *P. trichocarpa* roots measured by RT-qPCR showed an expression pattern nearly identical to the accumulation pattern of the PtTPS5 products measured in the same tissue (Figure 3C,D). This indicates that PtTPS5 likely produces (1*S*,5*S*,7*R*,10*R*)-guaia-4(15)-en-11-ol, (1*S*,7*R*,10*R*)-guaia-4-en-11-ol, and traces of hedycaryol in *P. cactorum*-infected *P. trichocarpa* roots.



**Figure 3.** Accumulation of (1*S*,5*S*,7*R*,10*R*)-guaia-4(15)-en-11-ol and (1*S*,7*R*,10*R*)-guaia-4-en-11-ol and gene expression of poplar *PtTPS5* in uninfected and *Phytophthora cactorum*-infected *Populus trichocarpa* roots. (A) Terpenes were extracted with hexane from pulverized root material or from an assay containing recombinant PtTPS5 and (*E,E*)-FPP and analyzed using GC-MS. 1, elemol; 2, (1*S*,5*S*,7*R*,10*R*)-guaia-4(15)-en-11-ol; 3, (1*S*,7*R*,10*R*)-guaia-4-en-11-ol. (B) Mass spectra of (1*S*,5*S*,7*R*,10*R*)-guaia-4(15)-en-11-ol and (1*S*,7*R*,10*R*)-guaia-4-en-11-ol found in *P. cactorum*-infected roots and in enzyme assays of PtTPS5 (peaks 2 and 3). (C) Accumulation of PtTPS5 products in *P. trichocarpa* roots. Means and SE are shown ( $n = 7-8$ ). Wilcoxon rank sum test ( $T = 36.00$ ,  $p < 0.001$ ). (D) Relative expression of *PtTPS5* determined by RT-qPCR. Means and SE are shown ( $n = 7-8$ ). Student's *t*-test ( $t = 7.626$ ,  $p < 0.001$ ). TIC = total ion count; \*\*\*,  $p \leq 0.001$ .

### 3. Discussion

In this study, we showed that infection of poplar roots by the generalist oomycete *P. cactorum* resulted in the induced accumulation of a number of potential defense compounds including terpenoids, aromatic compounds and fatty acids. Two of these compounds were exclusively produced in infected roots and could be identified as (1*S*,5*S*,7*R*,10*R*)-guaia-4(15)-en-11-ol and (1*S*,7*R*,10*R*)-guaia-4-en-11-ol, with the first one being a novel sesquiterpenoid (Table 1; Figures 2 and 3). A recently reported terpene synthase, PtTPS5 [12], was found to form both sesquiterpene alcohols as major products and minor amounts of hedyacryol *in vitro* (Figure 3). Since *P. trichocarpa* possesses no other terpene synthase with high similarity to PtTPS5 [12], and *PtTPS5* is the only *TPS* gene significantly induced upon *P. cactorum* infection in roots (Supplemental Table S1), we conclude that the pathogen-induced accumulation of (1*S*,5*S*,7*R*,10*R*)-guaia-4(15)-en-11-ol, (1*S*,7*R*,10*R*)-guaia-4-en-11-ol, and hedyacryol is likely due to PtTPS5 activity *in vivo*. Infection of strawberry (*Fragaria vesca*) roots with *P. cactorum* has been shown to induce massive changes in the transcriptome, including the upregulation of the complete mevalonate pathway, two FPP synthase genes, and four putative sesquiterpene synthase genes with similarity to germacene D synthase [25]. Moreover, Yadav and colleagues reported that the infection of *Medicago truncatula* roots with the oomycete *Aphanomyces euteiches* led also to the expression of a sesquiterpene synthase gene [26]. The encoded enzyme MtTPS10 was shown to produce a blend of sesquiterpenes with the alcohol himachalol as the major component. Down regulation of *MtTPS10* resulted in increased susceptibility to the oomycete and a mixture of isolated MtTPS10 products inhibited mycelial growth and *A. euteiches* zoospore germination. However, since himachalol could not be detected in *A. euteiches*-infected roots, MtTPS10 alcohols are likely converted to other terpenoids as speculated by the authors [26]. Indeed, conversion of sesquiterpenes into polar compounds such as aldehydes and acids upon pathogen infection has been described in a number of plants. Pathogen-infected maize, for example, produces the sesquiterpene hydrocarbon  $\beta$ -macrocarpene, which is further converted to antimicrobial sesquiterpene acids called zealexins [7,27]. Kauralexins, another group of terpene acid phytoalexins found in maize, are produced from the diterpene *ent*-kaurene [6], and the sesquiterpene  $\delta$ -cadinene acts as precursor for the formation of gossypol and other sesquiterpenoid phytoalexins in cotton [28,29]. In contrast to himachalol in infected *Medicago* roots, (1*S*,5*S*,7*R*,10*R*)-guaia-4(15)-en-11-ol and (1*S*,7*R*,10*R*)-guaia-4-en-11-ol accumulated in oomycete-infected poplar roots and thus could function as defense compounds themselves. However, considering the findings from the other plant systems described above, it is tempting to speculate that they might also be converted to other so far unknown antimicrobial defense compounds. Metabolism of terpenes often involves diverse hydroxylation and oxidation steps. Such reactions are in general catalyzed by cytochrome P450 monooxygenases or dioxygenases [10,30,31]. Our poplar transcriptome analysis revealed a number of putative P450 and dioxygenase genes that were strongly upregulated upon *P. cactorum* infection (Supplemental Table S1). Testing their enzymatic activity with PtTPS5 products as substrate will be a worthwhile aim for further studies.

Free fatty acids have been described to be involved in plant defense against various pathogens and herbivores [32–34]. They often act as signaling compounds or as precursors for signaling compounds [35], but can also directly impair the attacker [34]. Beside the two sesquiterpene alcohols (1*S*,5*S*,7*R*,10*R*)-guaia-4(15)-en-11-ol and (1*S*,7*R*,10*R*)-guaia-4-en-11-ol, we identified a number of fatty acids including myristic acid, pentadecanoic acid, palmitic acid, oleic acid, and stearic acid, that accumulated in substantial amounts in *P. cactorum*-infected roots (Table 1). Since pentadecanoic acid, palmitic acid, oleic acid, and stearic acid could also be detected in hexane extracts made from *P. cactorum* mycelium grown in the absence of roots (Supplemental Figure S2), it is likely that their increased accumulation in infected poplar roots is mainly caused by the oomycete itself. However, myristic acid was not found in *P. cactorum* mycelium and is produced by the poplar roots. Myristic acid has been shown to possess antimicrobial activity against diverse pathogenic

fungi [34] and might act as a defense against the oomycete *P. cactorum*. Moreover, the related myristaldehyde has been reported as main constituent of many antimicrobial essential oils [36] and its oomycete-induced upregulation also indicates a function in poplar defense against pathogens.

## 4. Materials and Methods

### 4.1. Biological Material

Western balsam poplar (*Populus trichocarpa*, clone Muhle-Larsen, P&P Baumschule, Eitelborn, Germany) trees were propagated from monoclonal stem cuttings and grown under summer conditions in the greenhouse (24 °C, 60% rel. humidity, 16 h/8 h light/dark cycle) in hydroculture medium until they reached about 0.15 m in height. The hydroculture medium contained 7.05 g NaNO<sub>3</sub>, 3.05 g Fertyl Basis 1 (Planta Düngemittel GmbH, Regenstauf, Germany), 1.36 g MgSO<sub>4</sub>, 0.04 g FeSO<sub>4</sub> · 7H<sub>2</sub>O, and 0.05 g Titriplex® V in a total volume of 5 L H<sub>2</sub>O.

*Phytophthora cactorum* (Oomycetes) was obtained from the Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures GmbH (Braunschweig, Germany). The generalist root pathogen was grown and sub-cultured via mycelial inoculation in petri dishes containing tomato juice medium. A 1.5 L quantity of medium contained 300 mL tomato juice ("Bio" quality from Netto supermarket), 4.5 g CaCO<sub>3</sub> (Roth, Karlsruhe, Germany), and 11.25 g agar-agar, filled to full volume with triple distilled water (adjusted to pH 7.2) at room temperature.

### 4.2. *Phytophthora Cactorum* Treatment

Prior to the onset of the experiment, *P. cactorum* was freshly sub-cultured from mycelium and incubated in the dark at 25 °C. After seven days, plates were washed with ddH<sub>2</sub>O and the suspension obtained contained the *P. cactorum* sporangia. The number of sporangia was determined with a counting chamber and adjusted to a concentration of  $3.78 \times 10^5$  sporangia per 50 mL poplar hydroculture medium. The sporangia solution was stored for 30 min at 4 °C to induce the release of the zoospores. Each poplar tree was either placed in clean 50 mL poplar hydroculture medium (control;  $n = 7$ ) or in 50 mL poplar hydroculture medium containing the above determined amount of *P. cactorum* sporangia (*P. cactorum*-infected;  $n = 8$ ). Poplar trees were further grown for five days under summer conditions as described above (Section 4.1). Poplar hydroculture medium (50 mL) containing the same amount of *P. cactorum* sporangia (*P. cactorum* mycelium;  $n = 4$ ) was cultivated for five days as described for the poplar trees. After five days of inoculation, poplar root material (average root weight of  $0.41 \text{ g} \pm 0.05$  (control) and  $0.38 \text{ g} \pm 0.06$  (*P. cactorum*-infected)) was harvested, immediately flash-frozen in liquid nitrogen, and stored at  $-80$  °C until further processing. The *P. cactorum* mycelium samples were centrifuged at  $15,000 \times g$  for 5 min, and the supernatant removed. The remaining mycelium was flash frozen in liquid nitrogen, and stored at  $-80$  °C until further processing.

### 4.3. Hexane Extraction of Root Tissue and GC-MS/GC-FID Analysis

To determine the accumulation of non-polar compounds in poplar roots, 100 mg of ground root powder was extracted in a GC glass vial with 400 µL hexane including 10 ng/µL nonyl acetate as an internal standard. The extracts were shaken for one hour at 900 rpm and incubated overnight at room temperature. After centrifugation for 10 min at  $5000 \times g$ , the supernatant was taken and subsequently analyzed via gas chromatography-mass spectrometry (GC-MS) and gas chromatography-flame ionization detection (GC-FID). The extraction of non-polar compounds from *P. cactorum* mycelium was performed as described above for the root tissue, except that 50 mg of the mycelium and 200 µL hexane were used.

Qualitative and quantitative analysis of non-polar compounds in (non-) infected *P. trichocarpa* roots and *P. cactorum* mycelium was conducted using a 6890 Series gas chromatograph (Agilent Technologies, Santa Clara, CA, USA) coupled to an Agilent 5973

quadrupole mass selective detector (interface temp, 270 °C; quadrupole temp, 150 °C; source temp, 230 °C; electron energy, 70 eV) or a flame ionization detector (FID) operated at 300 °C, respectively. The constituents of the hexane extracts were separated using a ZB5 column (Phenomenex, Aschaffenburg, Germany, 30 m × 0.25 mm × 0.25 µm) and He (MS) or H<sub>2</sub> (FID) as carrier gas. The sample (1 µL) was injected without split at an initial oven temperature of 45 °C. The temperature was held for 2 min and then increased to 280 °C with a gradient of 6 °C min<sup>-1</sup>, and then further increased to 300 °C with a gradient of 60 °C min<sup>-1</sup> and a hold of 2 min. Compounds were identified by comparing their retention times and mass spectra to those of authentic standards (Supplemental Tables S2 and S3), or to reference spectra in the Wiley and National Institute of Standards and Technology Libraries.

#### 4.4. Methanol Extraction of Root Tissue and HPLC-UV, LC-MS/MS Analysis of Methanol Extracts

Metabolites were extracted from 40 mg fresh plant material by adding 1 mL 100% methanol (MeOH) containing 0.8 mg/mL phenyl-β-D-glucopyranoside (Sigma Aldrich, St. Louis, MO, USA) and 40 ng/mL D<sub>6</sub>-abscisic acid (D<sub>6</sub>-ABA) (Santa Cruz Biotechnology, Dallas, TX, USA) as internal standards. Samples were shaken for 30 sec in a paint shaker (Scandex, Büdelsdorf, Germany) and afterwards for 30 min at 200 rpm on a horizontal shaker (IKA Labortechnik, Staufen, Germany). After centrifugation, the supernatants were split for high performance liquid chromatography (HPLC)-UV and liquid chromatography-tandem mass spectrometry (LC-MS/MS) measurements.

Salicinoid analysis and quantification was performed by HPLC-UV (200 nm) as described previously in Böckler et al. [37] for the compounds salicin, salicortin, tremulacin, and homaloside D, and for 6'-O-benzoylsalicortin as described in Lackner et al. [38]. Chromatographic separation was achieved on an Agilent 1100 Series LC system (Agilent Technologies), using an EC 250/4.6 Nucleodur Sphinx column (RP 5 µm, Macherey-Nagel, Düren, Germany), with water and acetonitrile as mobile phases A and B, respectively. The mobile phase flow rate was 1 mL/min. The elution profile is listed in Supplemental Table S4 as gradient A. Salicinoids were quantified relative to the signal of the internal standard phenyl-β-D-glucopyranoside, by applying experimentally determined response factors [37,38].

The compounds salirepin, salicin-7-sulfate, and salirepin-7-sulfate were analyzed and quantified by LC-MS/MS as follows and as previously described in Lackus et al. [39]. Chromatographic separation was achieved using an Agilent 1260 infinity II LC system (Agilent Technologies) equipped with a Zorbax Eclipse XDB-C18 column (50 × 4.6 mm, 1.8 µm, Agilent Technologies), using aqueous formic acid (0.05% (v/v)) and acetonitrile as mobile phases A and B, respectively. The mobile phase flow rate was 1.1 mL/min. The elution profile is listed in Supplemental Table S4 as gradient B. The column temperature was maintained at 20 °C. The LC system was coupled to a QTRAP 6500<sup>®</sup> tandem mass spectrometer (AB Sciex, Darmstadt, Germany) equipped with a turbospray ion source, operated in negative ionization mode. The ion spray voltage was maintained at -4500 eV and the turbo gas temperature was set at 700 °C. Nebulizing gas was set at 60 psi, curtain gas at 40 psi, heating gas at 60 psi and collision gas at medium level. Multiple reaction monitoring (MRM) was used to monitor analyte parent ion → product ion formation, and respective parameters are listed in Supplemental Table S5. Sulfated salicinoids and salirepin were quantified relative to the signal of the internal standard D<sub>6</sub>-ABA, by applying experimentally determined response factors [39]. Analyst 1.6.3 software (Applied Biosystems, Darmstadt, Germany) was used for data acquisition and processing.

#### 4.5. RNA Extraction and Reverse Transcription

Total RNA was isolated from frozen and ground plant material using the InviTrap Spin Plant RNA Kit (Invitek, Berlin, Germany) according to the manufacturer's instructions. RNA concentration was assessed using a spectrophotometer (NanoDrop 2000c, Thermo Fisher Scientific, Waltham, MA, USA). RNA was treated with DNaseI (Thermo Fisher

Scientific) prior to cDNA synthesis. Single-stranded cDNA was prepared from 1 µg of DNase-treated RNA using SuperScript<sup>TM</sup> III reverse transcriptase and oligo (dT<sub>12-18</sub>) primers (Invitrogen, Carlsbad, CA, USA).

#### 4.6. Heterologous Expression of PtTPS5 and Enzyme Assays

PtTPS5 was previously characterized by Irmisch et al. [12]. Based on its sequence deposited in GenBank with the accession number KF776503, PtTPS5 was synthesized and cloned into pET100/D-TOPO vector (Thermo Fisher Scientific). The *Escherichia coli* strain BL21 Star<sup>TM</sup> (DE3) (Thermo Fisher Scientific) was used for heterologous expression. The culture was grown at 37 °C, induced at an OD<sub>600</sub> = 0.6 with 1 mM IPTG, and subsequently placed at 18 °C and grown for another 20 h. The cells were collected by centrifugation and disrupted by a 4 × 20 s treatment with a sonicator (Bandelin UW2070, Berlin, Germany) in chilled extraction buffer (10 mM Tris-HCl (pH 7.5), 1 mM dithiothreitol, 10% (v/v) glycerol). Cell fragments were removed by centrifugation at 14,000 g and the supernatant was further processed via an Illustra NAP-5 gravity flow desalting column (GE Healthcare, Chicago, IL, USA) and eluted in extraction buffer.

Enzyme assays were performed in a Teflon-sealed, screw-capped 1 mL GC glass vial containing 50 µL of the heterologously expressed protein and 50 µL assay buffer containing 50 µM (E,E)-FPP substrate and 20 mM MgCl<sub>2</sub>. Assays were overlaid with 100 µL hexane and incubated for 60 min at 30 °C. One microliter of the hexane phase was injected into the GC-MS and the analysis was conducted using the same analytical parameters and equipment as described above for the analysis of poplar root hexane extracts. However, chromatographic separation was achieved with an initial oven temperature of 45 °C hold for 2 min, which was then increased to 180 °C with a gradient of 6 °C min<sup>-1</sup>, and then further increased to 300 °C with a gradient of 60 °C min<sup>-1</sup> and a hold of 2 min.

#### 4.7. RNA Sequencing and RT-qPCR Analysis

Total RNA was extracted from root material as described above, TruSeq RNA-compatible libraries were prepared, and PolyA enrichment was performed before sequencing eight transcriptomes of *P. trichocarpa*, four biological replicates (individual trees) each for the control and the oomycete treatments, on an IlluminaHiSeq 3000 sequencer (Max Planck Genome Centre, Cologne, Germany) with 45 Mio reads per library, 150 base pair, single end. Trimming of the obtained Illumina reads and mapping to the poplar gene model version 3.0 (<https://phytozome.jgi.doe.gov/pz/portal.html>) were performed with the program CLC Genomics Workbench (Qiagen Bioinformatics, Hilden, Germany) (mapping parameter: length fraction, 0.7; similarity fraction, 0.9; max number of hits, 25). Empirical analysis of digital gene expression (EDGE) implemented in the program CLC Genomics Workbench was used for gene expression analysis.

For RT-qPCR analysis, cDNA was prepared as described above and diluted 1:10 with water. Primers for gene expression analysis of PtTPS5 and *Ypt1* were used as described in Irmisch et al. [12] and Schena et al. [18], respectively. *Ubiquitin (UBQ)*, *actin*, *elongation factor 1 alpha (EF1α)*, *histone superfamily protein H3 (HIS)*, and *tubulin (TUB)* were tested as reference genes [40–42]. Primer sequences can be found in Supplemental Table S6. Comparison of ΔCq values and the corresponding standard deviation revealed *HIS* as the most suitable reference gene for expression analysis in *P. trichocarpa* samples (Supplemental Table S7). Gene expression analysis was performed with an initial incubation at 95 °C for 3 min followed by 40 cycles of amplification (95 °C for 10 s, 60 °C for 10 s). For all measurements, plate reads were taken at the end of the extension step of each cycle and data for the melting curves were recorded at the end of cycling from 60 °C to 95 °C. All samples were run on the same PCR machine (Bio-Rad CFX Connect<sup>TM</sup> Real-Time PCR Detection System (Bio-Rad Laboratories, Hercules, CA, USA)) in an optical 96-well plate, using Brilliant<sup>®</sup> III SYBR<sup>®</sup> Green QPCR Master Mix (Stratagene, San Diego, CA, USA). Expression analysis was conducted for eight biological replicates in technical triplicates.

#### 4.8. Compound Isolation and Structure Elucidation

The expression strain *E. coli* BL21 was transformed with the plasmid construct for PtTPS5 expression by electroporation. The cells were plated on LB agar containing ampicillin ( $100 \text{ mg mL}^{-1}$ ) and incubated at  $37^\circ\text{C}$  overnight. A single colony was selected from the plate and incubated in 10 mL of liquid LB medium at  $37^\circ\text{C}$  overnight. The fresh culture was sequentially used to inoculate larger culture volumes ( $1 \text{ mL L}^{-1}$ , 8 L in total), followed by cultivation until an  $\text{OD}_{600}$  of 0.4–0.6 was reached. The cultures were cooled to  $18^\circ\text{C}$  and IPTG solution ( $400 \text{ mM}$ ,  $1 \text{ mL L}^{-1}$ ) was added to induce protein expression. The cultures were grown overnight and then cells were harvested by centrifugation ( $3.600 \times g$ , 40 min). The pelleted cells were resuspended in binding buffer ( $10 \text{ mL L}^{-1}$  culture;  $20 \text{ mM Na}_2\text{HPO}_4$ ,  $500 \text{ mM NaCl}$ ,  $20 \text{ mM imidazole}$ ,  $1 \text{ mM MgCl}_2$ ,  $\text{pH} = 7.4$ ) and lysed by ultra-sonication ( $7 \times 1 \text{ min}$ ). The supernatant obtained by centrifugation ( $11.000 \times g$ , 10 min) contained the target protein for enzyme incubations.

The enzymatic assay was conducted in a total volume of 160 mL, containing 80 mL of enzyme preparation (with a protein concentration of  $1.3 \text{ mg mL}^{-1}$  as determined by Bradford assay), 80 mg (0.185 mmol) FPP trisammonium salt dissolved in 10 mL water, 304 mg (3.2 mmol)  $\text{MgCl}_2$  in 1.2 mL water (for a final concentration of 20 mM) and 68.8 mL incubation buffer (10 mM Tris-HCl, 1 mM dithiothreitol, 10% glycerol,  $\text{pH} 7.5$ ). The incubation was performed at  $28^\circ\text{C}$  overnight. The reaction mixture was extracted with pentane ( $3 \times 150 \text{ mL}$ ), the extract was dried with  $\text{MgSO}_4$  and evaporated under reduced pressure to give 17 mg crude product. Purification by column chromatography on silica gel (pentane/ether = 4:1) and then HPLC ( $\text{H}_2\text{O}/\text{methanol} = 25:75$ ;  $5:0 \text{ mL min}^{-1}$ ; Smartline HPLC series; KNAUER Eurospher II 100-5  $\text{C}_{18}$ ,  $5 \mu\text{m}$ ,  $250 \times 8 \text{ mm}$ ) yielded (1*S*,5*S*,7*R*,10*R*)-guaia-4(15)-en-11-ol and (1*S*,7*R*,10*R*)-guaia-4-en-11-ol as colorless oils.

(1*S*,5*S*,7*R*,10*R*)-Guaia-4(15)-en-11-ol. (1). Yield: 0.9 mg (0.004 mmol, 2%). TLC (pentane/ether = 4:1):  $R_f = 0.17$ . Optical rotation:  $[\alpha]_{\text{D}}^{20} = +47.8$  ( $c$  0.09,  $\text{C}_6\text{D}_6$ ). HRMS (EI):  $m/z = 222.1978$  (calc. for  $[\text{C}_{15}\text{H}_{26}\text{O}]^+$  222.1978). GC (HP5-MS):  $I = 1660$ . MS (EI, 70 eV):  $m/z$  (%) = 222 (0.3), 204 (29), 189 (24), 175 (2), 161 (24), 149 (22), 133 (11), 121 (19), 107 (31), 91 (41), 81 (54), 67 (25), 59 (100), 53(13), 41 (35). IR (diamond ATR):  $\tilde{\nu}/\text{cm}^{-1} = 2953$  (m), 2923 (s), 2854 (m), 1714 (w), 1650 (w), 1456 (m), 1376 (m), 1260 (m), 1094 (s), 1020 (s), 873 (m), 800 (s).

(1*S*,7*R*,10*R*)-Guaia-4-en-11-ol (2). Yield: 0.6 mg (0.003 mmol, 2%). TLC (pentane/ether = 4:1):  $R_f = 0.21$ . Optical rotation:  $[\alpha]_{\text{D}}^{20} = +21.7$  ( $c$  0.06,  $\text{C}_6\text{D}_6$ ). HRMS (EI):  $m/z = 222.1975$  (calc. for  $[\text{C}_{15}\text{H}_{26}\text{O}]^+$  222.1978). GC (HP5-MS):  $I = 1661$ . MS (EI, 70 eV):  $m/z$  (%) = 222 (4), 204 (81), 189 (69), 175 (7), 161 (75), 147 (27), 133 (29), 119 (37), 105 (61), 91 (81), 79 (68), 67 (30), 59 (100), 51(5), 41 (55). IR (diamond ATR):  $\tilde{\nu}/\text{cm}^{-1} = 2954$  (s), 2923 (s), 2854 (s), 1723 (w), 1670 (w), 1459 (m), 1376 (m), 1260 (w), 1096 (w), 1025 (w), 800 (w).

Isotopic labelling experiments were performed to determine the absolute configurations of (1*S*,5*S*,7*R*,10*R*)-guaia-4(15)-en-11-ol and (1*S*,7*R*,10*R*)-guaia-4-en-11-ol. For the reactions with DMAPP (1 mg in 1 mL water) and (*Z*)-(4- $^{13}\text{C}$ ,4- $^2\text{H}$ )IPP or (*E*)-(4- $^{13}\text{C}$ ,4- $^2\text{H}$ )IPP (1 mg in 1 mL water), protein preparations of PtTPS5 (1 mL) and FPPS (1 mL, [22]),  $\text{MgCl}_2$  (19 mg, final concentration 20 mM) and incubation buffer (6 mL) were added. For the reactions with (*R*)-(1- $^{13}\text{C}$ ,1- $^2\text{H}$ )IPP or (*S*)-(1- $^{13}\text{C}$ ,1- $^2\text{H}$ )IPP (1 mg in 1 mL water), protein preparations of PtTPS5 (1 mL), FPPS (1 mL, [22]), IDI (1 mL, [21]),  $\text{MgCl}_2$  (19 mg, final concentration 20 mM) and incubation buffer (6 mL) were added. After incubation with shaking at  $28^\circ\text{C}$  overnight, the reaction mixtures were extracted with  $\text{C}_6\text{D}_6$ , the extracts were dried with  $\text{MgSO}_4$  and analyzed by NMR and GC-MS.

#### 4.9. Statistical Analysis

Throughout the manuscript, data are presented as means  $\pm$  SE. Statistical analysis was performed with SigmaPlot 11.0 for Windows (Systat Software Inc., San Jose, CA, USA) and is described in the figure and table legends for the respective experiments. Whenever necessary, the data were log transformed to meet statistical assumptions such as normality and homogeneity of variances.

#### 4.10. Accession Numbers

Raw reads of the RNAseq experiment were deposited in the NCBI Sequence Read Archive (SRA) under the BioProject accession PRJNA660564 ‘Oomycete-induced changes in the root transcriptome of poplar’.

**Supplementary Materials:** The following are available online, Figure S1: *Phytophthora cactorum* infection induces the accumulation of sesquiterpenes in *Populus trichocarpa* roots. (A) Pulverized root material was extracted with hexane and the extracts were analyzed using GC-MS. 1, elemol; 2, unidentified sesquiterpene alcohol 1; 3, unidentified sesquiterpene alcohol 2; 4, myristaldehyde. (B) *P. cactorum* infection was verified by RT-qPCR analysis of *Ypt1* (*Phytophthora*-specific Ras-related protein, Schena et al. [18]) gene expression. Means and SE are shown ( $n = 7-8$ ). Asterisks indicate statistical significance as assessed by Wilcoxon rank sum test (\*\*\*) ( $p < 0.001$ ), ( $T = 36.00$ ,  $p < 0.001$ ). Figure S2: Representative GC-MS chromatograms of hexane extracts made from untreated *P. trichocarpa* roots, *Phytophthora cactorum*-treated *P. trichocarpa* roots, and *P. cactorum* mycelium. 1, 1-hexanol; 2, 1,8-cineole\*; 3, benzyl alcohol\*; 4, salicylaldehyde\*; 5, (E)-4-nonenal; 6, 2-phenylethanol\*; 7,  $\alpha$ -terpineole\*; 8, contamination (softener); 9, elemol; 10, (1S,5S,7R,10R)-guaia-4(15)-en-11-ol + (1S,7R,10R)-guaia-4-en-11-ol; 11, myristaldehyde\*; 12, myristic acid\*; 13, unidentified compound; 14, pentadecanoic acid\*; 15, palmitic acid\*; 16, oleic acid\*; 17, stearic acid\*; 18, unidentified compound; IS, internal standard (nonylacetate). Compounds marked by asterisks were identified using authentic standards. Figure S3: Determination of the absolute configuration of (1S,5S,7R,10R)-guaia-4(15)-en-11-ol by enantioselective deuteration. Partial HSQC spectra of A) unlabeled guaia-4(15)-en-11-ol (region for C2, C6 and C8), guaia-4(15)-en-11-ol obtained from B) (R)-(1-13C,1-2H)IPP and C) (S)-(1-13C,1-2H)IPP, D) unlabeled guaia-4(15)-en-11-ol (region for C3 and C9), and guaia-4(15)-en-11-ol obtained from E) (E)-(4-13C,4-2H)IPP and F) (Z)-(4-13C,4-2H)IPP. Taken together, these data establish the absolute configuration of (1S,5S,7R,10R)-guaia-4(15)-en-11-ol. Supplemental Table S1: Gene expression values (RPKM) and statistical parameters for all genes significantly upregulated (fold change > 5.0) upon *Phytophthora cactorum* infection in *Populus trichocarpa* roots. Supplemental Table S2: Compounds used as standards for GC-MS analysis. Supplemental Table S3: Experimentally determined Kovats retention indices. Supplemental Table S4: HPLC gradients used for separation and analysis of metabolites. Supplemental Table S5: Parameters used for LC-MS/MS analysis. Details of the HPLC gradient are given in Supplemental Table S4. CE, collision energy; DP, declustering potential; Q1, quadrupole 1; Q3, quadrupole 3. Supplemental Table S6: Primers used in this study. Supplemental Table S7: Ubiquitin (UBQ), *actin*, *elongation factor 1 alpha* (EF1 $\alpha$ ), *histone superfamily protein H3* (HIS), and *tubulin* (TUB) were tested as reference genes for RT-qPCR.

**Author Contributions:** T.G.K., N.D.L., J.S.D. and J.G. designed research. N.D.L., J.M. and H.X. carried out the experimental work. N.D.L., J.M., H.X. and J.S.D. analyzed data. T.G.K., N.D.L. and J.S.D. wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was funded by the Max Planck Society and by the Deutsche Forschungsgemeinschaft (DI1536/7-2).

**Data Availability Statement:** All data generated or analyzed during this study are included in the main text or supplement of this article. Raw sequences of the RNAseq experiment were deposited in the NCBI Sequence Read Archive under the BioProject accession PRJNA660564.

**Acknowledgments:** We thank all gardeners of the MPICE for rearing poplar plants.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Sample Availability:** Samples of the compounds described in this study are available from the authors (J.S.D and T.G.K).

## References

1. Tholl, D. Biosynthesis and biological functions of terpenoids in plants. In *Biotechnology of Isoprenoids*; Schrader, J., Bohlmann, J., Eds.; Springer: Cham, Switzerland, 2015; pp. 63–106.
2. Unsicker, S.B.; Kunert, G.; Gershenson, J. Protective perfumes: The role of vegetative volatiles in plant defense against herbivores. *Curr. Opin. Plant Biol.* **2009**, *12*, 479–485. [[CrossRef](#)]
3. Junker, R.R.; Gershenson, J.; Unsicker, S.B. Floral odor bouquet loses its ant repellent properties after inhibition of terpene biosynthesis. *J. Chem. Ecol.* **2011**, *37*, 1323–1331. [[CrossRef](#)]



4. Zhou, W.; Kügler, A.; McGale, E.; Haverkamp, A.; Knaden, M.; Guo, H.; Beran, F.; Yon, F.; Li, R.; Lackus, N.; et al. Tissue-specific emission of (*E*)- $\alpha$ -bergamotene helps resolve the dilemma when pollinators are also herbivores. *Curr. Biol.* **2017**, *27*, 1336–1341. [[CrossRef](#)] [[PubMed](#)]
5. Brooks, C.J.; Watson, D.G. Terpenoid phytoalexins. *Nat. Prod. Rep.* **1991**, *8*, 367–389. [[CrossRef](#)] [[PubMed](#)]
6. Schmelz, E.A.; Kaplan, F.; Huffaker, A.; Dafoe, N.J.; Vaughan, M.M.; Ni, X.Z. Identity, regulation, and activity of inducible diterpenoid phytoalexins in maize. *Proc. Natl. Acad. Sci. USA* **2011**, *108*. [[CrossRef](#)] [[PubMed](#)]
7. Huffaker, A.; Kaplan, F.; Vaughan, M.M.; Dafoe, N.J.; Ni, X.; Rocca, J.R.; Alborn, H.T.; Teal, P.E.; Schmelz, E.A. Novel acidic sesquiterpenoids constitute a dominant class of pathogen-induced phytoalexins in maize. *Plant Physiol.* **2011**, *156*, 2082–2097. [[CrossRef](#)]
8. Degenhardt, J.; Köllner, T.G.; Gershenzon, J. Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. *Phytochemistry* **2009**, *70*, 1621–1637. [[CrossRef](#)] [[PubMed](#)]
9. Dudareva, N.; Pichersky, E.; Gershenzon, J. Biochemistry of plant volatiles. *Plant Physiol.* **2004**, *135*, 1893–1902. [[CrossRef](#)]
10. Bathe, U.; Tissier, A. Cytochrome P450 enzymes: A driving force of plant diterpene diversity. *Phytochemistry* **2019**, *161*, 149–162. [[CrossRef](#)]
11. Danner, H.; Böckler, G.A.; Irmisch, S.; Yuan, J.S.; Chen, F.; Gershenzon, J.; Unsicker, S.B.; Köllner, T.G. Four terpene synthases produce major compounds of the gypsy moth feeding-induced volatile blend of *Populus trichocarpa*. *Phytochemistry* **2011**, *72*. [[CrossRef](#)]
12. Irmisch, S.; Jiang, Y.; Chen, F.; Gershenzon, J.; Köllner, T.G. Terpene synthases and their contribution to herbivore-induced volatile emission in western balsam poplar (*Populus trichocarpa*). *BMC Plant Biol.* **2014**, *14*, 270. [[CrossRef](#)] [[PubMed](#)]
13. Irmisch, S.; Müller, A.T.; Schmidt, L.; Günther, J.; Gershenzon, J.; Köllner, T.G. One amino acid makes the difference: The formation of *ent*-kaurene and 16 $\alpha$ -hydroxy-*ent*-kaurane by diterpene synthases in poplar. *BMC Plant Biol.* **2015**, *15*, 262. [[CrossRef](#)] [[PubMed](#)]
14. Lackus, N.D.; Lackner, S.; Gershenzon, J.; Unsicker, S.B.; Köllner, T.G. The occurrence and formation of monoterpenes in herbivore-damaged poplar roots. *Sci. Rep.* **2018**, *8*, 17936. [[CrossRef](#)] [[PubMed](#)]
15. Lackus, N.D.; Petersen, N.P.; Nagel, R.; Schmidt, A.; Irmisch, S.; Gershenzon, J.; Köllner, T.G. Identification and characterization of *trans*-isopentenyl diphosphate synthases involved in herbivory-induced volatile terpene formation in *Populus trichocarpa*. *Molecules* **2019**, *24*, 2408. [[CrossRef](#)]
16. Jones, R.; Sutherland, M. Hedycaryol, the precursor of elemol. *Chem. Commun.* **1968**, *20*, 1229–1230. [[CrossRef](#)]
17. Böckler, G.A.; Gershenzon, J.; Unsicker, S.B. Phenolic glycosides of the Salicaceae and their role as anti-herbivore defenses. *Phytochemistry* **2011**, *72*, 1497–1509. [[CrossRef](#)]
18. Schena, L.; Duncan, J.M.; Cooke, D.E.L. Development and application of a PCR-based ‘molecular tool box’ for the identification of *Phytophthora* species damaging forests and natural ecosystems. *Plant Pathol.* **2008**, *57*, 64–75. [[CrossRef](#)]
19. Tissandie, L.; Viciania, S.; Brevard, H.; Meierhenrich, U.J.; Filippi, J.J. Towards a complete characterisation of guaiacwood oil. *Phytochemistry* **2018**, *149*, 64–81. [[CrossRef](#)]
20. Xu, H.; Dickschat, J.S. Germacrene A—A central intermediate in sesquiterpene biosynthesis. *Chemistry* **2020**. [[CrossRef](#)]
21. Rinkel, J.; Dickschat, J.S. Addressing the chemistry of germacrene A by isotope labeling experiments. *Org. Lett.* **2019**, *21*, 2426–2429. [[CrossRef](#)]
22. Rabe, P.; Rinkel, J.; Nubbemeyer, B.; Köllner, T.G.; Chen, F.; Dickschat, J.S. Terpene cyclases from social amoebae. *Angew. Chem. Int. Ed.* **2016**, *55*, 15420–15423. [[CrossRef](#)] [[PubMed](#)]
23. Cornforth, J.W.; Cornforth, R.H.; Popják, G.; Yengoyan, L. Studies on the biosynthesis of cholesterol. XX. Steric course of decarboxylation of 5-pyrophosphomevalonate and of the carbon to carbon bond formation in the biosynthesis of farnesyl pyrophosphate. *J. Biol. Chem.* **1966**, *241*, 3970–3987. [[CrossRef](#)]
24. Lauterbach, L.; Rinkel, J.; Dickschat, J.S. Two bacterial diterpene synthases from *Allokutzneria albata* produce bonnadiene, phomopsene, and allokutznerene. *Angew. Chem. Int. Ed.* **2018**, *57*, 8280–8283. [[CrossRef](#)] [[PubMed](#)]
25. Toljamo, A.; Blande, D.; Karenlampi, S.; Kokko, H. Reprogramming of strawberry (*Fragaria vesca*) root transcriptome in response to *Phytophthora cactorum*. *PLoS ONE* **2016**, *11*, e0161078. [[CrossRef](#)] [[PubMed](#)]
26. Yadav, H.; Dreher, D.; Athmer, B.; Porzel, A.; Gavrín, A.; Baldermann, S.; Tissier, A.; Hause, B. Medicago terpene synthase 10 is involved in defense against an oomycete root pathogen. *Plant Physiol.* **2019**, *180*, 1598–1613. [[CrossRef](#)] [[PubMed](#)]
27. Köllner, T.G.; O’Maille, P.E.; Gatto, N.; Boland, W.; Gershenzon, J.; Degenhardt, J. Two pockets in the active site of maize sesquiterpene synthase TPS4 carry out sequential parts of the reaction scheme resulting in multiple products. *Arch. Biochem. Biophys.* **2006**, *448*, 83–92. [[CrossRef](#)]
28. Davis, G.D.; Essenberg, M. (+)- $\delta$ -cadinene is a product of sesquiterpene cyclase activity in cotton. *Phytochemistry* **1995**, *39*, 553–567. [[CrossRef](#)]
29. Chen, X.-Y.; Chen, Y.; Heinsteinst, P.; Davisson, V.J. Cloning, expression, and characterization of (+)- $\delta$ -cadinene synthase: A catalyst for cotton phytoalexin biosynthesis. *Arch. Biochem. Biophys.* **1995**, *324*, 255–266. [[CrossRef](#)]
30. Mao, H.; Liu, J.; Ren, F.; Peters, R.J.; Wang, Q. Characterization of CYP71Z18 indicates a role in maize zealexin biosynthesis. *Phytochemistry* **2016**, *121*, 4–10. [[CrossRef](#)]
31. Tian, X.; Ruan, J.-X.; Huang, J.-Q.; Yang, C.-Q.; Fang, X.; Chen, Z.-W.; Hong, H.; Wang, L.-J.; Mao, Y.-B.; Lu, S.; et al. Characterization of gossypol biosynthetic pathway. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E5410. [[CrossRef](#)]
32. Upchurch, R.G. Fatty acid unsaturation, mobilization, and regulation in the response of plants to stress. *Biotechnol. Lett.* **2008**, *30*, 967–977. [[CrossRef](#)] [[PubMed](#)]

33. Rojas, C.M.; Senthil-Kumar, M.; Tzin, V.; Mysore, K.S. Regulation of primary plant metabolism during plant-pathogen interactions and its contribution to plant defense. *Front. Plant Sci.* **2014**, *5*. [[CrossRef](#)] [[PubMed](#)]
34. Pohl, C.H.; Kock, J.L.; Thibane, V.S. Antifungal free fatty acids: A review. *Sci. Microb. Pathog. Commun. Curr. Res. Technol. Adv.* **2011**, *3*, 61–71.
35. Kachroo, A.; Kachroo, P. Fatty acid-derived signals in plant defense. *Annu. Rev. Phytopathol.* **2009**, *47*, 153–176. [[CrossRef](#)] [[PubMed](#)]
36. Kalemba, D.; Kunicka, A. Antibacterial and antifungal properties of essential oils. *Curr. Med. Chem.* **2003**, *10*, 813–829. [[CrossRef](#)]
37. Böckler, G.A.; Gershenzon, J.; Unsicker, S.B. Gypsy moth caterpillar feeding has only a marginal impact on phenolic compounds in old-growth black poplar. *J. Chem. Ecol.* **2013**, *39*, 1301–1312. [[CrossRef](#)]
38. Lackner, S.; Lackus, N.D.; Paetz, C.; Köllner, T.G.; Unsicker, S.B. Aboveground phytochemical responses to belowground herbivory in poplar trees and the consequence for leaf herbivore preference. *Plant Cell Environ.* **2019**. [[CrossRef](#)]
39. Lackus, N.D.; Müller, A.; Kröber, T.D.U.; Reichelt, M.; Schmidt, A.; Nakamura, Y.; Paetz, C.; Luck, K.; Lindroth, R.L.; Constabel, C.P.; et al. The occurrence of sulfated salicinoids in poplar and their formation by sulfotransferase 1. *Plant Physiol.* **2020**, *183*, 137–151. [[CrossRef](#)]
40. Ramírez-Carvajal, G.A.; Morse, A.M.; Davis, J.M. Transcript profiles of the cytokinin response regulator gene family in *Populus* imply diverse roles in plant development. *New Phytol.* **2008**, *177*, 77–89. [[CrossRef](#)]
41. Xu, M.; Zhang, B.; Su, X.; Zhang, S.; Huang, M. Reference gene selection for quantitative real-time polymerase chain reaction in *Populus*. *Anal. Biochem.* **2011**, *408*, 337–339. [[CrossRef](#)]
42. Wang, H.L.; Chen, J.; Tian, Q.; Wang, S.; Xia, X.; Yin, W. Identification and validation of reference genes for *Populus euphratica* gene expression analysis during abiotic stresses by quantitative real-time pcr. *Physiol. Plant* **2014**, *152*, 529–545. [[CrossRef](#)] [[PubMed](#)]

## Appendix I

### **1,2- or 1,3-Hydride Shifts: What Controls Guaiane Biosynthesis?**

*Chem. Eur. J.* **2021**, *27*, 9758

DOI: 10.1002/chem.202101371

# 1,2- or 1,3-Hydride Shifts: What Controls Guaiane Biosynthesis?

Houchao Xu,<sup>[a]</sup> Bernd Goldfuss,<sup>\*,[b]</sup> and Jeroen S. Dickschat<sup>\*,[a]</sup>

Dedicated to Prof. Wittko Francke, an exceptional natural product chemist, who passed away on 27th December 2020.

**Abstract:** A systematic computational study addressing the entire chemical space of guaianes in conjunction with an analysis of all known compounds shows that 1,3-hydride shifts are rare events in guaiane biosynthesis. As demonstrated here, 1,3-hydride shifts towards guaianes can only be realized for two stereochemically well defined out of numerous possible stereoisomeric skeletons. One example is given by the mechanism of guaia-4(15)-en-11-ol synthase from California poplar, an enzyme that yields guaianes with unusual stereochemical properties. The general results from DFT calculations were experimentally verified through isotopic-labeling experiments with guaia-4(15)-en-11-ol synthase.

During the past two decades many terpene synthases (TPSs) have been characterized, mainly from plants,<sup>[1,2]</sup> bacteria<sup>[3]</sup> and fungi.<sup>[4,5]</sup> These remarkable enzymes convert acyclic, achiral polyisoprenoid diphosphates into structurally complex, often polycyclic, chiral and enantiomerically enriched terpenes. These transformations involve just a single enzyme catalyzed reaction and proceed through cationic cascade reactions inside a hydrophobic cavity of the TPS. Because of their transient nature the cationic intermediates along the cascade cannot be observed spectroscopically, but especially isotopic labeling experiments<sup>[6,7]</sup> and DFT or QM/MM calculations<sup>[8–13]</sup> have helped to develop a deep mechanistic understanding of TPS catalysis. Also structure based site-directed mutagenesis can give valuable insights,<sup>[14–16]</sup> especially if an enzyme variant leads to an aberrant product formed by deprotonation of a cationic intermediate, giving indirect evidence for its existence. In some

cases terpene cyclizations proceed through a neutral (deprotonated) intermediate that can be reactivated by reprotonation for further downstream cyclization steps; these neutral intermediates can often be observed as minor products, as they can leak from the enzyme's active site. It is, however, difficult to distinguish in these cases between true intermediates and shunt products, because instead of a deprotonation-reprotonation sequence a direct intramolecular or water/enzyme mediated proton transfer could bypass such a hypothetical neutral "intermediate". While keeping this in mind, for simplification we will no longer differentiate here between neutral "intermediate" and "shunt product", or only where it is relevant.

Germacrene A (1) and hedycaryol (2) belong to the most important intermediates of sesquiterpene biosynthesis and numerous compounds derive from them,<sup>[17]</sup> likely because their fairly strained ten-membered ring is sufficiently reactive for further protonation induced cyclizations (Scheme 1).

In almost all cases the resulting structures can be explained by reprotonation of a double bond with a well explainable face selectivity, that is, attack of the proton from the enzyme exposed face and not at the inner face of the macrocycle. Starting from different ring conformations (DU, UU, DD and UD, referring to Me14 and Me15 down=D or up=U) the reprotonation at C4 leads to the four stereoisomeric intermediates A–D, while the four intermediates E–H can be reached through reprotonations at C10 (plus their enantiomers from the antipodes of 1 and 2). As a result of the *E*-configured double bonds in 1 and 2 the C4 reprotonations always lead to a *trans* orientation of Me15 and H5, while for C10 reprotonations Me14 and H1 are always *trans*. We have recently reported about the oomycete infection induced PtTPS5 from *Populus trichocarpa* (California poplar) that converts farnesyl diphosphate (FPP) into (1*S*,7*R*,10*R*)-guaia-4-en-11-ol (3) and (1*S*,5*S*,7*R*,10*R*)-guaia-4(15)-en-11-ol (4), besides minor amounts of 2.<sup>[18]</sup> The double bond positioning in 3 and 4 indicates a cationic precursor with the charge residing at C4, which could be reached directly through cyclization of 2 upon C10 protonation, but none of the intermediates E–H fulfills the stereochemical requirements of the observed products.

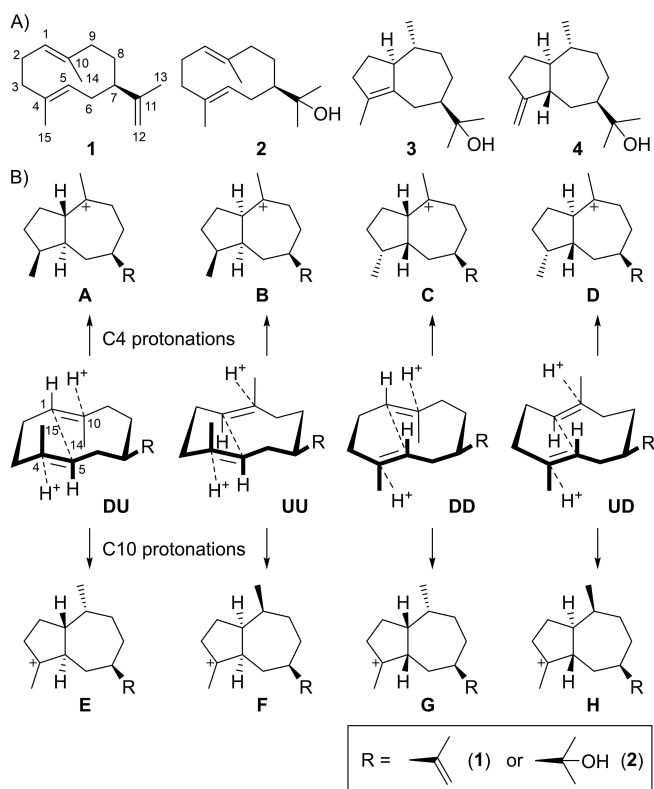
A systematic analysis of the reachable chemical space revealed that such a situation is very rare among guaiane sesquiterpenes. For this purpose, the possible structures of guaadienes and guaienols were identified as the three different deprotonation products of each of the intermediates A–H (Schemes S1 and S2 in the Supporting Information). Further compounds can be reached through 1,2- or 1,3-hydride

[a] H. Xu, Prof. Dr. J. S. Dickschat  
 Kekulé-Institute for Organic Chemistry and Biochemistry  
 University of Bonn  
 Gerhard-Domagk-Straße 1, 53121 Bonn (Germany)  
 E-mail: dickschat@uni-bonn.de

[b] Prof. Dr. B. Goldfuss  
 Institute for Organic Chemistry, University of Cologne  
 Greinstraße 4, 50939 Cologne (Germany)  
 E-mail: goldfuss@uni-koeln.de

Supporting information for this article is available on the WWW under <https://doi.org/10.1002/chem.202101371>

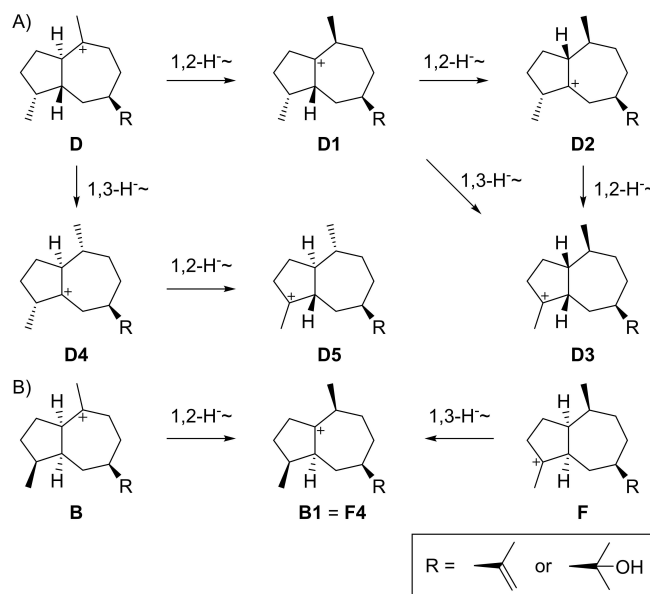
© 2021 The Authors. Published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.



**Scheme 1.** Biosynthesis of guaiane sesquiterpenes. A) Structures of germacrene A (1) and of the PtTP55 products 2–4. B) Possible cyclization reactions from 1 and 2 to different stereoisomers of the guaiane skeleton.

migration or their combinations and different deprotonation events (Schemes S3–S14). The systematics of this approach is summarized for the D series in Scheme 2A which includes the precursors for 3 (D4 or D5) and 4 (D5). In some cases the same compounds can be formed through alternative pathways, such as D3 and its deprotonation products can hypothetically arise from D1 by two sequential 1,2-hydride shifts or one 1,3-hydride migration (but in other cases the order of steps is relevant, for example, starting from D the sequence of 1,2- plus 1,3-hydride migration leads to D3, while the reverse order of 1,3- plus 1,2-hydride transfer leads to another stereoisomer D5). This analysis also turned out that some compounds can be obtained from different initial cyclization products, for example, a 1,2-hydride shift from B or a 1,3-hydride shift from F both lead to B1 = F4 (Scheme 2B). All structures are summarized together with the information about their potential precursors A–H in Figures S1–S9.

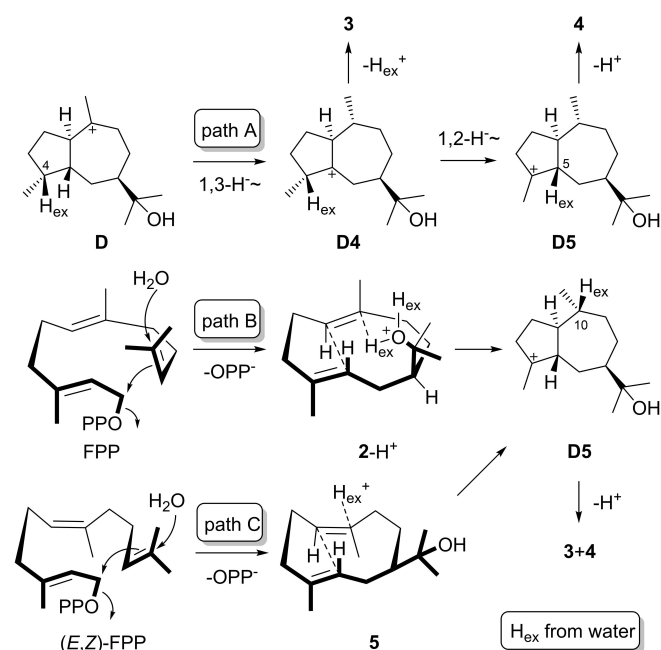
These considerations have so far neglected whether the proposed hydride migrations can indeed be realized or not. While it seems rationale to assume that 1,2-hydride shifts may be possible in every case, some of the 1,3-hydride shifts that were taken into account might be prevented by steric constraints, that is, the hydride to be shifted may point away from the empty p orbital at the cationic center, so that no significant orbital overlap can be achieved. To gain deeper insights DFT calculations for all eight series starting from A–H,



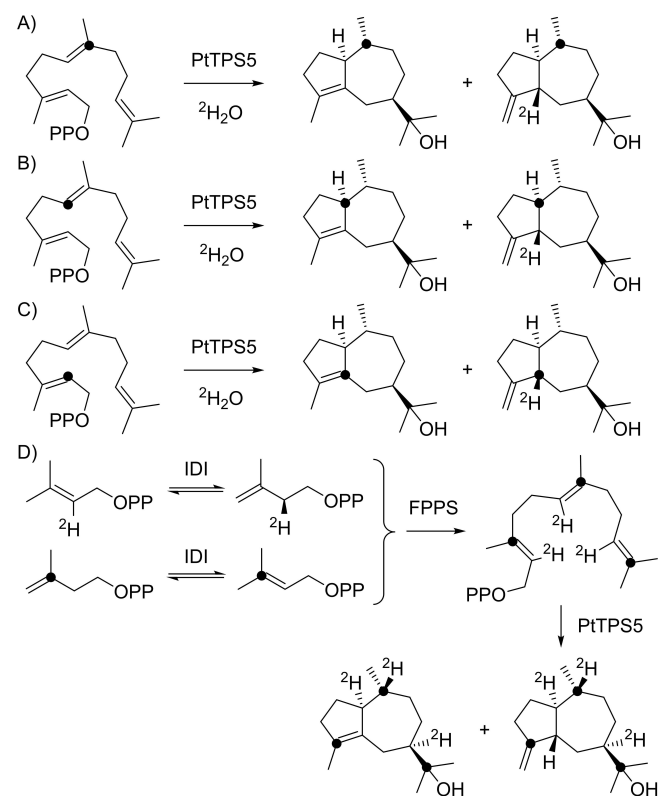
**Scheme 2.** Exploring the reachable chemical space of guaiane sesquiterpenes. A) Cationic intermediates from the initially formed bicyclic intermediate, here exemplified for the D series. B) Some cationic intermediates can be reached from different starting points.

with both substituents of an isopropenyl or a hydroxyisopropyl group, were carried out for all (corresponding) hydride migrations as in Scheme 2A (Figures S10–S25). As expected, low to moderate transition state (TS) barriers between 0.76 kcal/mol (H-TS1 in Figure S25) and 11.14 kcal/mol (H-TS3 in Figure S16) were obtained for all 1,2-hydride shifts, with an average TS barrier of 5.13 kcal/mol. In contrast, several 1,3-hydride shifts could not be realized, including those from E1 to E3 and the corresponding intermediates in the F, G and H series (Figures S18–S25), while for the A–D series high TS barriers were found for this step (Figures S10–S17). Because this step can be substituted by two sequential 1,2-hydride shifts in all cases, 1,3-hydride shifts are, if possible at all, likely not relevant here. Furthermore, hydride shifts could not be realized for the steps from B to B4 and from C to C4 (Figures S12–S15), while the corresponding steps showed high TS barriers in the E–H series (Figures S18–S25). Only for the A and D series this step with TS barriers between 7.27 and 9.72 kcal/mol is feasible (Figures S10, S11, S16 and S17). These 1,3-hydride shifts, where possible, open the path towards guaiane stereoisomers that cannot be reached through another sequence of hydride shifts. Going back with these insights to the known natural products and their possible mechanisms of formations (cf. precursor cations and their color code in Figures S1–S9), it becomes clear that the PtTP55 product 4 is the only known guaiane that must be generated with participation of a 1,3-hydride shift, at least if the so far discussed simple mechanistic models apply, whereas in many other cases an optional 1,3-hydride shift can be substituted by two energetically more feasible 1,2-hydride transfers. The only other known compounds for which 1,3-hydride shifts could be relevant are 3, and it is logical to assume a common biosynthetic mechanism for 3 and 4 by PtTP55, and

(1*S*,7*R*,10*R*)-guai-4-en-11-ol from *Bulnesia sarmientoi* (Figure S6).<sup>[19]</sup> In fact, for the latter compound only future experimental work will give clarification, as it can be formed from **A** through 1,3-hydride shift (9.72 kcal/mol, Figure S10) and depro-



**Scheme 3.** Biosynthetic hypotheses for the PtTPS5 products **3** and **4**.

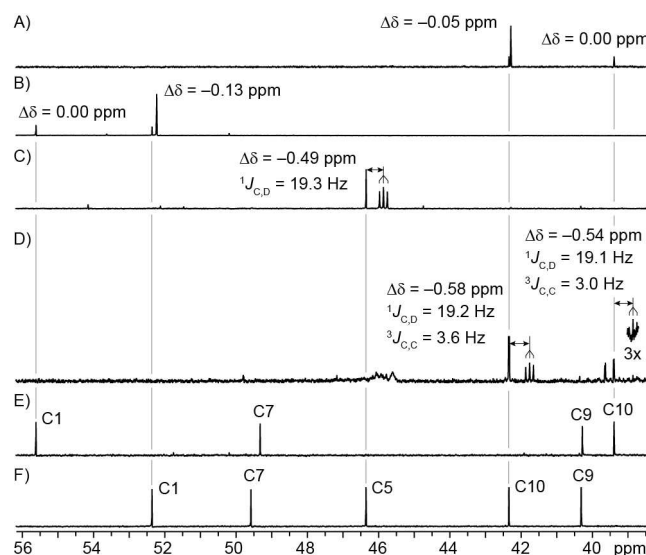


**Scheme 4.** Labeling experiments on the cyclization mechanism of PtTPS5.

tonation, or from **D** through two sequential 1,2-hydride shifts (highest TS is 8.54 kcal/mol, Figure S17) and deprotonation.

Are mechanistic alternatives to the so far considered path **A** of Scheme 3 for the biosynthesis of **3** and **4** possible? The cyclization of FPP and capture with water could lead to protonated hedycaryol (**2-H<sup>+</sup>**), with subsequent direct intramolecular proton transfer to C10 from the inner sphere, which can directly result in cyclization to **D5** (path **B**). In this case, the observed PtTPS5 product **2** would be identified as a shunt product rather than an intermediate towards **3** and **4**. Or if (*2E,6Z*)-FPP would be the substrate of PtTPS5, this could be cyclized to **5**, a stereoisomer of **2**, that could be followed by outer sphere protonation to induce direct cyclization to **D5** (path **C**). This hypothesis is unlikely, because normal (*2E,6E*)-FPP is efficiently converted into **3** and **4** by PtTPS5, and there is no good mechanistic explanation for a *6E/6Z* double bond isomerization in FPP.

To distinguish between path **A** and path **B** isotopic labeling experiments were performed (Scheme 4 and Figure 1). While path **A** must operate with proton incorporation from water at C4 with subsequent migration to C5 in **4** or loss by deprotonation in **3**, path **B** should give proton incorporation at C10 of both products. Incubation of (*7-<sup>13</sup>C*)FPP<sup>[20]</sup> in deuterium oxide gave strongly enhanced singlets for C10 of **3** and **4** in the <sup>13</sup>C NMR spectrum, with a small upfield shift for **4** indicating deuterium incorporation two positions away (Figure 1A). The same experiment with (*6-<sup>13</sup>C*)FPP<sup>[20]</sup> gave singlets for C1 of **3** and **4**, again with an upfield shift for **4**, in line with deuterium incorporation at a neighboring position (Figure 1B), while (*2-<sup>13</sup>C*)FPP<sup>[20]</sup> in <sup>2</sup>H<sub>2</sub>O gave a singlet for C4 of **3** and an upfield shifted triplet for **4**, giving direct evidence for deuterium incorporation at C5 (Figure 1C), in line with path **A** and conflicting path **B**. The 1,3-hydride shift in the biosynthesis of **3**



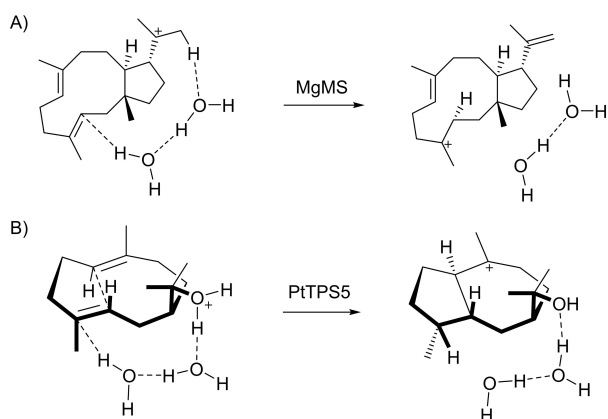
**Figure 1.** Labeling experiments on the cyclization mechanism of PtTPS5. Partial <sup>13</sup>C NMR spectra for labeled **3** and **4** obtained from **A**) (*7-<sup>13</sup>C*)FPP, **B**) (*6-<sup>13</sup>C*)FPP or **C**) (*2-<sup>13</sup>C*)FPP in D<sub>2</sub>O, **D**) obtained from [*3,7,11-<sup>13</sup>C*]<sub>3</sub>,*2,6,10-<sup>2</sup>H*]<sub>3</sub> FPP, and for **E**) unlabeled **3** and **F**) unlabeled **4**.

and **4** was demonstrated by incubation of ( $2\text{-}^2\text{H}$ )DMAPP<sup>[21]</sup> and ( $3\text{-}^{13}\text{C}$ )IPP<sup>[22]</sup> with isopentenyl diphosphate isomerase from *Escherichia coli*<sup>[23]</sup> and FPP synthase (FPPS) from *Streptomyces coelicolor*.<sup>[24]</sup> This will yield a mixture of isotopomers of [ $3,7,11\text{-}^{13}\text{C}_3,2,6,10\text{-}^2\text{H}_3$ ]FPP in which each terpene unit carries either a  $^{13}\text{C}$ - or a  $^2\text{H}$ -labeling, but not both simultaneously. Further conversion with PtTPS5 gave corresponding mixtures of isotopomers of **3** and **4** (all eight FPP isotopomers in this mixture and their conversion by PtTPS5 are shown in Scheme S15). For the relevant carbon C10 that only gives a strong signal in the  $^{13}\text{C}$ NMR if it is  $^{13}\text{C}$ -labeled itself, either upfield shifted triplet signals for a  $^1J_{\text{CD}}$  coupling with deuterium or doublets for a  $^3J_{\text{CC}}$  coupling with  $^{13}\text{C}$  (C4) were observed, depending on whether the third unit of FPP was derived from ( $2\text{-}^2\text{H}$ )DMAPP or ( $3\text{-}^{13}\text{C}$ )IPP. The triplet coupling with deuterium unequivocally established the 1,3-hydride shift in the biosynthesis of **3** and **4** (the triplets are not explainable by two sequential 1,2-hydride shifts, because then the deuterium and

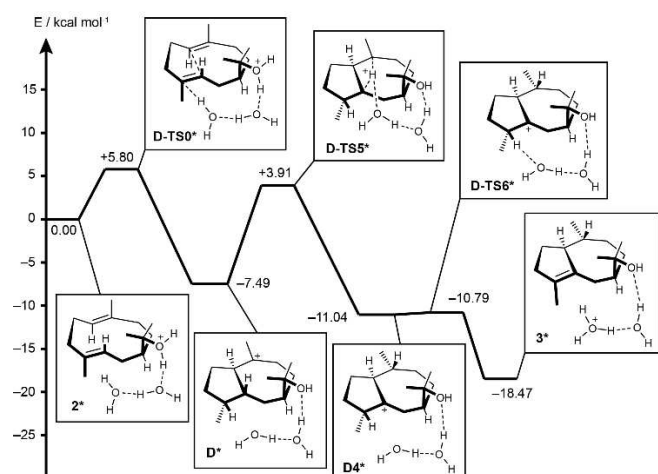
the  $^{13}\text{C}$ -labelings must be incorporated into the same terpene unit, which is not possible from the precursors used).

After having established path A experimentally and computationally, a possible mechanism for the protonation-induced cyclization of **2** by PtTPS5 was investigated in more detail. Here the question is what could be the source of the proton to induce the second cyclization? For a similar step by the fungal myrothec-15(17)-en-7-ol synthase from *Myrothecium gramineum* (MgMS) recently a proton transfer mediated through two water molecules to induce further cyclization events was suggested based on DFT calculations (Scheme 5A).<sup>[25]</sup> Such a mechanism could also be of interest for catalysis by PtTPS5 (Scheme 5B). To investigate this hypothesis DFT calculations were started from protonated hedycaryol (Figure 2). Water-mediated protonation at C4 required bridging by two molecules of water, leading through **D-TS0\*** with a TS barrier of 5.80 kcal/mol to **D\*** (one molecule of water was not sufficient to realize this step). The 1,3-hydride transfer can also be assisted by the water network, but the barrier for **D-TS5\*** (11.40 kcal/mol) is not better than without water (8.13 kcal/mol, Figure S17). However, the presence of water can explain a very smooth deprotonation of **D4\*** through **D-TS6\*** to the product **3\*** that is with 0.25 kcal/mol nearly barrierless. Instead of water, also active site residues or diphosphate could be involved in mediating proton transfers or act as a base in the final deprotonation.

In conclusion, we have shown that many guaiane skeletons can easily be reached by 1,2-hydride migrations. It is surprising that many theoretically possible structures have not been discovered from natural sources. This could mean that some structures are privileged in nature, but there is also another possible explanation: today countless studies rely only on GC/MS-based compound identification, even without the use of reference standards. It is well known that the stereoisomers of terpenes can have similar mass spectra and retention indices, and if one compound has been reported hundreds of times, it may be tempting to claim to have found the same "privileged structure", when in fact it is one of the missing compounds. In contrast to 1,2-hydride migrations, 1,3-hydride shifts are exceptional events in guaiane biosynthesis. A deep analysis of the eight stereoisomeric series **A–H** demonstrated that, in many cases, 1,3-hydride shifts are sterically impossible or they are associated with high barriers, making their participation very unlikely. However, in these cases, 1,3-hydride transfers cannot fully be excluded because relevant barriers could be lowered by the enzyme. For a very few cases, 1,3-hydride shifts must be considered as there is no other obvious solution to the formation of the observed skeletons; this includes the PtTPS5 products **3** and **4**. Here, the barriers for the 1,3-hydride shifts are comparably low and seem to be realizable, as verified in this study experimentally through isotopic labeling. Furthermore, our DFT calculations show that a conceptually interesting water-mediated proton transfer could be involved in the terpene cyclization to **3** and **4** by triggering the second cyclization event, but for deeper insights QM/MM calculations based on a crystal structure would be needed. Systematic explorations of the reachable chemical space of guaiane sesquiterpenes have demonstrated that for some known



**Scheme 5.** Water-mediated proton transfer in the terpene cyclizations A) to myrothec-15(17)-en-7-ol by MgMS, and B) to **3** and **4** by PtTPS5.



**Figure 2.** Energy profile (Gibbs energies, 298 K, mPW1PW91/6-311 + G(d,p)//B97D3/6-31g (d,p)) for the PtTPS5 mechanism including a water-mediated proton transfer.

compounds such as (1S,4S,5S,7S)-guai-9-en-11-ol from *B. sarmentoi* (Figure S1)<sup>[19]</sup> and guaia-5,11-diene from *Cymbastela hooperi* (Figure S9)<sup>[26]</sup> only one biosynthetic mechanism is plausible, while for most of the known compounds mechanistic alternatives can apply. In none of these cases has the cyclization mechanism been studied to distinguish between these alternatives; this opens up an interesting playground for future terpene biosynthesis work.

## Acknowledgements

This work was supported by the DFG (DI1536/7-2) and by the computing center of the University of Cologne (RRZK), providing CPU time on the DFG-funded supercomputer CHEOPS. We thank Tobias Köllner and Jonathan Gershenzon for the expression plasmid for PtTPS5. Open access funding enabled and organized by Projekt DEAL.

## Conflict of Interest

The authors declare no conflict of interest.

**Keywords:** carbocations · DFT calculations · hydride shifts · isotopes · terpenes

- [1] J. Degenhardt, T. G. Köllner, J. Gershenzon, *Phytochemistry* **2009**, *70*, 1621–1637.
- [2] Q. Jia, G. Li, T. G. Köllner, J. Fu, X. Chen, W. Xiong, B. J. Crandall-Stotler, J. L. Bowman, D. J. Weston, Y. Zhang, L. Chen, Y. Xie, F.-W. Li, C. J. Rothfels, A. Larsson, S. W. Graham, D. W. Stevenson, G. K.-S. Wong, J. Gershenzon, F. Chen, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 12328–12333.
- [3] J. S. Dickschat, *Nat. Prod. Rep.* **2016**, *33*, 87–110.
- [4] M. B. Quin, C. M. Flynn, C. Schmidt-Dannert, *Nat. Prod. Rep.* **2014**, *31*, 1449–1473.
- [5] A. Minami, T. Ozaki, C. Liu, H. Oikawa, *Nat. Prod. Rep.* **2018**, *35*, 1330–1346.
- [6] J. S. Dickschat, *Angew. Chem. Int. Ed.* **2019**, *58*, 15964–15976; *Angew. Chem.* **2019**, *131*, 16110–16123.
- [7] J. S. Dickschat, *Eur. J. Org. Chem.* **2017**, 4872–4882.
- [8] D. J. Tantillo, *Nat. Prod. Rep.* **2011**, *28*, 1035–1053.
- [9] S. H. Hare, D. J. Tantillo, *Beilstein J. Org. Chem.* **2016**, *12*, 377–390.
- [10] D. J. Tantillo, *Angew. Chem. Int. Ed.* **2017**, *56*, 10040–10045; *Angew. Chem.* **2017**, *129*, 10172–10178.
- [11] Y. Freud, T. Ansbacher, D. T. Major, *ACS Catal.* **2017**, *7*, 7653–7657.
- [12] K. Raz, R. Driller, N. Dimos, M. Ringel, T. Brück, B. Loll, D. T. Major, *J. Am. Chem. Soc.* **2020**, *142*, 21562–21574.
- [13] X. Tang, F. Zhang, T. Zeng, W. Li, S. Yin, R. Wu, *ACS Chem. Biol.* **2020**, *15*, 2820–2832.
- [14] M. Seemann, G. Z. Zhai, J. W. de Kraker, C. M. Paschall, D. W. Christianson, D. E. Cane, *J. Am. Chem. Soc.* **2002**, *124*, 7681–7689.
- [15] P. Baer, P. Rabe, K. Fischer, C. A. Citron, T. A. Klapschinski, M. Groll, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2014**, *53*, 7652–7656; *Angew. Chem.* **2014**, *126*, 7783–7787.
- [16] T. Tomita, S.-Y. Kim, K. Teramoto, A. Meguro, T. Ozaki, A. Yoshida, Y. Motoyoshi, N. Mori, K. Ishigami, H. Watanabe, M. Nishiyama, T. Kuzuyama, *ACS Chem. Biol.* **2017**, *12*, 1621–1628.
- [17] H. Xu, J. S. Dickschat, *Chem. Eur. J.* **2020**, *26*, 17318–17341.
- [18] N. D. Lackus, J. Morawetz, H. Xu, J. Gershenzon, J. S. Dickschat, T. G. Köllner, *Molecules* **2021**, *26*, 555.
- [19] L. Tissandie, S. Viciano, H. Brevard, U. J. Meierhenrich, J.-J. Filippi, *Phytochemistry* **2018**, *149*, 64–81.
- [20] P. Rabe, L. Barra, J. Rinkel, R. Riclea, C. A. Citron, T. A. Klapschinski, A. Janusko, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2015**, *54*, 13448–13451; *Angew. Chem.* **2015**, *127*, 13649–13653.
- [21] J. Rinkel, P. Rabe, X. Chen, T. G. Köllner, F. Chen, J. S. Dickschat, *Chem. Eur. J.* **2017**, *23*, 10501–10505.
- [22] P. Rabe, J. Rinkel, E. Dolja, T. Schmitz, B. Nubbemeyer, T. H. Luu, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2017**, *56*, 2776–2779; *Angew. Chem.* **2017**, *129*, 2820–2823.
- [23] F. M. Hahn, A. P. Hurlburt, C. D. Poulter, *J. Bacteriol.* **1999**, *181*, 4499–4504.
- [24] P. Rabe, J. Rinkel, B. Nubbemeyer, T. G. Köllner, F. Chen, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2016**, *55*, 15420–15423; *Angew. Chem.* **2016**, *128*, 15646–15649.
- [25] F. L. Lin, L. Lauterbach, J. Zhou, Y. H. Wang, J. M. Lv, G. D. Chen, D. Hu, H. Gao, X. S. Yao, J. S. Dickschat, *ACS Catal.* **2020**, *10*, 4306–4312.
- [26] G. M. König, A. D. Wright, *J. Org. Chem.* **1997**, *62*, 3837–3840.

Manuscript received: April 16, 2021  
 Accepted manuscript online: April 30, 2021  
 Version of record online: May 26, 2021



## Appendix J

### The enzyme mechanism of patchoulol synthase

*Beilstein J. Org. Chem.* **2022**, *18*, 13

DOI: [10.3762/bjoc.18.2](https://doi.org/10.3762/bjoc.18.2)



## The enzyme mechanism of patchoulol synthase

Houchao Xu<sup>1</sup>, Bernd Goldfuss<sup>2</sup>, Gregor Schnakenburg<sup>3</sup> and Jeroen S. Dickschat\*<sup>1</sup>

### Full Research Paper

Open Access

Address:

<sup>1</sup>Kekulé-Institute of Organic Chemistry and Biochemistry, University of Bonn, Gerhard-Domagk-Straße 1, 53121 Bonn, Germany,  
<sup>2</sup>Department of Chemistry, University of Cologne, Greinstraße 4, 50939 Cologne, Germany and <sup>3</sup>Institute of Inorganic Chemistry, University of Bonn, Gerhard-Domagk-Straße 1, 53121 Bonn, Germany

Email:

Jeroen S. Dickschat\* - dickschat@uni-bonn.de

\* Corresponding author

Keywords:

biosynthesis; DFT calculations; enzyme mechanisms; isotopes; terpenes

*Beilstein J. Org. Chem.* **2022**, *18*, 13–24.

<https://doi.org/10.3762/bjoc.18.2>

Received: 29 October 2021

Accepted: 17 December 2021

Published: 03 January 2022

Associate Editor: B. Nay

© 2022 Xu et al.; licensee Beilstein-Institut.

License and terms: see end of document.

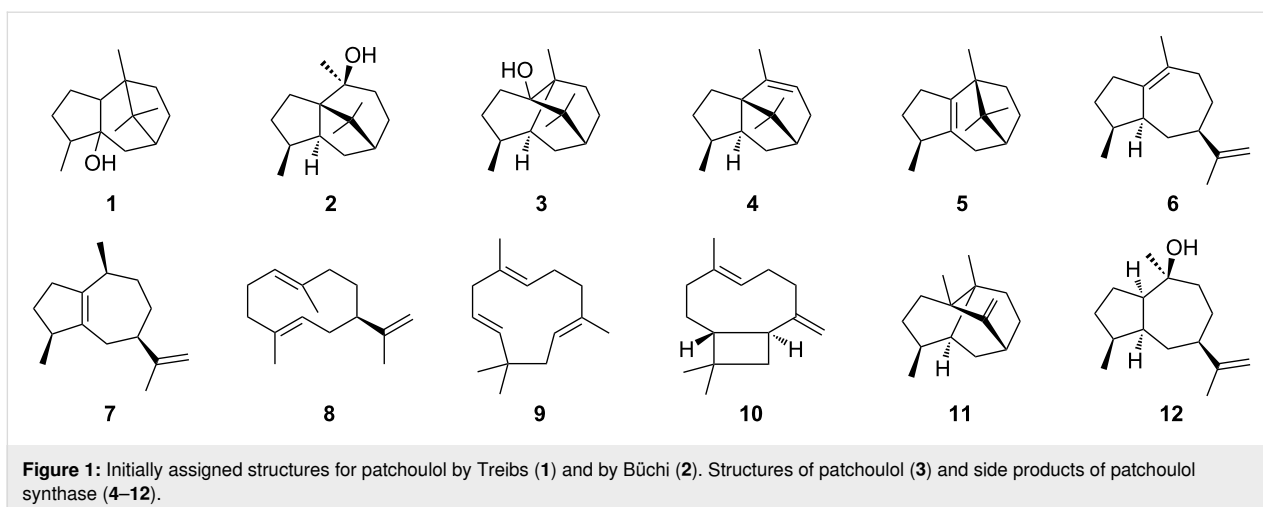
### Abstract

Different mechanisms for the cyclisation of farnesyl pyrophosphate to patchoulol by the patchoulol synthase are discussed in the literature. They are based on isotopic labelling experiments, but the results from these experiments are contradictory. The present work reports on a reinvestigation of patchoulol biosynthesis by isotopic labelling experiments and computational chemistry. The results are in favour of a pathway through the neutral intermediates germacrene A and  $\alpha$ -bulnesene that are both reactivated by protonation for further cyclisation steps, while previously discussed intra- and intermolecular hydrogen transfers are not supported. Furthermore, the isolation of the new natural product (2*S*,3*S*,7*S*,10*R*)-guaia-1,11-dien-10-ol from patchouli oil is reported.

### Introduction

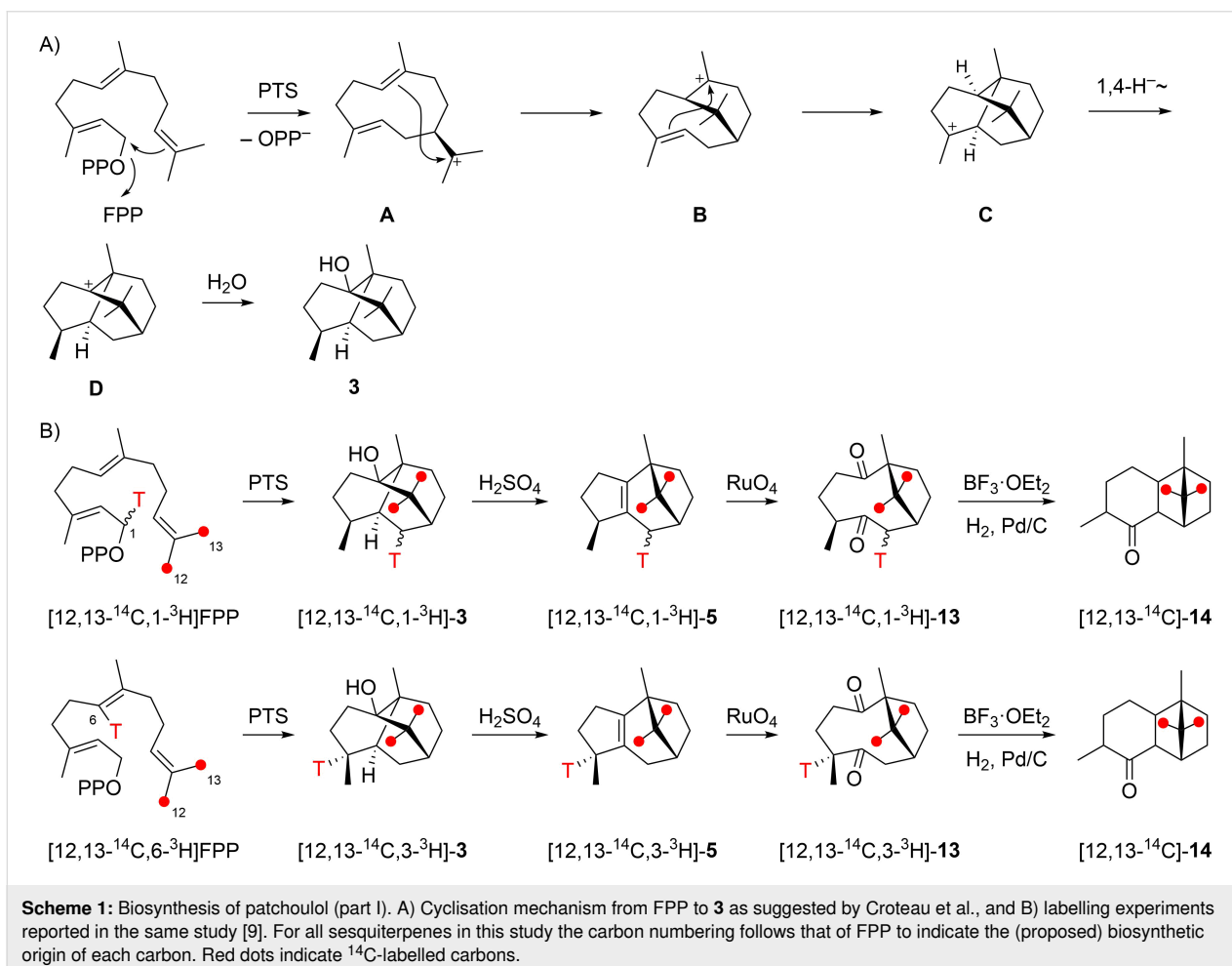
Patchouli oil, the essential oil of the shrub *Pogostemon cablin*, has a pleasant woody odour and is of high economic value for the perfumery and cosmetics industries. It is mainly composed of sesquiterpenes with patchoulol as the main compound (ca. 40%) [1,2]. Pure patchoulol is a crystalline material that has first been described by Gal in 1869 [3]. Its planar structure was initially described as that of compound **1** (Figure 1) by Treibs [4], and later reassigned to structure **2** based on a total synthesis from camphor by Büchi [5]. Because of an unexpected rearrangement this structural assignment was still erroneous, and

the correct structure **3** was finally established by X-ray analysis of its chromic acid diester [6]. The patchoulol synthase (PTS) has been purified from plant leaves and shown to convert farnesyl diphosphate (FPP) into compound **3** and several biogenetically related terpene hydrocarbons including  $\alpha$ -patchoulene (**4**),  $\beta$ -patchoulene (**5**),  $\alpha$ -bulnesene (**6**) and  $\alpha$ -guaiene (**7**) (Figure 1) [7]. The enzyme was subsequently made available by cDNA gene cloning, revealing germacrene A (**8**),  $\alpha$ -humulene (**9**), (*E*)- $\beta$ -caryophyllene (**10**), seychellene (**11**) and pogostol (**12**) as further side products [8].



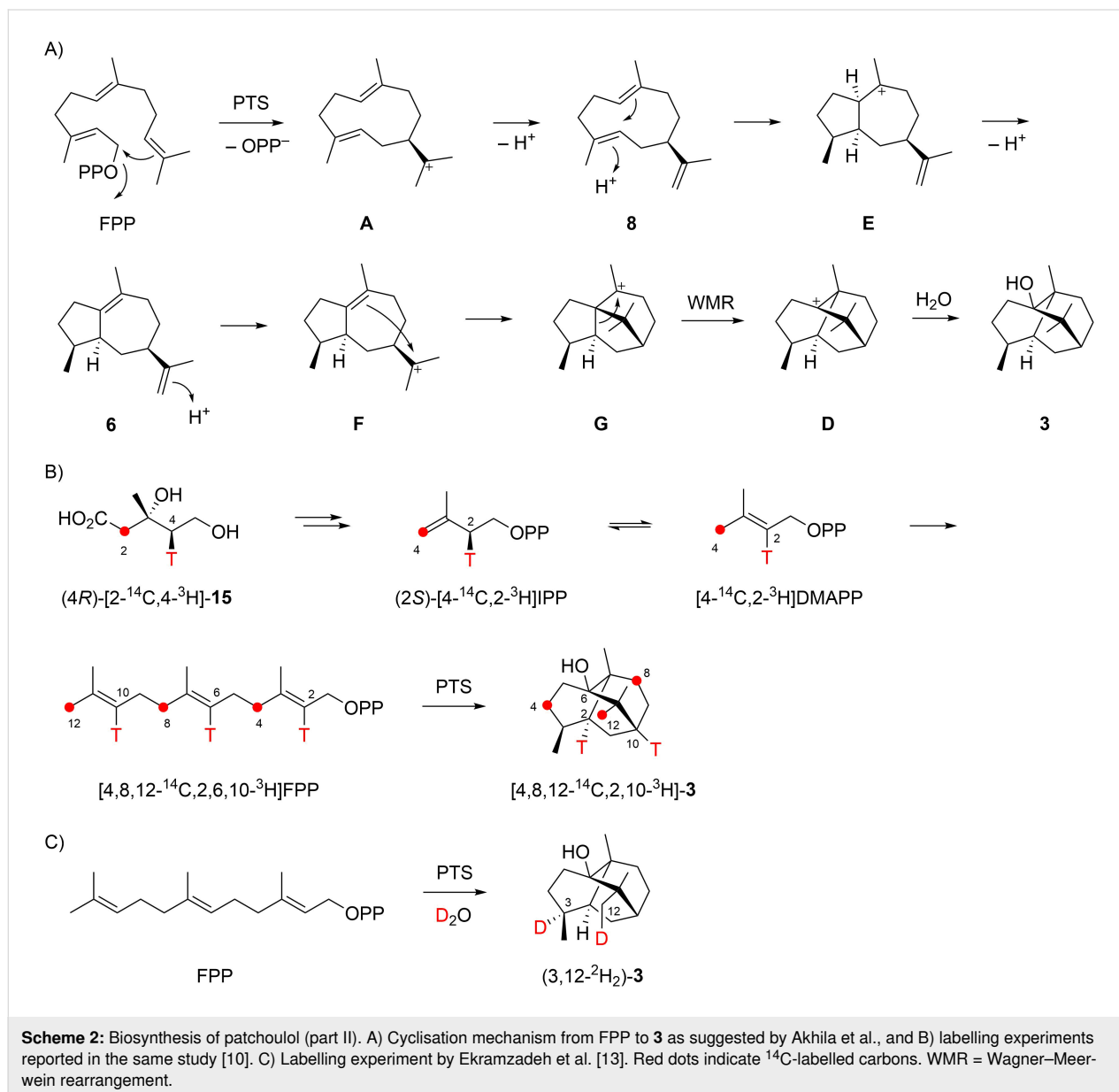
The biosynthetic mechanism of the formation of compound **3** was investigated by several groups through isotopic labelling experiments. In 1987, Croteau et al. have suggested a pathway through 1,10-cyclisation of FPP to the (*E,E*)-germacradienyl cation (**A**), followed by direct cyclisation reactions to **B** and **C**,

a 1,4-hydride shift to **D** and capture with water to yield **3** (Scheme 1A) [9]. This mechanism was supported by radioactive labelling experiments with [12,13-<sup>14</sup>C,1-<sup>3</sup>H]FPP and [12,13-<sup>14</sup>C,6-<sup>3</sup>H]FPP, whose enzymatic conversion with PTS into **3** proceeded with full retainment of the labelling in both



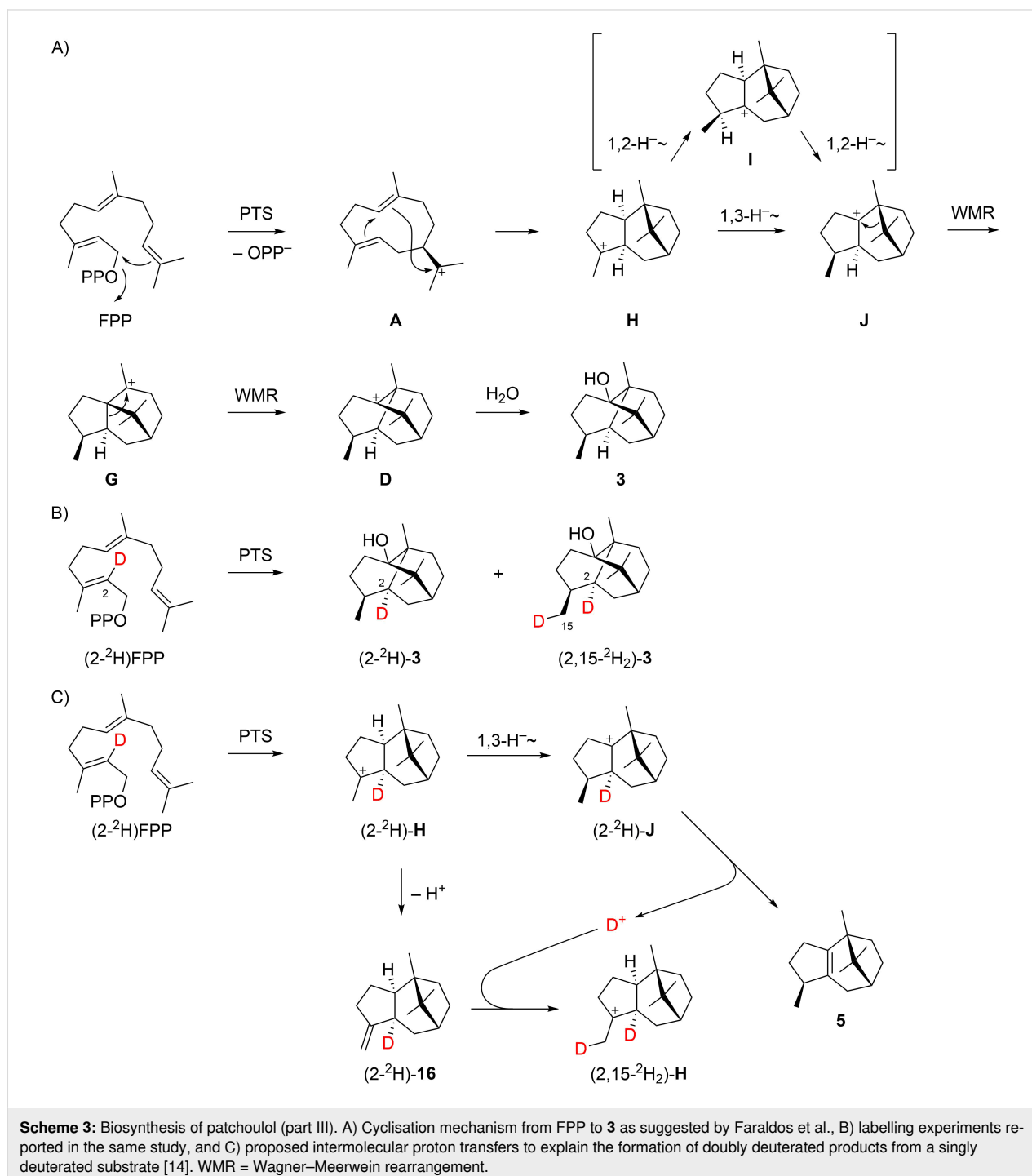
cases (Scheme 1B). Subsequent chemical degradation through acid catalysed conversion into **5**, oxidative cleavage to the diketone **13**,  $\text{BF}_3 \cdot \text{OEt}_2$  mediated ring closure by aldol reaction and catalytic hydrogenation gave **14**. For both experiments a full retainment of labelling was reported for all intermediates until **13**, while a loss of tritium was observed for **14** with both substrates. From these experiments it was concluded that the hydrogen H6 must migrate into another position, as realised by the 1,4-hydride shift from **C** to **D**. The loss of  $^3\text{H}$  in the experiment with  $[12,13\text{-}^{14}\text{C},1\text{-}^3\text{H}]\text{FPP}$  was expected for the aldol reaction of **13**, but is more difficult to understand in the experiment with  $[12,13\text{-}^{14}\text{C},6\text{-}^3\text{H}]\text{FPP}$ . In this case the loss of  $^3\text{H}$  was explained by an exchange against  $^1\text{H}$  during catalytic hydrogenation [9].

One year later, Akhila et al. proposed an alternative biosynthetic mechanism that also starts with a cyclisation of FPP to **A** (Scheme 2A) [10], but then a subsequent deprotonation to **8**, an important neutral intermediate in the biosynthesis of many sesquiterpenes [11], is assumed. A reprotonation-induced cyclisation leads to **E** that is again deprotonated to **6**, followed by another reprotonation to **F**, cyclisation to **G** and Wagner–Meerwein rearrangement to **D**, the same final intermediate as suggested by Croteau. This mechanism was supported by feeding experiments with  $(4R)\text{-}[2\text{-}^{14}\text{C},4\text{-}^3\text{H}]\text{mevalonic acid}$  (**15**) that is converted through IPP and DMAPP into FPP (Scheme 2B). According to the FPP biosynthesis as established by Cornforth and co-workers, these reactions should proceed with full retainment of all labellings [12]. For isolated **3** a loss of one of the three  $^3\text{H}$



atoms was reported that is explainable by the deprotonation step from **E** to **6** [10], but contradicts the retainment of this hydrogen as reported by Croteau [9]. Further support for Akhila's mechanism was provided by Ekramzadeh et al., who observed the uptake of two deuterium atoms at C3 and C12 in an incubation of FPP with PTS in deuterium oxide buffer that explain the protonations of the neutral intermediates **8** and **6** (Scheme 2C) [13].

In 2010, Faraldos et al. published a third mechanism that also starts with a cyclisation of FPP to **A** (Scheme 3A) [14]. Similar to Croteau's mechanism, **A** is directly further cyclised to **H**, followed by a 1,3-hydride shift to **J** (equivalent to the 1,4-hydride migration from **C** to **D** in Scheme 1A), and a Wagner–Meerwein rearrangement to **G**. The final steps are identical to those in Akhila's mechanism (Scheme 2A). This work also reported on a labelling experiment with (2-<sup>2</sup>H)FPP that was enzymati-



cally converted with PTS with incorporation of deuterium at C2 of **3** (Scheme 3B). This result ruled out that the 1,3-hydride shift from **H** to **J** must be replaced by two sequential 1,2-hydride transfers via **I**, but cannot discriminate between the Croteau's and Akhila's mechanistic alternatives. In addition, the formation of doubly labelled (2,15-<sup>2</sup>H<sub>2</sub>)-**3** from (2-<sup>2</sup>H)FPP was reported, which was explained by an unusual intramolecular deuterium transfer. Herein, the deuterium is released from (2-<sup>2</sup>H)-**J** in the deprotonation step to **5** (or other enzyme products losing the same hydrogen in the terminal deprotonation). Deprotonation of (2-<sup>2</sup>H)-**H** was suggested to produce the unknown sesquiterpene (2-<sup>2</sup>H)-**16** that may take up the deuterium released in the formation of **5** (and similar compounds) to give (2,15-<sup>2</sup>H<sub>2</sub>)-**H** (Scheme 3C).

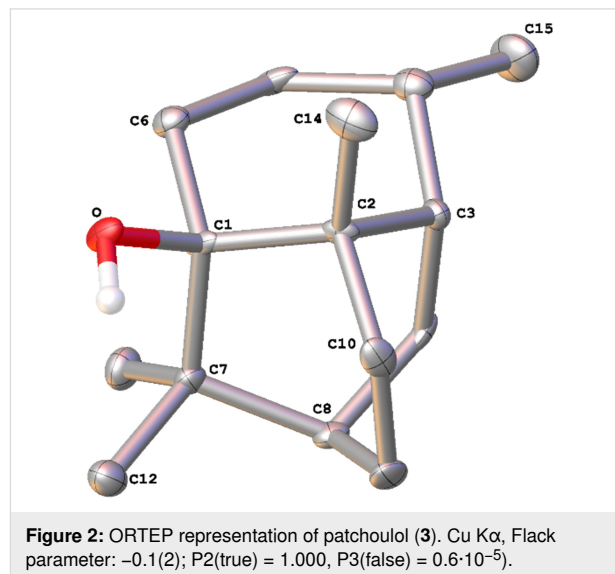
Notably, none of the proposed mechanisms in Schemes 1–3 can explain the reported results from all labelling experiments and some of the reported findings are even contradictory. For this reason, we have reinvestigated the enzyme mechanism of PTS in isotopic labelling experiments through methods recently developed in our laboratory that make use of <sup>13</sup>C and <sup>2</sup>H-substituted terpene precursors, and by DFT calculations. The general strategy in these experiments is to use substrates or substrate combinations so that deuterium migrations end at <sup>13</sup>C-labelled carbons, resulting in triplet signals in the <sup>13</sup>C NMR spectra [15,16]. Moreover, deuterium atoms ending in neighbouring positions of <sup>13</sup>C-labelled carbons become evident from slight upfield shifted <sup>13</sup>C NMR signals. These experiments and the DFT calculations were not only carried out in a way to gain support for one mechanism, but also to disprove some of the earlier reports in order to resolve the contradictions in the literature.

## Results and Discussion

### Absolute configurations of patchouliol and pogostol

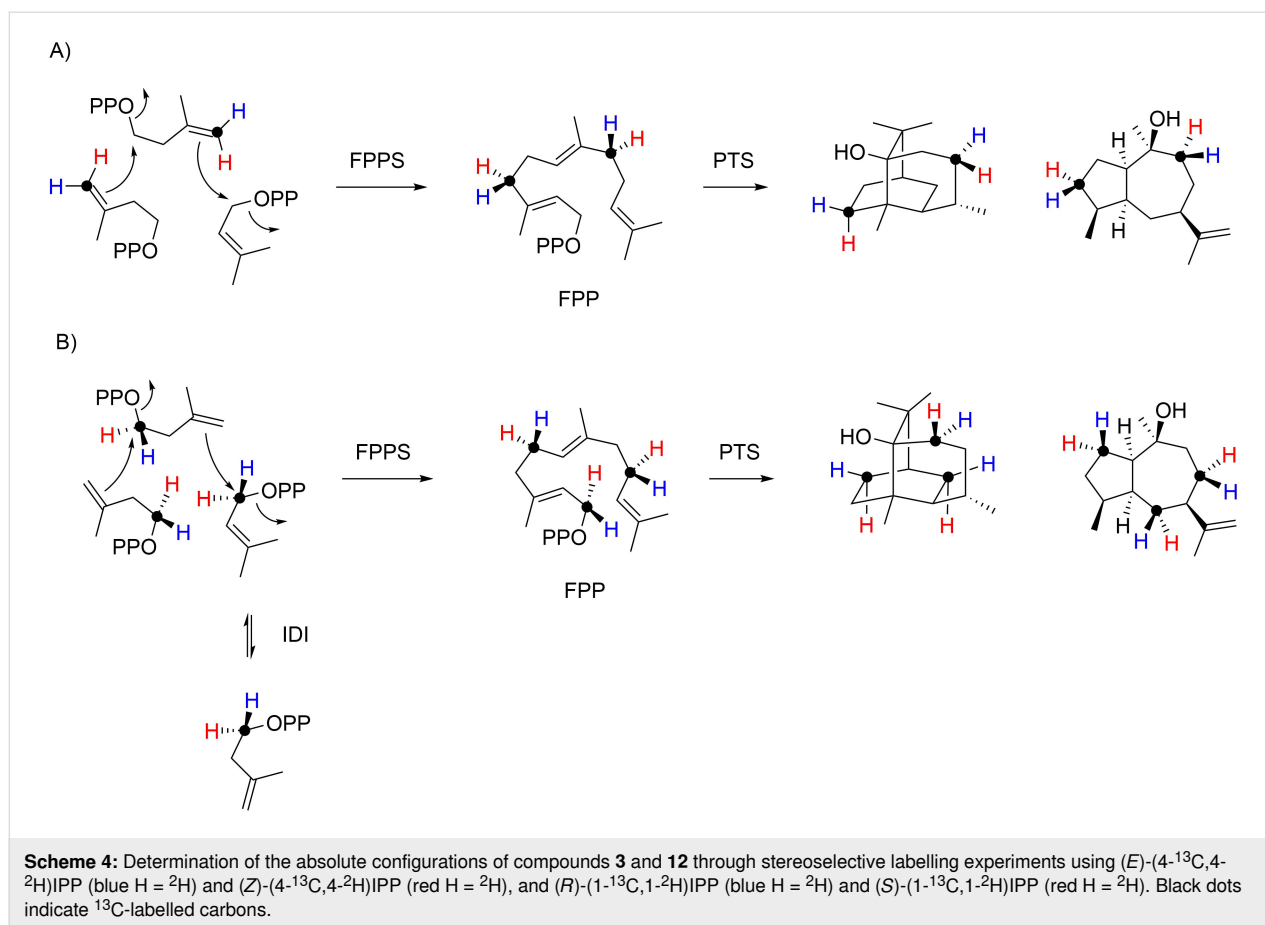
In order to reinvestigate the biosynthesis of patchouliol (**3**) the synthetic gene for patchouliol synthase from *P. cablin* was cloned and expressed in *Escherichia coli* [8]. The purified protein (Figure S1 in Supporting Information File 1) converted FPP into **3** as the main product, besides several side products (see Figure S2 in Supporting Information File 1). A reference sample of **3** was isolated from patchouli oil and its structure was confirmed by NMR spectroscopy (Supporting Information File 1, Table S1 and Figures S3–S10). So far, only the X-ray structure of the chromate diester [6] and a Mo K $\alpha$  structure of **3** were reported (CCDC no. 1491695) [17], but these data did not allow to conclude on the absolute configuration of compound **3**. We now obtained **3** as a crystalline material and performed an X-ray structural analysis through anomalous dispersion using

Cu K $\alpha$  irradiation (Table S2 in Supporting Information File 1), resulting in the structure of **3** with the absolute configuration as shown in Figure 2.



The absolute configuration of **3** was furthermore independently confirmed through a stereoselective deuteration strategy (Scheme 4; all labelling experiments of this study are summarised in Supporting Information File 1, Table S3). Using dimethylallyl diphosphate (DMAPP) and (*E*)- and (*Z*)-(4-<sup>13</sup>C,4-<sup>2</sup>H)isopentenyl diphosphate (IPP) [18] in conjunction with FPP synthase (FPPS) from *Streptomyces coelicolor* [19] and PTS (Supporting Information File 1, Figure S11), stereogenic centres of known configuration are introduced at the deuterated carbons. The NOESY-based assignment of the diastereotopic hydrogens at these carbons for the unlabelled compound then allows to conclude on the absolute configuration of alcohol **3**. A second set of experiments made use of (*R*)- and (*S*)-(1-<sup>13</sup>C,1-<sup>2</sup>H)IPP [20] that were enzymatically converted with isopentenyl diphosphate isomerase (IDI) from *E. coli* [20,21], FPPS, and PTS (Figure S12 in Supporting Information File 1). The additional <sup>13</sup>C-labellings in these experiments serve for a sensitive monitoring of deuterium incorporation through HSQC spectroscopy. All X-ray and labelling experiments confirmed the absolute configuration of **3** as reported previously.

It is reasonable to assume that pogostol (**12**) as a side product of PTS has the absolute configuration as shown in Figure 1, but surprisingly its absolute configuration has never been formally established. Moreover, pogostol registered under the CAS number 21698-41-9 is even assigned the opposite absolute configuration as expected from these biosynthetic considerations. After a recent correction [22] of its initially reported relative configuration [23] that was shown to be erroneous by total synthesis



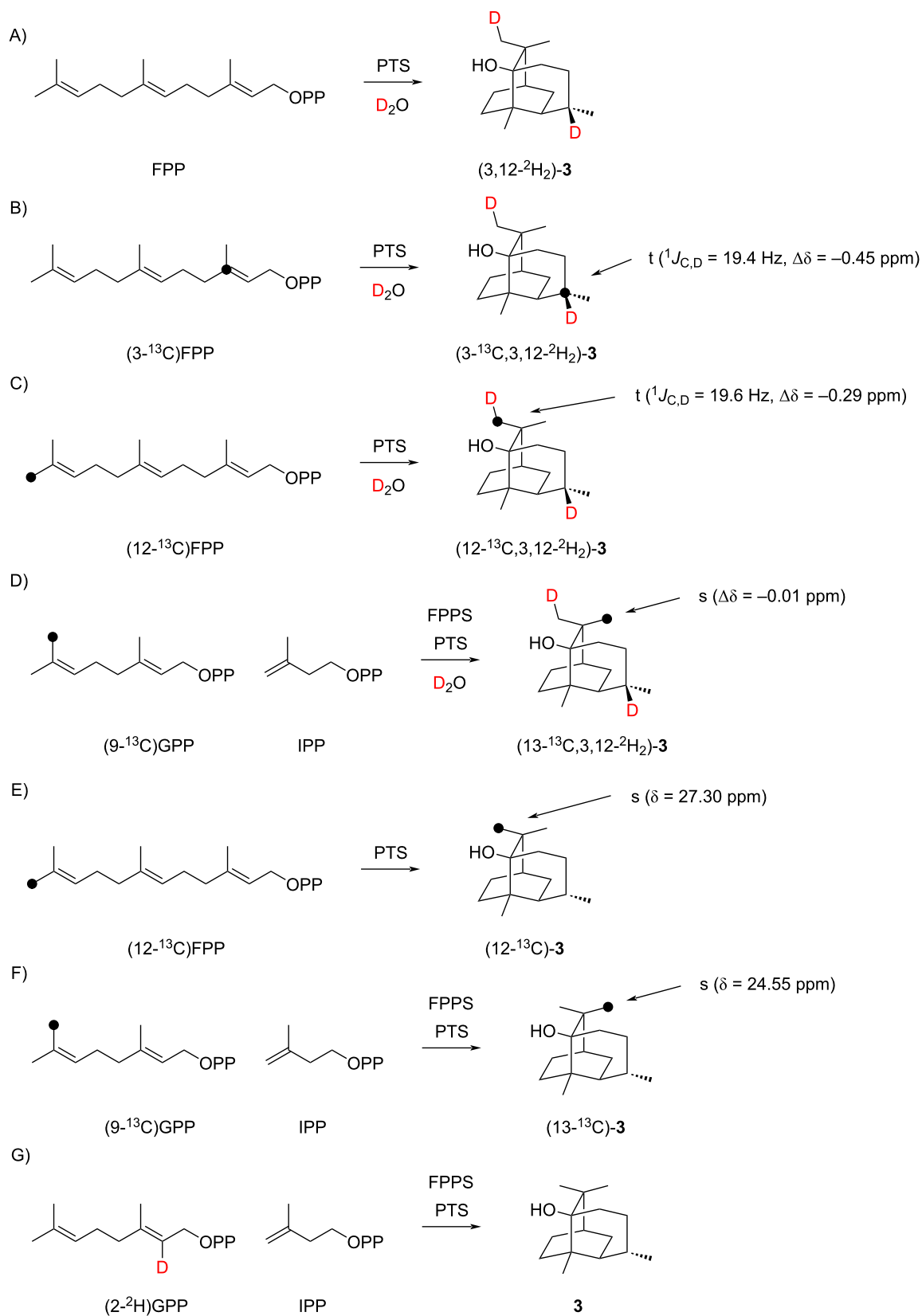
[24], we now address the problem of the absolute configuration of **12** experimentally. For this purpose, compound **12** was re-isolated from patchouli oil and its NMR data were fully assigned (see Supporting Information File 1, Table S4 and Figures S13–S20). Since **12** is also a side product of PTS, the data obtained from the above described labelling experiments were then used to determine the absolute configuration of pogostol as (*1R,4S,5S,7R,10S*)-**12** (Scheme 4, and Figures S21 and S22 in Supporting Information File 1).

### Investigations on patchoulol biosynthesis by labelling experiments

The cyclisation mechanism from FPP to patchoulol (**3**) was investigated in isotopic labelling experiments. Our aim was not only to obtain results that can support one of the three mechanisms under discussion in the literature, but because of the partially contradictory findings also to perform experiments that may disprove some of the proposed mechanisms, in order to obtain a refined understanding of the biosynthesis of compound **3**.

In a first experiment, repeating earlier findings by Ekramzadeh et al. [13], the uptake of deuterium during an incubation of FPP

with PTS in deuterium oxide buffer was investigated, revealing incorporation of two deuterium atoms into **3** (Scheme 5A and Figure S23 in Supporting Information File 1). This result is in agreement with the mechanism proposed by Akhila et al. (Scheme 2) [10], but not with the alternative mechanisms of Scheme 1 and Scheme 3, and therefore the next experiments focussed on gaining further evidence for the mechanism of Scheme 2. The site of incorporation for the deuterium uptake was evident from incubations of (3-<sup>13</sup>C)FPP and (12-<sup>13</sup>C)FPP [25] in deuterium oxide (Scheme 5B and 5C). The <sup>13</sup>C NMR analysis of the obtained products showed slightly upfield-shifted triplets for C3 ( $\Delta\delta = -0.45$  ppm,  $J = 19.4$  Hz) and C12 ( $\Delta\delta = -0.29$  ppm,  $J = 19.6$  Hz) as a result of <sup>1</sup>J<sub>C,D</sub> couplings (see Figure S24B and S24C in Supporting Information File 1), again in full agreement with the mechanism by Akhila et al. [10]. A control experiment with (13-<sup>13</sup>C)FPP, enzymatically prepared from (9-<sup>13</sup>C)GPP [26], and IPP with FPPS, resulted in a singlet with a very small upfield shift ( $\Delta\delta = -0.01$  ppm) in the <sup>13</sup>C NMR (Scheme 5D and Figure S24D in Supporting Information File 1), indicating a deuterium incorporation two positions away from C13 and a clear stereochemical course for the geminal methyl groups C12 and C13 of FPP. This was also confirmed through <sup>13</sup>C-labelling experiments with (12-<sup>13</sup>C)FPP



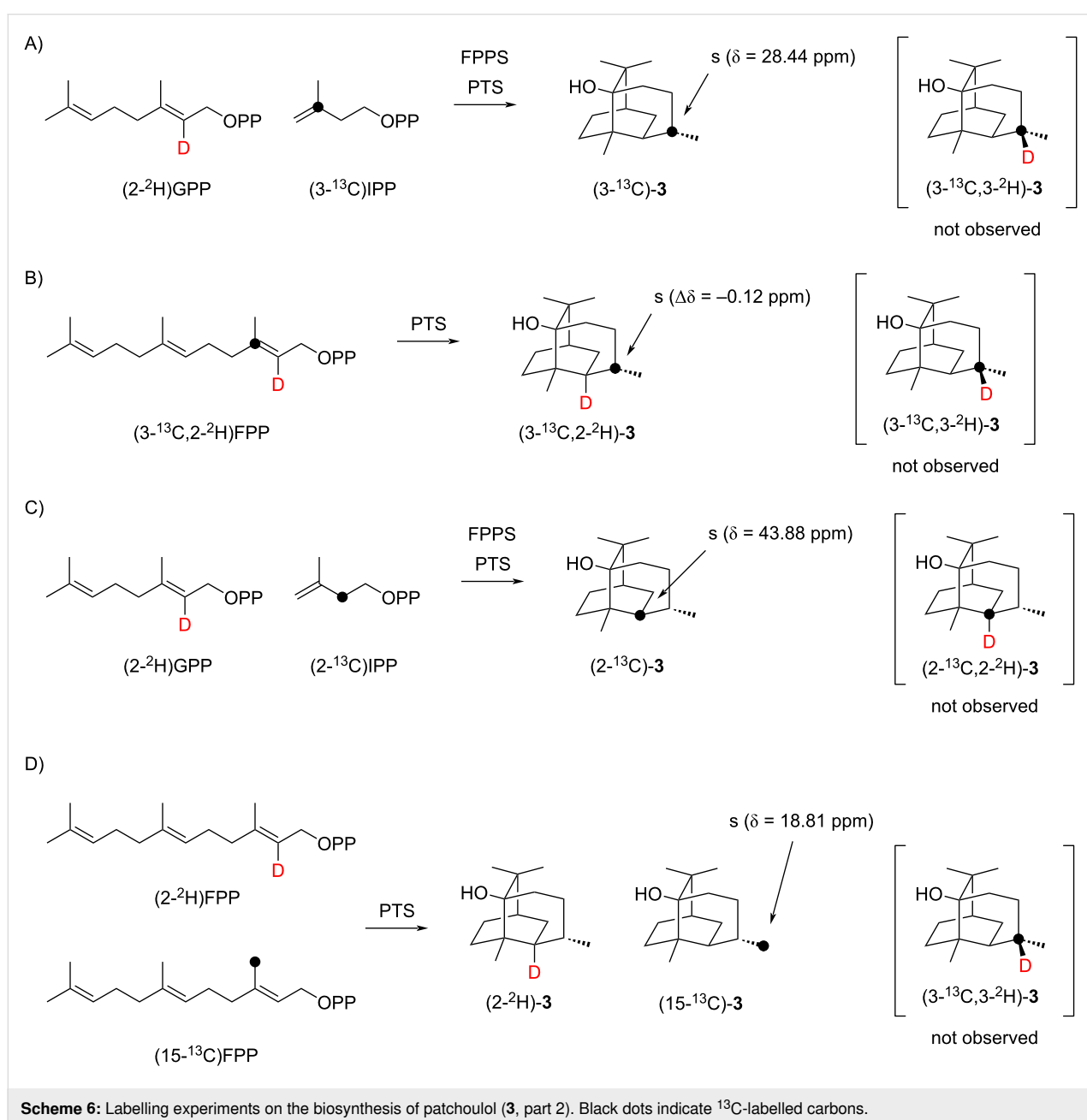
**Scheme 5:** Labelling experiments on the biosynthesis of patchouliol (3, part 1). Black dots indicate <sup>13</sup>C-labelled carbons.



and (9-<sup>13</sup>C)GPP plus IPP in non-deuterated aqueous environment (Scheme 5E and 5F and Figure S24E and S24F in Supporting Information File 1). These experiments together with a detailed inspection of the NOESY spectrum of **3** also indicated that the assigned sites of incorporation of labellings from C12 of FPP by Akhila et al. (Scheme 2B) and by Ekramzadeh et al. (Scheme 2C) must be corrected, i.e., the carbons in **3** derived from the geminal Me groups C12 and C13 of FPP must be exchanged. Finally, the mechanism proposed by Akhila et al. includes a deprotonation step from C6 of FPP towards the neutral intermediate **6**. The enzymatic conversion of (2-<sup>2</sup>H)GPP [26] and IPP with FPPS into (6-<sup>2</sup>H)FPP and its subsequent

cyclisation with PTS resulted in the formation of non-labelled **3**, in agreement with this deprotonation step (Scheme 5G and Figure S25 in Supporting Information File 1).

Notably, not only the double deuterium uptake into **3** from D<sub>2</sub>O, but also the loss of deuterium from (6-<sup>2</sup>H)FPP contradicts the mechanisms of Scheme 1 and Scheme 3 that both propose a migration of hydrogen from C6 to C3, either through a 1,4- or a 1,3-hydride shift. An additional experiment with (2-<sup>2</sup>H)GPP, (3-<sup>13</sup>C)IPP [27], FPPS, and PTS produced a clear singlet in the <sup>13</sup>C NMR spectrum for C3 of compound **3** (Scheme 6A and Figure S26 in Supporting Information File 1). Thus, there is no

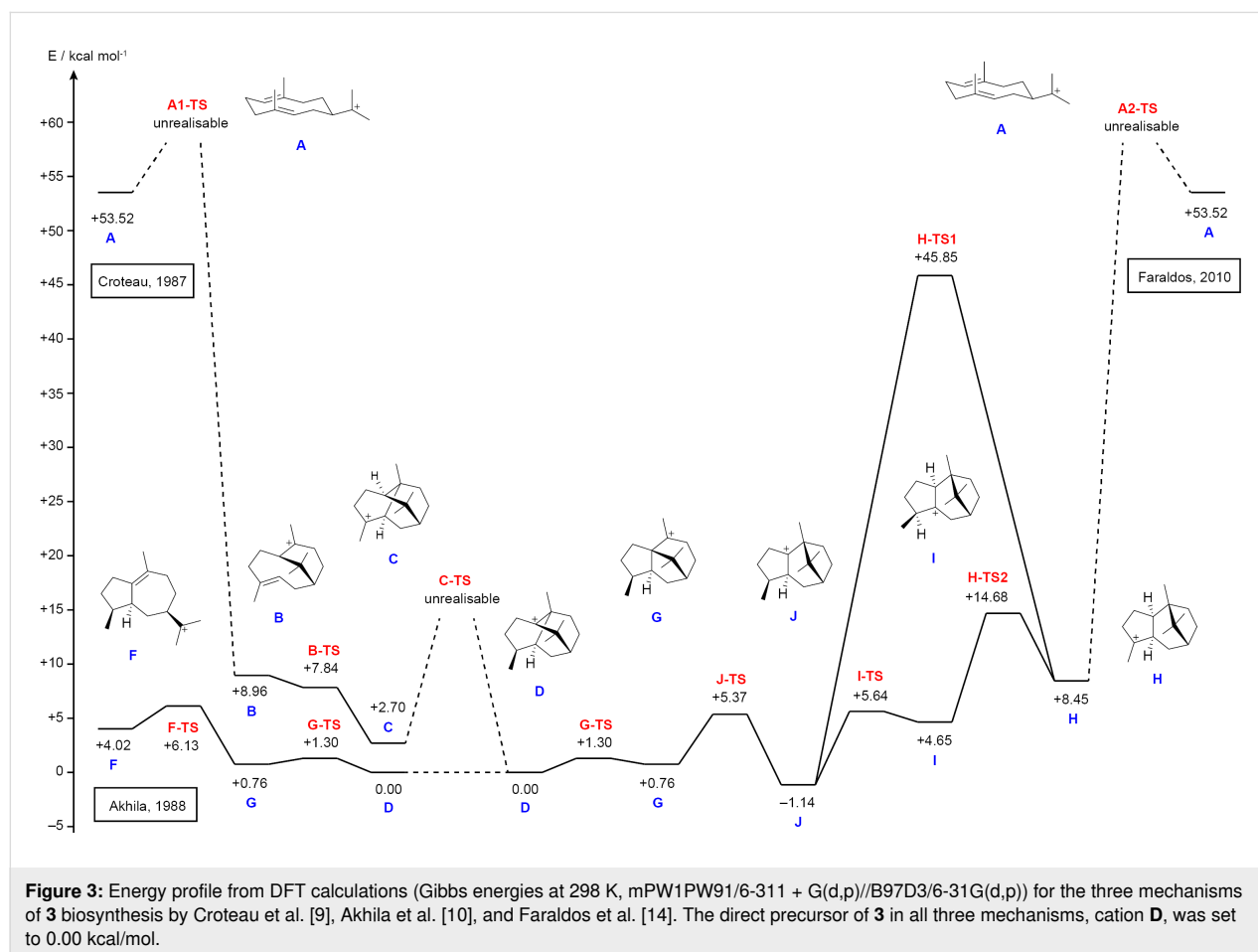


evidence, also not for a minor participation, for the proposed 1,4- or 1,3-hydride shifts of Scheme 1 and Scheme 3. Despite their proposal of a 1,3-hydride transfer for the conversion of **H** to **J** (Scheme 3), Faraldos et al. have pointed out that instead two sequential 1,2-hydride migrations through **I** would be easier to understand [14]. To investigate whether such alternative 1,2-hydride shifts take part, incubation experiments were performed with (3- $^{13}\text{C}$ ,2- $^2\text{H}$ )FPP [28] plus PTS, and with (2- $^2\text{H}$ )GPP and (2- $^{13}\text{C}$ )IPP [27] plus FPPS and PTS (Scheme 6B and 6C), but in both cases the product analysis by  $^{13}\text{C}$  NMR spectroscopy showed only singlet signals for C3 and C2 of **3**, respectively (in the first case associated with a small upfield shift of  $\Delta\delta = -0.12$  ppm as a result of deuterium in the neighbouring position to C3, Figure S27 in Supporting Information File 1). Moreover, no triplet signals indicative for a direct  $^{13}\text{C}$ - $^2\text{H}$  bond were observed, ruling out the participation of two sequential 1,2-hydride shifts in the **H** to **J** transformation. To re-investigate the suggested intermolecular proton exchange in the biosynthesis of **3** (Scheme 3C) [14], an incubation experiment with the mixed substrates (2- $^2\text{H}$ )FPP [27] and (15- $^{13}\text{C}$ )FPP [24] was performed (Scheme 6D). Their conversion with PTS only resulted in a singlet for labelled C15 of **3** in the

$^{13}\text{C}$  NMR spectrum, but no upfield-shifted triplet (Figure S28, Supporting Information File 1), demonstrating that the hypothetical intermolecular proton shift does not take place. Instead, an additional incorporation of deuterium into the substrate during synthesis, contaminating the target compound (2- $^2\text{H}$ )FPP with some (2,15- $^2\text{H}_2$ )FPP, seems to be the more likely explanation for the deuteration of **3** at C15 observed by Faraldos et al. [14]. This can also much better explain the deuterium content observed by GC-MS in the PTS products that were proposed to transfer deuterium to **3**. Following the mechanism of Scheme 3C, compounds such as **5**, if indeed obtained from pure (2- $^2\text{H}$ )FPP, should not show any residual deuterium content, if they donate their deuterium to **3**. The analytical data in reference [14] in fact show that **5** does contain deuterium, only one deuterium atom less than in **3**, but this deuterium loss for **5** is best explained by the terminal deprotonation step from C6.

### Investigations on patchouliol biosynthesis by DFT calculations

The biosynthesis of **3** was also investigated by DFT calculations (Figure 3). For the mechanism proposed by Croteau et al. (Scheme 1) [9], the cyclisation of the (*E,E*)-germacradienyl



cation (**A**) to **B** could not be realised. The further reaction of **B** to **C** is barrierless, but the proposed 1,4-hydride shift to **D** is geometrically impossible and also cannot be realised by computations.

For the formation of the bicyclic cation **E** according to Akhila et al. (Scheme 2) [10] DFT calculations have been performed previously by us as part of a general study on guaiane sesquiterpenes from germacrene A (**8**) [29]. After reprotonation of the neutral intermediate **6** to **F** the next cyclisation to **G** and Wagner–Meerwein rearrangement to **D** can be realised with low TS barriers.

Also for the cyclisation of **A** to **H** as suggested by Faraldos et al. [14] the DFT calculations showed a strong steric repulsion that cannot be realised computationally. The 1,3-hydride shift from **H** to **J** is associated with a very high TS barrier (37.4 kcal/mol), while the sequence of two 1,2-hydride migrations via **I** to **J** would indeed be much easier. The final transformations involving two Wagner–Meerwein rearrangements through **G** and **D** can proceed smoothly.

### Isolation of guaia-1,11-dien-1-ol from patchouli oil

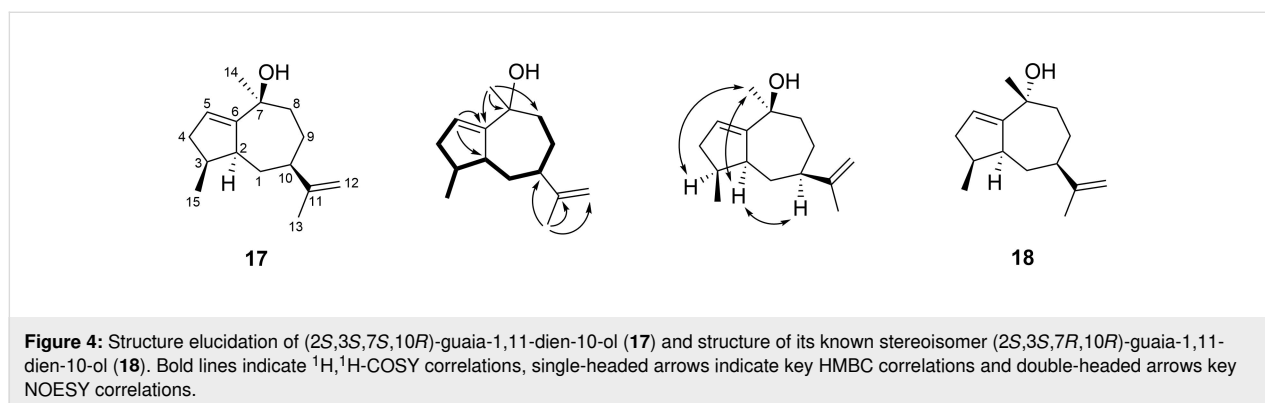
Fractionation of patchouli oil by column chromatography resulted in the isolation of the new natural product **17** [HRMS–ESI ( $m/z$ ): 221.1904 [ $M + H$ ]<sup>+</sup>, calculated for  $C_{15}H_{25}O^+$  221.1900 and  $[\alpha]_D^{25} = -7.7$ , ( $c$  0.26, benzene)] whose structure was elucidated by NMR spectroscopy (Table 1 and Figures S29–S35 in Supporting Information File 1). The  $^{13}C$  NMR spectrum showed signals for 15 carbons, including three Me groups, four olefinic carbons (two quaternary, one CH and one  $CH_2$ ), and a tertiary alcohol, suggesting the structure of an oxidised (dehydrogenated) bicyclic sesquiterpene alcohol (Figure 4). The  $^1H$ ,  $^1H$ -COSY spectrum revealed one large contiguous spin system C-2-3-4(15)-5-6-7-8-9. HMBC correlations from  $H_3$ -13 to C-7, C-11, and C-12 indicated an isopropenyl group attached to C-7, while additional HMBC

correlations from  $H_3$ -14 to C-8, C-9, and C-10 and from H-2 to C-3, C-5, and C-10 completed the planar structure of **17**. Key NOESY correlations from H-4 and H-5 to  $H_3$ -14 and from H-5 to H-7 placed these groups on one hemisphere of the molecule, revealing the structure of guaia-1,11-dien-10-ol. Based on the very likely biosynthetic relationship to the products of patchoulol synthase (especially **12**, Figure 1), the absolute con-

**Table 1:** NMR data of compound **17** (700 MHz,  $C_6D_6$ ).

C <sup>a</sup>	type	$^{13}C$	$^1H$
1	$CH_2$	35.62	1.75 (m, $H_\alpha$ ) 1.13 (ddd, 12.4, 12.4, 12.4, $H_\beta$ )
2	CH	47.84	2.06 (dddd, $J = 12.3, 7.3, 2.1, 2.1$ )
3	CH	40.02	2.23 (dddq, $J = 10.1, 7.2, 7.2, 7.2$ )
4	$CH_2$	38.13	2.15 (ddd, $J = 15.6, 7.7, 3.0, H_\alpha$ ) 1.89 (dddd, $J = 15.6, 10.1, 1.8, 1.8, H_\beta$ )
5	CH	124.12	5.69 (dd, $J = 2.9, 1.8$ )
6	$C_q$	159.17	–
7	$C_q$	73.38	–
8	$CH_2$	42.68	1.76 (m, $H_\beta$ ) 1.47 (m, $H_\alpha$ )
9	$CH_2$	29.20	1.53 (m, $H_\alpha$ ) 1.46 (m, $H_\beta$ )
10	CH	50.80	1.87 (m)
11	$C_q$	151.90	–
12	$CH_2$	108.87	4.80 (m, $H_Z$ ) 4.74 (m, $H_E$ )
13	$CH_3$	20.70	1.65 (dd, $J = 1.3, 0.8$ )
14	$CH_3$	32.45	1.26 (s)
15	$CH_3$	15.47	0.91 (d, $J = 7.0$ )

<sup>a</sup>Carbon numbering as shown in Figure 4; <sup>b</sup>multiplicities are indicated by s = singlet, d = doublet, q = quartet, m = multiplet; coupling constants  $J$  are given in hertz.



figuration was tentatively assigned as (4*S*,5*S*,7*R*,10*S*)-**17**. A stereoisomer of this compound, (4*S*,5*S*,7*R*,10*R*)-**18**, was reported before from *Hyptis suaveolens* [30] and has been obtained by synthesis from  $\alpha$ -bulnesene (**6**) with an optical rotation of  $[\alpha]_{\text{D}}^{29} = -79.2$  (*c* 0.25, CHCl<sub>3</sub>) [31,32]. The negative optical rotation of **17** ( $[\alpha]_{\text{D}}^{25} = -7.7$  (*c* 0.26, C<sub>6</sub>D<sub>6</sub>)) in comparison to the negative optical rotation of **18** further supports the tentatively assigned absolute configuration for **17**.

## Conclusion

Different contradictory mechanisms for patchoulol biosynthesis have been discussed in the literature. The present study resolves this situation through isotopic labelling experiments. These experiments support the passage of two neutral intermediates, germacrene A and  $\alpha$ -bulnesene, that become reactivated by reprotonations, as shown by incubation experiments in deuterium oxide buffer. These observations are in line with the proposed mechanisms by Akhila et al. [10] and Ekramzadeh et al. [13], with minor corrections regarding the stereochemical course of the geminal Me groups of FPP. Because it is possible that multiple mechanisms operate simultaneously, we also performed experiments to exclude other proposals made by Croteau et al. [9] and Faraldos et al. [14]. For this purpose, <sup>13</sup>C-labelled substrates were used in conjunction with deuterium labelling. These substrates have the advantage that the incorporation of labelling can be detected and localised through <sup>13</sup>C NMR spectroscopy with very high sensitivity, but no hints for critical steps such as 1,3- or 1,4-hydride shifts or intermolecular deuterium transfers as suggested in these studies were obtained. The results from labelling experiments are furthermore fully supported by DFT calculations. Our computational work also demonstrated that the mechanisms by Croteau et al. [9] and Faraldos et al. [14] are difficult to understand, while the mechanism by Akhila et al. [10] can proceed via low transition state barriers. As discussed above, the mistake in the mechanistic work by Faraldos et al. [14] seems to reside in an impure starting material (2-<sup>2</sup>H)FPP containing additional deuterium at C15, but it is difficult to understand the results by Croteau et al. [9]. As a general comment we can only state, how difficult it was to perform the old work using radioactive labellings, especially in terms of localising the site of incorporation by chemical degradations. It should be emphasised how fascinating and how deep the insights of many of such studies are. Today <sup>13</sup>C and <sup>2</sup>H-labellings in conjunction with NMR and MS-based analysis can be used, with strong advantages over radioactive labellings, not only from a safety perspective, but also with respect to the ease of data interpretation. Overall, our study gives another example of terpene biosynthesis through neutral intermediates, and more specifically another example of sesquiterpene biosynthesis through the widespread biosynthetic intermediate germacrene A [11].

## Supporting Information

### Supporting Information File 1

Experimental details, characterisation data and copies of spectra.

[<https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-18-2-S1.pdf>]

## Acknowledgements

We thank Andreas Schneider for HPLC purifications and the computing centre of the University of Cologne (RRZK), providing CPU time on the DFG-funded supercomputer CHEOPS.

## Funding

This work was funded by the DFG (DI1536/7-2).

## ORCID® iDs

Houchao Xu - <https://orcid.org/0000-0002-4480-2035>

Gregor Schnakenburg - <https://orcid.org/0000-0001-6489-2106>

Jeroen S. Dickschat - <https://orcid.org/0000-0002-0102-0631>

## References

- Swamy, M. K.; Sinniah, U. R. *Molecules* **2015**, *20*, 8521–8547. doi:10.3390/molecules20058521
- van Beek, T. A.; Joulain, D. *Flavour Fragrance J.* **2018**, *33*, 6–51. doi:10.1002/ffj.3418
- Gal, H. *Justus Liebigs Ann. Chem.* **1869**, *150*, 374–376. doi:10.1002/jlac.18691500310
- Treibs, W. *Justus Liebigs Ann. Chem.* **1949**, *564*, 141–151. doi:10.1002/jlac.19495640207
- Büchi, G.; MacLeod, W. D. *J. Am. Chem. Soc.* **1962**, *84*, 3205–3206. doi:10.1021/ja00875a047
- Dobler, M.; Dunitz, J. D.; Gubler, B.; Weber, H. P.; Büchi, G.; Padilla, O. J. *Proc. Chem. Soc., London* **1963**, 383.
- Munck, S. L.; Croteau, R. *Arch. Biochem. Biophys.* **1990**, *282*, 58–64. doi:10.1016/0003-9861(90)90086-e
- Deguerry, F.; Pastore, L.; Wu, S.; Clark, A.; Chappell, J.; Schalk, M. *Arch. Biochem. Biophys.* **2006**, *454*, 123–136. doi:10.1016/j.abb.2006.08.006
- Croteau, R.; Munck, S. L.; Akoh, C. C.; Fisk, H. J.; Satterwhite, D. M. *Arch. Biochem. Biophys.* **1987**, *256*, 56–68. doi:10.1016/0003-9861(87)90425-5
- Akhila, A.; Sharma, P. K.; Thakur, R. S. *Phytochemistry* **1988**, *27*, 2105–2108. doi:10.1016/0031-9422(88)80105-5
- Xu, H.; Dickschat, J. S. *Chem. – Eur. J.* **2020**, *26*, 17318–17341. doi:10.1002/chem.202002163
- Cornforth, J. W.; Cornforth, R. H.; Popják, G.; Yengoyan, L. *J. Biol. Chem.* **1966**, *241*, 3970–3987. doi:10.1016/s0021-9258(18)99800-5
- Ekramzadeh, K.; Brämer, C.; Frister, T.; Fohrer, J.; Kirschning, A.; Scheper, T.; Beutel, S. *Biotechnol. Prog.* **2020**, *36*, e2935. doi:10.1002/btpr.2935

14. Faraldos, J. A.; Wu, S.; Chappell, J.; Coates, R. M. *J. Am. Chem. Soc.* **2010**, *132*, 2998–3008. doi:10.1021/ja909251r
15. Rabe, P.; Pahirulzaman, K. A. K.; Dickschat, J. S. *Angew. Chem., Int. Ed.* **2015**, *54*, 6041–6045. doi:10.1002/anie.201501119
16. Dickschat, J. S. *Angew. Chem., Int. Ed.* **2019**, *58*, 15964–15976. doi:10.1002/anie.201905312
17. Inoue, Y. *IUCrData* **2017**, *2*, x170189. doi:10.1107/s2414314617001894
18. Lauterbach, L.; Rinkel, J.; Dickschat, J. S. *Angew. Chem., Int. Ed.* **2018**, *57*, 8280–8283. doi:10.1002/anie.201803800
19. Rabe, P.; Rinkel, J.; Nubbemeyer, B.; Köllner, T. G.; Chen, F.; Dickschat, J. S. *Angew. Chem., Int. Ed.* **2016**, *55*, 15420–15423. doi:10.1002/anie.201608971
20. Rinkel, J.; Dickschat, J. S. *Org. Lett.* **2019**, *21*, 2426–2429. doi:10.1021/acs.orglett.9b00725
21. Hahn, F. M.; Hurlburt, A. P.; Poulter, C. D. *J. Bacteriol.* **1999**, *181*, 4499–4504. doi:10.1128/jb.181.15.4499-4504.1999
22. Amand, S.; Langenfeld, A.; Blond, A.; Dupont, J.; Nay, B.; Prado, S. *J. Nat. Prod.* **2012**, *75*, 798–801. doi:10.1021/np2009913
23. Weyerstahl, P.; Marschall, H.; Splittgerber, U.; Wolf, D. *Flavour Fragrance J.* **2000**, *15*, 153–173. doi:10.1002/1099-1026(200005/06)15:3<153::aid-ffj885>3.0.co;2-5
24. Booker-Milburn, K. I.; Jenkins, H.; Charmant, J. P. H.; Mohr, P. *Org. Lett.* **2003**, *5*, 3309–3312. doi:10.1021/ol035373u
25. Rabe, P.; Barra, L.; Rinkel, J.; Riclea, R.; Citron, C. A.; Klapschinski, T. A.; Janusko, A.; Dickschat, J. S. *Angew. Chem., Int. Ed.* **2015**, *54*, 13448–13451. doi:10.1002/anie.201507615
26. Bian, G.; Rinkel, J.; Wang, Z.; Lauterbach, L.; Hou, A.; Yuan, Y.; Deng, Z.; Liu, T.; Dickschat, J. S. *Angew. Chem., Int. Ed.* **2018**, *57*, 15887–15890. doi:10.1002/anie.201809954
27. Rabe, P.; Rinkel, J.; Dolja, E.; Schmitz, T.; Nubbemeyer, B.; Luu, T. H.; Dickschat, J. S. *Angew. Chem., Int. Ed.* **2017**, *56*, 2776–2779. doi:10.1002/anie.201612439
28. Klapschinski, T. A.; Rabe, P.; Dickschat, J. S. *Angew. Chem., Int. Ed.* **2016**, *55*, 10141–10144. doi:10.1002/anie.201605425
29. Xu, H.; Goldfuss, B.; Dickschat, J. S. *Chem. – Eur. J.* **2021**, *27*, 9758–9762. doi:10.1002/chem.202101371
30. Singh, G.; Upadhyay, R. K. *Fitoterapia* **1994**, *65*, 186–187.
31. Ishihara, M.; Tsuneya, T.; Shiga, M.; Uneyama, K. *Phytochemistry* **1991**, *30*, 563–566. doi:10.1016/0031-9422(91)83727-3
32. Ishihara, M.; Masatsugu, Y.; Uneyama, K. *Tetrahedron* **1992**, *48*, 10265–10276. doi:10.1016/s0040-4020(01)88332-0

## License and Terms

This is an open access article licensed under the terms of the Beilstein-Institut Open Access License Agreement (<https://www.beilstein-journals.org/bjoc/terms>), which is identical to the Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0>). The reuse of material under this license requires that the author(s), source and license are credited. Third-party material in this article could be subject to other licenses (typically indicated in the credit line), and in this case, users are required to obtain permission from the license holder to reuse the material.

The definitive version of this article is the electronic one which can be found at: <https://doi.org/10.3762/bjoc.18.2>

## Appendix K

### **Stereochemical characterisation of the non-canonical $\alpha$ -humulene synthase from *Acremonium strictum***

*Org. Biomol. Chem.* **2021**, *19*, 8482

DOI: 10.1039/D1OB01769A

## COMMUNICATION



Cite this: *Org. Biomol. Chem.*, 2021, **19**, 8482

Received 9th September 2021,  
Accepted 13th September 2021

DOI: 10.1039/d1ob01769a

rsc.li/obc

## Stereochemical characterisation of the non-canonical $\alpha$ -humulene synthase from *Acremonium strictum*†

Houchao Xu,<sup>a</sup> Carsten Schotte,<sup>b</sup> Russell J. Cox <sup>b</sup> and Jeroen S. Dickschat <sup>\*a</sup>

The non-canonical fungal  $\alpha$ -humulene synthase was investigated through isotopic labelling experiments for its stereochemical course regarding inversion or retention at C-1, the face selectivity at C-11, and the stereoselectivity of the final deprotonation. A new and convenient desymmetrisation strategy was developed to enable a full stereochemical analysis of the catalysed steps to the achiral  $\alpha$ -humulene product from stereoselectively labelled farnesyl diphosphate.

$\alpha$ -Humulene (**1**) (Fig. 1) was first isolated from hops (*Humulus lupulus*) in 1895 by Chapman.<sup>1</sup> The structure of **1** was under debate for a long time and, after publication of the constitutional formula by Dev,<sup>2</sup> finally assigned correctly with all-*E* olefin configurations by Sutherland and Waters in 1961.<sup>3</sup> Its biosynthesis from farnesyl diphosphate (FPP) only requires a simple 1,11-cyclisation and deprotonation, and  $\alpha$ -humulene synthases catalysing this process have been reported from the plants *Zingiber zerumbet*<sup>4</sup> and *Aquilaria crassna*.<sup>5</sup> In *Z. zerumbet* **1** is subsequently oxidised by a cytochrome P450 monooxygenase and an alcohol dehydrogenase to zerumbone (**2**).<sup>6</sup> More recently, the non-canonical  $\alpha$ -humulene synthases AsR6 from *Acremonium strictum*<sup>7</sup> and PycR6 from an uncharacterised fungus<sup>8</sup> were reported, while the closely related enzyme EupR3 from a *Phaeosphaeriaceae* sp. produces (2*Z*)-humulene (**3**).<sup>8</sup> EupfG from *Penicillium janthinellum* likely converts FPP into the same product **3**, which was indirectly shown by characterisation of an oxidation product.<sup>9</sup> In fungi **1** and **3** serve as biosynthetic precursors towards tropolone sesquiterpenoids such as pycnidone (**4**)<sup>10,11</sup> and eupenifeldin (**5**)<sup>12</sup> that arise from a hetero-Diels Alder reaction with assistance of a Diels-Alderase.<sup>9</sup>

Because of its  $C_s$  symmetry, **1** is one of the few known achiral terpenes, which makes it particularly challenging to follow the stereochemical course of the terpene cyclase reaction with respect to (1) inversion or retention at C1, (2) attack at C-11 from the *Si* or the *Re* face, and (3) the abstraction of the 9-*pro-R* or 9-*pro-S* hydrogen in the final deprotonation step. Here we describe isotopic labelling experiments to address these problems for the fungal humulene synthase AsR6 and a desymmetrisation strategy for **1** to follow the results from these experiments.

The question of inversion *versus* retention of configuration at C-1 of FPP during the terpene cyclisation by AsR6 was addressed using the enantioselectively deuterated substrates (*R*)- and (*S*)-(1-<sup>13</sup>C,1-<sup>2</sup>H)FPP.<sup>13</sup> Their enzymatic conversion yields (1-<sup>13</sup>C,1-<sup>2</sup>H)-**1** with incorporation of deuterium into enantiotopic positions (Scheme 1A). While inversion of con-

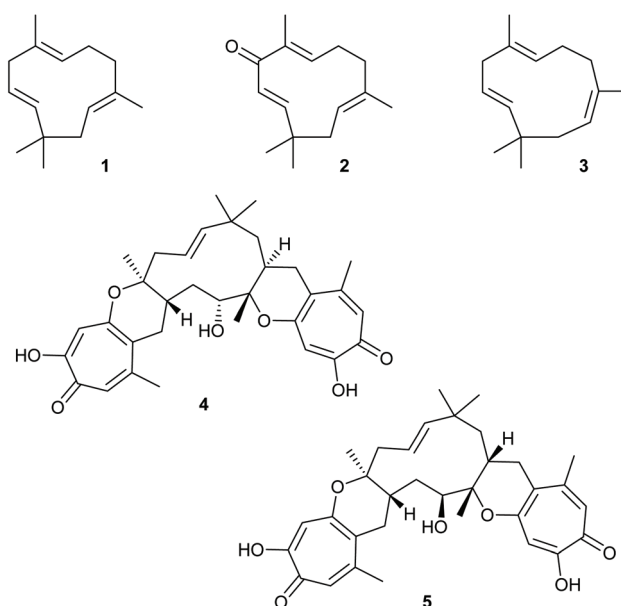
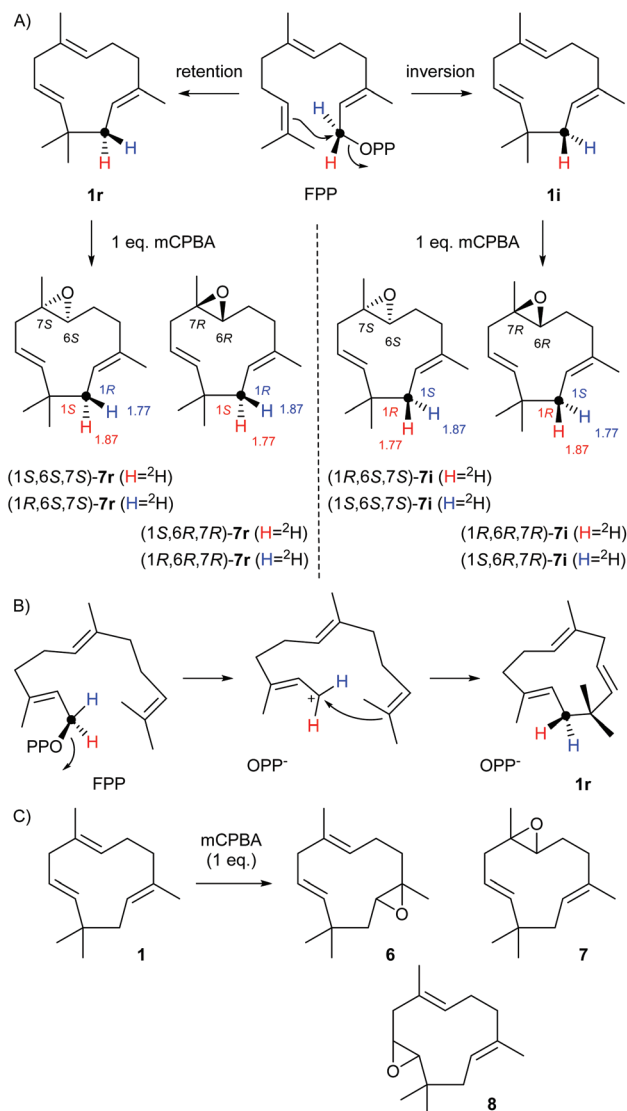


Fig. 1 Structures of compounds 1–5.

<sup>a</sup>Kekulé-Institute for Organic Chemistry and Biochemistry, University of Bonn, 53121 Bonn, Germany. E-mail: dickschat@uni-bonn.de

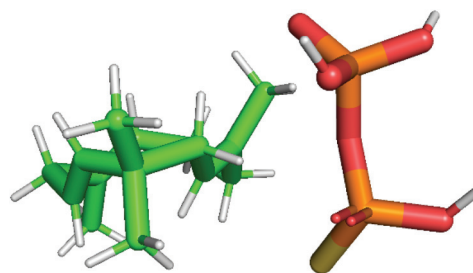
<sup>b</sup>Institute of Organic Chemistry, University of Hannover, Schneiderberg 38, 30167 Hannover, Germany

† Electronic supplementary information (ESI) available: Experimental details, tabulated NMR data and spectra of humulene epoxides and results from labelling experiments. See DOI: 10.1039/d1ob01769a



**Scheme 1** The stereochemical course of the cyclisation of FPP to **1** by AsR6 for C-1. (A) Enzymatic conversion of  $(R)$ - $(1-^{13}\text{C}, 1-^2\text{H})$ FPP (blue H =  $^2\text{H}$ ) and  $(S)$ - $(1-^{13}\text{C}, 1-^2\text{H})$ FPP (red H =  $^2\text{H}$ ) into **1** with retention or inversion of configuration at C-1 and subsequent desymmetrisation by conversion into humulene epoxide II (**7**) with mCPBA. Black dots indicate  $^{13}\text{C}$ -labelled carbons. The stereochemical descriptors for C-1 shown in blue are for the labelled compounds from  $(R)$ - $(1-^{13}\text{C}, 1-^2\text{H})$ FPP and those in red are for the labelled compounds from  $(S)$ - $(1-^{13}\text{C}, 1-^2\text{H})$ FPP. Data at the labelled hydrogens indicate  $^1\text{H-NMR}$  shifts in ppm. (B) Mechanism explaining retention of configuration. (C) Epoxidation of **1** with mCPBA yielding racemic humulene epoxides **6–8**.

figuration can be achieved by a simple  $\text{S}_{\text{N}}2$  reaction at C-1 leading to **1i** (**i** for inversion), retention of configuration to **1r** (**r** for retention) is more difficult to explain. After the abstraction of diphosphate a conformational change of the resulting farnesyl cation could bring C-1 into close proximity of C-11, allowing its attack from the same side as the diphosphate has left (Scheme 1B). Such a reaction would be rather unusual, but the crystal structure of AsR6 with *in crystallo* generated **1** and thiolodiphosphate<sup>14</sup> shows an almost orthogonal orientation



**Fig. 2** Relative orientation of thiolodiphosphate and *in crystallo* formed **1** in the crystal structure of AsR6 (PDB code 7OC4).

of incoming (C-11) and leaving group (sulfur atom of thiolodiphosphate, Fig. 2), which does not point clearly to inversion or retention at C-1.

The AsR6 cyclisations of  $(R)$ - and  $(S)$ - $(1-^{13}\text{C}, 1-^2\text{H})$ FPP through inversion or retention lead to enantiomers of  $(1-^{13}\text{C}, 1-^2\text{H})$ -**1** which cannot be distinguished by NMR spectroscopy. Therefore, a desymmetrisation strategy was developed to distinguish between the 1-*pro-R* and 1-*pro-S* hydrogens in **1**. The epoxidation of **1** with one equivalent of *meta*-chloroperbenzoic acid (mCPBA) is known to yield a racemic mixture of all three epoxides, with humulene epoxide II (**7**) as the main product and humulene epoxides I (**6**) and III (**8**) as side products, and only little formation of bis- and tris-epoxides (Scheme 1C).<sup>15</sup> The epoxidation of enantioselectively deuterated  $(1-^{13}\text{C}, 1-^2\text{H})$ -**1** will for each of these epoxides produce two diastereomers from **1r** that are enantiomeric to the expected products from **1i**, as shown for the main product **7** (Scheme 1A).

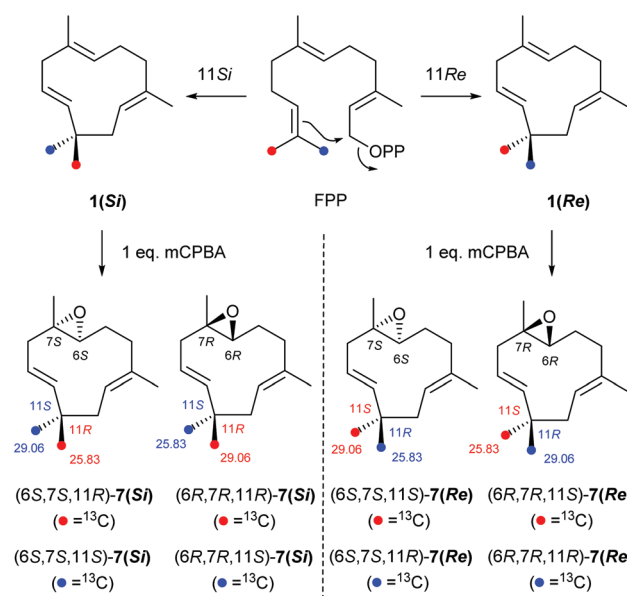
The obtained diastereoisomers, either  $(6S, 7S)$ -**7r** and  $(6R, 7R)$ -**7r** by retention or  $(6S, 7S)$ -**7i** and  $(6R, 7R)$ -**7i** by inversion at C-1, will behave in many aspects like enantiomers, as the configuration at the deuterated carbon C-1 will have no strong influence on the polarity in chromatographic separations or on the optical rotations, but NMR spectroscopy will allow discrimination between the incorporations of deuterium into the diastereotopic positions of the epoxides, showing the relative orientations of hydrogen and deuterium at C-1 with respect to the epoxide. The absolute configuration for each product can be determined from the optical rotations, as for natural  $(-)$ -**6** and  $(-)$ -**7** from *Z. zerumbet* the structures of  $(2R, 3R)$ -**6** and  $(6R, 7R)$ -**7** were assigned.<sup>16,17</sup> For natural  $(+)$ -**8** only the relative configuration has been determined.<sup>18</sup>

In a first step, all three humulene epoxides **6–8** were prepared by oxidation with mCPBA for a complete assignment of NMR data including assignments of the diastereotopic hydrogens and methyl groups (Tables S1–S3 and Fig. S1–S24†). For the main product **7** the separation of enantiomers was achieved by HPLC on a chiral stationary phase, yielding enantiomerically pure  $(6R, 7R)$ - $(-)$ -**7** and  $(6S, 7S)$ - $(+)$ -**7**, which allowed further experimental work to focus on this main epoxidation product. Then  $(R)$ - and  $(S)$ - $(1-^{13}\text{C}, 1-^2\text{H})$ FPP<sup>13</sup> (5 mg each) were enzymatically converted with AsR6. The enzyme products were extracted and mixed with unlabelled **1** (30 mg), followed by

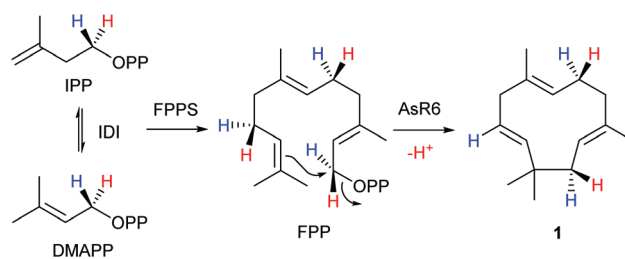


epoxidation with mCPBA and isolation of labelled  $[1-^{13}\text{C}, 1-^2\text{H}]$ -(6*R*,7*R*)-(-)-**7** and  $[1-^{13}\text{C}, 1-^2\text{H}]$ -(6*S*,7*S*)-(+)-**7**. The additional  $^{13}\text{C}$ -labelling in these compounds allows for a sensitive analysis by HSQC spectroscopy, revealing for each sample specific incorporation of deuterium into one of the diastereotopic hydrogen positions at C-1, *i.e.* (1*S*)-configuration for both stereoisomers of labelled **7** obtained from (*R*)-( $1-^{13}\text{C}, 1-^2\text{H}$ )FPP (incorporation of deuterium for (6*S*,7*S*)-**7** into  $\delta = 1.87$  ppm and for (6*R*,7*R*)-**7** into  $\delta = 1.77$  ppm) and (1*R*)-configuration for the two stereoisomers of labelled **7** from (*S*)-( $1-^{13}\text{C}, 1-^2\text{H}$ )FPP (incorporation of deuterium for (6*S*,7*S*)-**7** into  $\delta = 1.77$  ppm and for (6*R*,7*R*)-**7** into  $\delta = 1.87$  ppm, Fig. S25†). These results demonstrate inversion of configuration at C-1 during the cyclisation of FPP to **1** by AsR6.

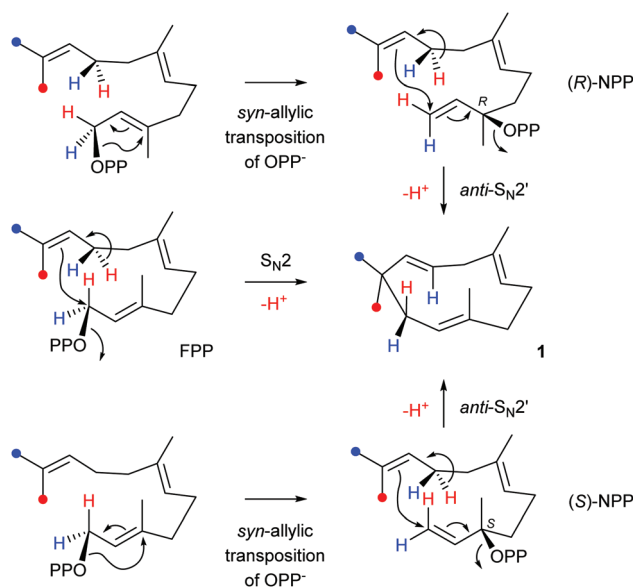
The face selectivity at C-11 during FPP cyclisation regarding *Re* or *Si* face attack was investigated by an analogous approach using (12- $^{13}\text{C}$ )FPP and (13- $^{13}\text{C}$ )FPP (Scheme 2). These compounds will be converted through 11*Si* cyclisation into **1**(*Si*) or 11*Re* cyclisation into **1**(*Re*), and subsequent epoxidation will yield the shown stereoisomers of **7**. Incubation of both labelled isotopomers (12- $^{13}\text{C}$ )FPP and (13- $^{13}\text{C}$ )FPP<sup>19</sup> with AsR6, mixing with unlabelled **1** and epoxidation with mCPBA was followed by separation of the diastereoisomers of labelled  $[12-^{13}\text{C}]$ -(6*R*,7*R*)-(-)-**7** and  $[13-^{13}\text{C}]$ -(6*S*,7*S*)-(+)-**7**. Their analysis by  $^{13}\text{C}$ -NMR spectroscopy revealed (11*R*)-configuration for both stereoisomers of labelled **7** obtained from (12- $^{13}\text{C}$ )FPP (incorporation of labelling for (6*S*,7*S*)-**7** into  $\delta = 25.83$  ppm and for (6*R*,7*R*)-**7** into  $\delta = 29.06$  ppm), and (11*S*)-configuration for the products from (13- $^{13}\text{C}$ )FPP (incorporation of labelling for (6*S*,7*S*)-**7** into  $\delta = 29.06$  ppm and for (6*R*,7*R*)-**7** into  $\delta =$



**Scheme 2** The stereochemical course of the cyclisation of FPP to **1** by AsR6 for C-11. Enzymatic conversion of (12- $^{13}\text{C}$ )FPP (blue dots =  $^{13}\text{C}$ ) and (13- $^{13}\text{C}$ )FPP (red dots =  $^{13}\text{C}$ ) into **1** with 11*Si* or 11*Re* attack. The stereochemical descriptors for C-11 shown in blue are for the labelled compounds from (12- $^{13}\text{C}$ )FPP and those in red are for the labelled compounds from (13- $^{13}\text{C}$ )FPP. Data at the labelled carbons indicate  $^{13}\text{C}$ -NMR shifts in ppm.



**Scheme 3** The stereochemical course of the cyclisation of FPP to **1** by AsR6 for the deprotonation from C-9. Enzymatic conversion of (*R*)-( $1-^2\text{H}$ )IPP (blue H =  $^2\text{H}$ ) and (*S*)-( $1-^2\text{H}$ )IPP (red H =  $^2\text{H}$ ) with IDI, FPPS and AsR6.



**Scheme 4** The stereochemical course of the cyclisation of FPP to **1** by AsR6. The results from labelling experiments are in accordance with direct  $\text{S}_{\text{N}}2$  cyclisation, but also with reactions through (*R*)- or (*S*)-NPP.

25.83 ppm, Fig. S26†), in agreement with *Re* face attack at C-11 in the FPP cyclisation by AsR6.

The stereochemical course for the final deprotonation step was followed using the substrates (*R*)- and (*S*)-( $1-^2\text{H}$ )isopentenyl diphosphate<sup>20</sup> (IPP) that were enzymatically equilibrated with (*R*)- and (*S*)-( $1-^2\text{H}$ )dimethylallyl diphosphate (DMAPP) by isopentenyl diphosphate isomerase (IDI) from *Escherichia coli*,<sup>21,22</sup> followed by enzymatic conversion with FPP synthase (FPPS) from *Streptomyces coelicolor*<sup>23</sup> and AsR6 (Scheme 3). GC/MS analysis of the products revealed specific loss of deuterium from (*S*)-( $1-^2\text{H}$ )IPP, *i.e.* loss of the 9-*pro-S* hydrogen of FPP (Fig. S27†).

## Conclusions

For the achiral sesquiterpene hydrocarbon  $\alpha$ -humulene (**1**) several stereochemical problems in its biosynthesis have been addressed through isotopic labelling experiments, using a new desymmetri-

sation strategy involving facile epoxidation and chromatographic separation of the stereoisomers. In contrast to the experiments described here, a chemical desymmetrisation was not necessary for the previously investigated stereochemical course of the biosynthesis of the achiral monoterpene ether 1,8-cineol, because in this case the introduction of stereoselective labelling does not yield enantiomers, but produces diastereoisomers, which allows for a direct analysis by NMR spectroscopy.<sup>24,25</sup>

The results show for the non-canonical fungal  $\alpha$ -humulene synthase AsR6 (1) inversion of configuration at C-1, (2) *Re* face attack at C-11, and (3) specific loss of the 9-*pro-R* hydrogen of FPP in the final deprotonation. These findings are in accordance with a direct cyclisation of FPP through an S<sub>N</sub>2 reaction, but according to Arigoni's model for cadalanes<sup>26</sup> also isomerisation of FPP to either enantiomer of nerolidyl diphosphate (NPP) by *syn*-allylic transposition of diphosphate, followed by a ring closure through *anti*-S<sub>N</sub>2' reaction is possible (Scheme 4). Here, different starting conformers of FPP need to be assumed, but for both enantiomers of NPP the results from the labelling experiments can be explained correctly.

Inversion of configuration at C-1 has been shown previously for pentalenene synthase from *Streptomyces exfoliatus*<sup>27</sup> and for presilphiperfolan-8 $\beta$ -ol synthase from *Botrytis cinerea*<sup>28</sup> that also both catalyse an initial 1,11-cyclisation of FPP. Moreover, the facial selectivity at C-11 for pentalenene synthase is the same as for AsR6, but has been described as "attack on the Si face of the 10,11 double bond"<sup>29</sup> (our assignment of *Re* and *Si* face refers to the attacked carbon C-11 and makes use of the sequence rule *Z* > *E* in the Cahn–Ingold–Prelog nomenclature, giving preference of the *Z*-Me group over the *E*-Me group bound to C-11).<sup>30</sup> The plant  $\alpha$ -humulene synthases have not been investigated for their stereochemical course, but it will be interesting to look into this, because plant terpene synthases often produce the enantiomers of bacterial and fungal terpene synthases.<sup>31–33</sup> For achiral **1** this could be reflected by an enantiomeric conformational fold of FPP in the active site that should lead to a different face selectivity at C-11. Finally, AsR6 removes the substrate 9-*pro-R* hydrogen to form  $\alpha$ -humulene as the final product of the reaction sequence. Pentalenene synthase also removes the 9-*pro-R* hydrogen to form an  $\alpha$ -humulene intermediate.<sup>27</sup> However, in this case, reprotonation of the newly formed *E*-9,10 olefin begins a cascade of reactions towards pentalenene itself that AsR6 does not catalyse. The structural basis for these mechanistic and stereochemical similarities and differences will form the basis for future studies.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

This work was funded by the DFG (DI1536/7-2). We thank Andreas Schneider (Bonn) for skillful technical assistance.

## Notes and references

- 1 A. C. Chapman, *J. Chem. Soc., Trans.*, 1895, **67**, 54.
- 2 S. Dev, *Tetrahedron Lett.*, 1959, **1**, 12.
- 3 M. D. Sutherland and O. J. Waters, *Aust. J. Chem.*, 1961, **14**, 596.
- 4 F. Yu, S. Okamoto, K. Nakasone, K. Adachi, S. Matsuda, H. Harada, N. Misawa and R. Utsumi, *Planta*, 2008, **227**, 1291.
- 5 Y. Kumeta and M. Ito, *J. Nat. Med.*, 2016, **70**, 452.
- 6 F. Yu, S. Okamoto, H. Harada, K. Yamasaki, N. Misawa and R. Utsumi, *Cell. Mol. Life Sci.*, 2011, **68**, 1033.
- 7 S. Schor, C. Schotte, D. Wibberg, J. Kalinowski and R. J. Cox, *Nat. Commun.*, 2018, **9**, 1963.
- 8 C. Schotte, L. Li, D. Wibberg, J. Kalinowski and R. J. Cox, *Angew. Chem.*, 2020, **59**, 23870.
- 9 Q. Chen, J. Gao, C. Jamieson, J. Liu, M. Ohashi, J. Bai, D. Yan, B. Liu, Y. Che, Y. Wang, K. N. Houk and Y. Hu, *J. Am. Chem. Soc.*, 2019, **141**, 14052.
- 10 G. H. Harris, K. Hoogsteen, K. C. Silverman, S. L. Raghoobar, G. F. Bills, R. B. Lingham, J. L. Smith, H. W. Dougherty, C. Cascales and F. Pelaez, *Tetrahedron*, 1993, **49**, 2139.
- 11 C. Y. Bemis, C. N. Ungarean, A. S. Shved, C. S. Jamieson, T. Hwang, K. S. Lee, K. N. Houk and D. Sarlah, *J. Am. Chem. Soc.*, 2021, **143**, 6006.
- 12 F. Mayerl, Q. Gao, S. Huang, S. E. Klohr, J. A. Matson, D. R. Gustavson, D. M. Pirnik, R. L. Berry, C. Fairchild and W. C. Rose, *J. Antibiot.*, 1993, **46**, 1082.
- 13 P. Rabe, J. Rinkel, E. Dolja, T. Schmitz, B. Nubbemeyer, T. H. Luu and J. S. Dickschat, *Angew. Chem., Int. Ed.*, 2017, **56**, 2776.
- 14 C. Schotte, P. Lukat, A. Deuschmann, W. Blankenfeldt and R. J. Cox, *Angew. Chem.*, 2021, **60**, 20308.
- 15 N. Zigon, M. Hoshino, S. Yoshioka, Y. Inokuma and M. Fujita, *Angew. Chem., Int. Ed.*, 2015, **54**, 9033.
- 16 N. P. Damodaran and S. Dev, *Tetrahedron*, 1968, **24**, 4123.
- 17 N. P. Damodaran and S. Dev, *Tetrahedron*, 1968, **24**, 4133.
- 18 A. S. Gupta and S. Dev, *Tetrahedron*, 1971, **27**, 635.
- 19 P. Rabe, L. Barra, J. Rinkel, R. Riclea, C. A. Citron, T. A. Klapschinski, A. Janusko and J. S. Dickschat, *Angew. Chem., Int. Ed.*, 2015, **54**, 13448.
- 20 J. Rinkel, L. Lauterbach and J. S. Dickschat, *Angew. Chem., Int. Ed.*, 2019, **58**, 452.
- 21 F. M. Hahn, A. P. Hurlburt and C. D. Poulter, *J. Bacteriol.*, 1999, **181**, 4499.
- 22 J. Rinkel and J. S. Dickschat, *Org. Lett.*, 2019, **21**, 2426.
- 23 P. Rabe, J. Rinkel, B. Nubbemeyer, T. G. Köllner, F. Chen and J. S. Dickschat, *Angew. Chem., Int. Ed.*, 2016, **55**, 15420.
- 24 M. L. Wise, M. Urbansky, G. L. Helms, R. M. Coates and R. Croteau, *J. Am. Chem. Soc.*, 2002, **124**, 8546.
- 25 J. Rinkel, P. Rabe, L. zur Horst and J. S. Dickschat, *Beilstein J. Org. Chem.*, 2016, **12**, 2317.
- 26 D. Arigoni, *Pure Appl. Chem.*, 1975, **41**, 219.
- 27 D. E. Cane, J. S. Oliver, P. H. M. Harrison, C. Abell, B. R. Hubbard, C. T. Kane and R. Lattman, *J. Am. Chem. Soc.*, 1990, **112**, 4513.

- 28 C.-M. Wang, R. Hopson, X. Lin and D. E. Cane, *J. Am. Chem. Soc.*, 2009, **131**, 8360.
- 29 D. E. Cane, T. Rossi, A. M. Tillman and J. P. Pachlatko, *J. Am. Chem. Soc.*, 1981, **103**, 1838.
- 30 R. S. Cahn, C. Ingold and V. Prelog, *Angew. Chem., Int. Ed. Engl.*, 1966, **5**, 385.
- 31 L. Ding, H. Goerls, K. Dornblut, W. Lin, A. Maier, H.-H. Fiebig and C. Hertweck, *J. Nat. Prod.*, 2015, **78**, 2963.
- 32 P. Rabe, T. Schmitz and J. S. Dickschat, *Beilstein J. Org. Chem.*, 2016, **12**, 1839.
- 33 H. Xu, J. Rinkel and J. S. Dickschat, *Org. Chem. Front.*, 2021, **8**, 1177.

## Appendix L

### **Biosynthesis of the Sesquiterpene Kitaviridene through Skeletal Rearrangement with Formation of a Methyl Group Equivalent**

*Org. Lett.* **2023**, *25*, 3330

DOI: [10.1021/acs.orglett.3c01211](https://doi.org/10.1021/acs.orglett.3c01211)

# Biosynthesis of the Sesquiterpene Kitaviridene through Skeletal Rearrangement with Formation of a Methyl Group Equivalent

Houchao Xu, Bernd Goldfuss, and Jeroen S. Dickschat\*



Cite This: *Org. Lett.* 2023, 25, 3330–3334



Read Online

ACCESS |



Metrics & More

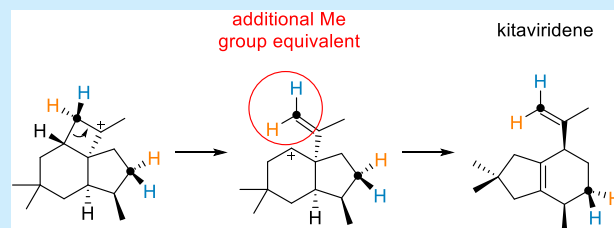


Article Recommendations

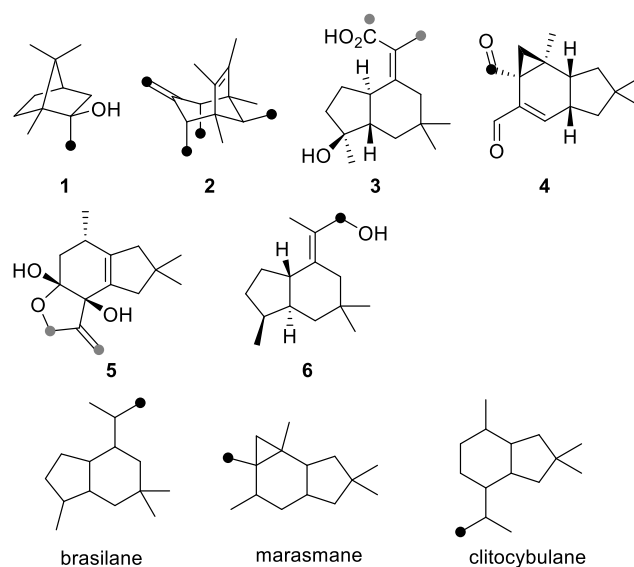


Supporting Information

**ABSTRACT:** A sesquiterpene synthase from *Kitasatospora viridis* was discovered and shown to produce kitaviridene, a sesquiterpene hydrocarbon with an additional methyl group equivalent in comparison to a regular sesquiterpene. Isotopic labeling experiments together with DFT calculations gave detailed insights into the cyclization cascade toward kitaviridene and explained the formation of the additional methyl group equivalent.



The biosynthesis of terpenes is a fascinating process that starts from only two  $C_5$  building blocks, the electrophile dimethylallyl diphosphate (DMAPP) and the nucleophile isopentenyl diphosphate (IPP). They can be fused by a prenyltransferase to the monoterpene precursor geranyl diphosphate (GPP),<sup>1</sup> and if further IPP units are attached, farnesyl diphosphate (FPP),<sup>2</sup> geranylgeranyl diphosphate (GGPP),<sup>3</sup> geranylarnesyl diphosphate (GFPP),<sup>4</sup> and farnesylfarnesyl diphosphate (FFPP)<sup>5</sup> can successively arise. These acyclic and achiral precursors composed of  $n$  units are then converted by type I terpene synthases into terpene hydrocarbons ( $C_{5n}H_{8n}$ ) or alcohols ( $C_{5n}H_{8n+2}O$ ) that are often polycyclic and may contain several stereogenic centers.<sup>6</sup> This metamorphosis requires just a single, yet highly complex enzymatic transformation that proceeds through a cationic cascade reaction that is initiated by the abstraction of diphosphate. The resulting allyl cation can then undergo cyclization reactions, Wagner–Meerwein rearrangements, and hydride or proton shifts. The cascade is terminated by deprotonation or quenching with water. The number of methyl group equivalents (MGEs, this includes usually methyl groups plus olefinic methylene groups that arise from methyl groups in deprotonation steps) is usually  $MGE = 3$  for a monoterpene and increases by one for each additional IPP unit; i.e., the general formula is  $MGE = n + 1$  for a terpene built from  $n$  isoprene units. Some cases are known that deviate from this rule (Figure 1), showing additional methyl group equivalents; e.g., 2-methylisoborneol (1)<sup>7</sup> contains four methyl groups, but this case is well understood, as the fourth methyl group is introduced from *S*-adenosylmethionine (SAM) by a methyltransferase.<sup>8</sup> Similarly, sodorifen (2)<sup>9</sup> is a methylated sesquiterpene with one SAM-derived methyl group,<sup>10</sup> but for this compound eight MGEs are observed, requiring the formation of three additional MGEs from chain methylene carbons of FPP during terpene cyclization. Further sesquiterpenes containing additional MGEs include the antibacterial

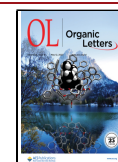


**Figure 1.** Structures of terpenoids with additional MGEs (marked with black dots; gray dots indicate alternative positions).

brasilane sesquiterpene xylarenic acid (3) from *Xylaria* sp.,<sup>11</sup> the antifungal and antibacterial marasmane derivative isovelleral (4) from the fungus *Lactarius vellereus*,<sup>12</sup> and clitocybulol A (5) from the fungus *Clitocybula oculus*.<sup>13</sup> A sesquiterpene synthase for trichobrasilenol (6), a representative of brasilane sesquiterpenes, has recently been reported

**Received:** April 13, 2023

**Published:** May 1, 2023



from *Trichoderma*, and its mechanism was elucidated through isotopic labeling experiments<sup>14</sup> and by computational chemistry,<sup>15</sup> showing that also in this case an additional MGE arises from a chain methylene group of FPP. To date, no enzymes are known that make a marasmane or a clitocybulane sesquiterpene. Here we report on the identification and mechanistic characterization of the kitaviridene synthase from *Kitasatospora viridis*, an enzyme that produces a sesquiterpene of the clitocybulane skeleton, and demonstrate the mechanism for the formation of its additional MGE.

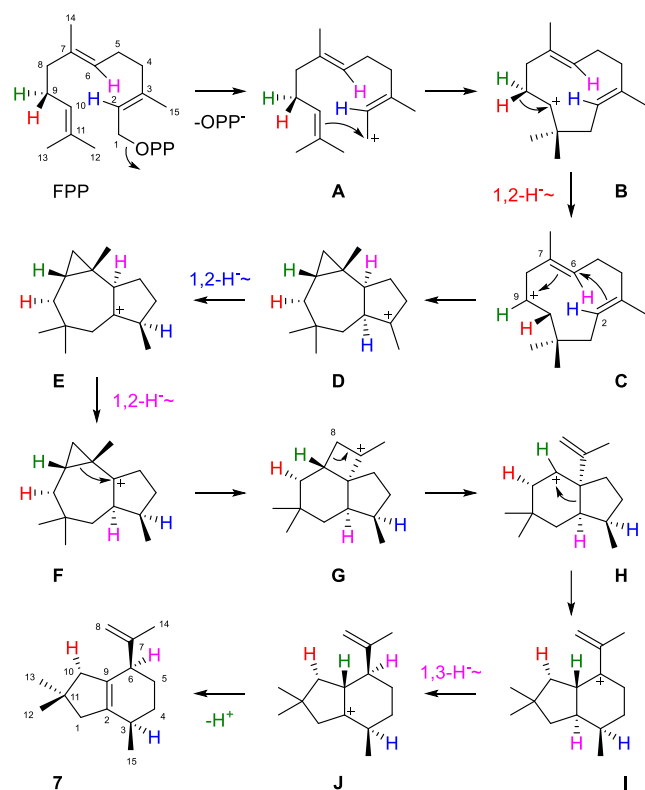
A gene from *Kitasatospora viridis* coding for a terpene synthase that does not cluster with other characterized terpene synthases in a phylogenetic tree (Figure S1) was cloned and expressed in *Escherichia coli*. The predicted amino acid sequence (Figure S2) showed the presence of all highly conserved motifs required for functionality,<sup>16–18</sup> including the aspartate-rich motif (<sup>81</sup>DDQFE), the pyrophosphate sensor (R175), the NSE triad (<sup>221</sup>NEVFSLEKE), and the RY pair (<sup>314</sup>RY), suggesting the enzyme could be functional. The purified recombinant protein (Figure S3) was incubated with GPP resulting in the production of several acyclic and monocyclic monoterpenes including linalool, geraniol, myrcene, (*E*)- and (*Z*)- $\beta$ -ocimene,  $\alpha$ -terpineol, limonene, and  $\alpha$ - and  $\beta$ -phellandrene (Figures S4 and S5, Table S2). FPP was converted into an unknown sesquiterpene hydrocarbon as the main product, besides small amounts of (*E*)- $\beta$ -farnesene,  $\alpha$ -humulene, (*E,E*)- $\alpha$ -farnesene, (*E,E*)-farnesol, and several unidentified trace compounds (Figures S4 and S5, Table S2), while GGPP and GFPP were not accepted as substrates. The sesquiterpene hydrocarbon was isolated, and its structure was elucidated by NMR spectroscopy (Table S3, Figures S6–S13), revealing the structure of a bicyclic compound with five MGEs that was named kitaviridene (7, Scheme 1). The

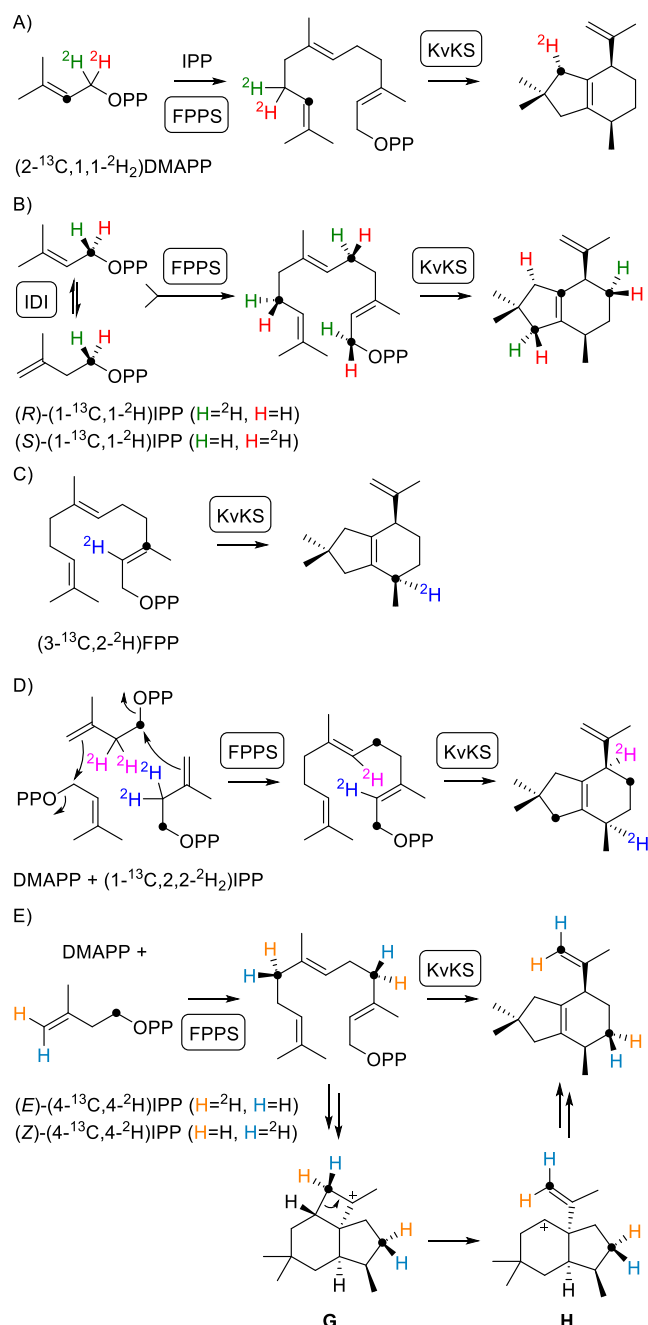
formation of 7 may be considered as the main function of the terpene synthase, while acyclic and simple monocyclic side products likely arise from GPP and FPP without tight enzyme control. Thus, the sesquiterpene synthase was identified as *Kitasatosporaviridis* Kitaviridene Synthase (KvKS), an enzyme for which the main product from FPP is a rare case of a sesquiterpene with an additional MGE.

The proposed cyclization mechanism to 7 starts with ionization of FPP to the farnesyl cation (A), followed by 1,11-cyclization to the (*E,E*)-humulyl cation (B) (Scheme 1). A subsequent 1,2-hydride shift to C, 7,9-, and 2,6-cyclization results in D that can further react by two sequential 1,2-hydride shifts to E and F. A ring expansion of the cyclopropane ring to G and opening of the cyclobutane ring may lead to H with the formation of a fifth MGE as an olefinic methylene group from C-8. Intermediate H may then undergo a Wagner–Meerwein rearrangement to I, a 1,3-hydride shift to J, and terminal deprotonation to yield 7.

To investigate the cyclization mechanism of KvKS experimentally, a series of isotopic labeling experiments were carried out (Table S4). The incubation of all 15 isotopomers of (<sup>13</sup>C)FPP with KvKS resulted in each of the 15 experiments in the incorporation of the <sup>13</sup>C label in the expected positions (Figure S14), which supported the mechanism of Scheme 1 and especially the skeletal rearrangements along the cyclization cascade. The 1,2-hydride shift from B to C was investigated using (2-<sup>13</sup>C,1,1-<sup>2</sup>H<sub>2</sub>)DMAPP and IPP in conjunction with FPP synthase (FPPS) from *Streptomyces coelicolor* and KvKS (Scheme 2A). In this reaction (10-<sup>13</sup>C,9,9-<sup>2</sup>H<sub>2</sub>)FPP is formed, and during the cyclization to 7 one deuterium is shifted from C-9 to C-10, as is evident from an upfield shifted triplet for C-10 in the <sup>13</sup>C NMR spectrum ( $\Delta\delta = -0.42$  ppm, <sup>1</sup>J<sub>C,D</sub> = 19.5 Hz, Figure S15). Its target position in 7 was identified by HSQC analysis, together with the NOESY-based assignment of hydrogens (Figure S6), as *trans* to the isopropenyl group (Figure S15D and E). The second deuterium is lost, confirming the terminal deprotonation from J to 7 (Figure S16A and B). To investigate which of the two hydrogens at C-9 of FPP is lost in this deprotonation step, (*R*)- and (*S*)-(1-<sup>13</sup>C,1-<sup>2</sup>H)IPP were converted with isopentenyl diphosphate isomerase (IDI) from *E. coli*, FPPS, and KvKS (Scheme 2B). This reaction yields FPP stereoselectively deuterated and <sup>13</sup>C-labeled at C-1, C-5, and C-9, and the GC/MS analysis of labeled 7 obtained in these experiments revealed the specific loss of deuterium from (*R*)-(1-<sup>13</sup>C,1-<sup>2</sup>H)IPP (Figure S16C, green hydrogen). The deuterium from (*S*)-(1-<sup>13</sup>C,1-<sup>2</sup>H)IPP is retained and is consequently identified as the one that shifts in the step from B to C (Figure S16D, red hydrogen). The 1,2-hydride shift from D to E (blue hydrogen) was investigated using (3-<sup>13</sup>C,2-<sup>2</sup>H)FPP (Scheme 2C), resulting in an upfield shifted triplet for C-3 of 7 ( $\Delta\delta = -0.51$  ppm, <sup>1</sup>J<sub>C,D</sub> = 19.2 Hz, Figure S17). Finally, the fate of the hydrogen connected to C-6 of FPP (purple hydrogen) was investigated. According to the biosynthetic model, H-6 first undergoes a 1,2-hydride shift from C-6 to C-2 in the step from E to F. This is later followed by a migration back from C-2 to C-6 in the step from I to J. A labeling experiment with DMAPP, (1-<sup>13</sup>C,2,2-<sup>2</sup>H<sub>2</sub>)IPP, FPPS, and KvKS generates (1,5-<sup>13</sup>C<sub>2</sub>,2,6-<sup>2</sup>H<sub>2</sub>)FPP that is subsequently cyclized to (1,5-<sup>13</sup>C<sub>2</sub>,3,6-<sup>2</sup>H<sub>2</sub>)-7 (Scheme 2D). The ultimate attachment of a deuterium at C-6 of 7 is revealed by a small upfield shift for C-5 that indicates a deuterium atom in a neighboring position (Figure S18).

Scheme 1. Cyclization Mechanism from FPP to 7



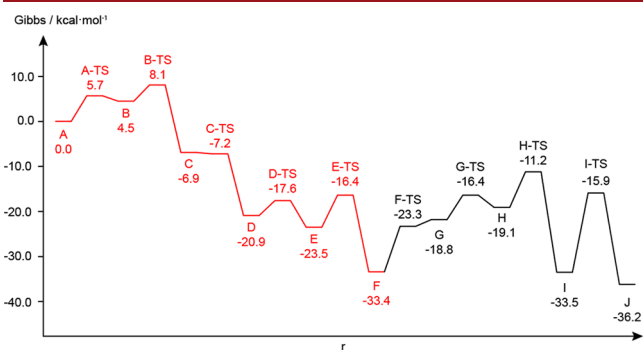
Scheme 2. Labeling Experiments for the Biosynthesis of 7<sup>a</sup>

<sup>a</sup>Black dots indicate <sup>13</sup>C-labeled carbons.

The absolute configuration of 7 was determined by chemical correlation using stereoselectively deuterated terpene precursors containing stereogenic anchors of known absolute configuration. After their conversion with KvKS, the relative orientation of the naturally present stereogenic centers with respect to the artificially introduced anchors can be determined, allowing to conclude on the absolute configuration of 7. The experiment shown in Scheme 2B using (*R*)- and (*S*)-(1-<sup>13</sup>C,1-<sup>2</sup>H)IPP not only can give insights into the fate of the C-9 hydrogens but also sets two stereogenic anchors at C-1 and C-5 for the absolute configuration determination. The additional <sup>13</sup>C-label in these probes allows for an efficient product analysis by HSQC spectroscopy (Figure S19). Two additional experiments were performed using DMAPP with

(*E*)- or (*Z*)-(4-<sup>13</sup>C,4-<sup>2</sup>H)IPP that were enzymatically converted into labeled 7 using FPPS and KvKS, resulting in a third anchor at C-4 (Scheme 2E, Figure S20). All four experiments established the absolute configuration of (3*R*,6*R*)-7, which is opposite to the absolute configuration of a hypothetical precursor toward 5. In addition, the experiments using (*E*)- or (*Z*)-(4-<sup>13</sup>C,4-<sup>2</sup>H)IPP revealed the stereochemical course for the fragmentation reaction from G to H: The 4*E* hydrogen of IPP (orange) ends up in the 8*Z* position of 7, and the 4*Z* hydrogen of IPP (light blue) converts into the 8*E* hydrogen of 7 (Scheme 2E and Figure S21). The interpretation of these experiments with the stereoselectively deuterated terpene precursors relies on the known stereochemical course for the formation of FPP from DMAPP and IPP with inversion of configuration at C-1 of the elongated allyl diphosphate and *Si* face attack at C-4 of IPP.<sup>19,20</sup>

Quantum chemical calculations were performed to further investigate the feasibility of the reactions toward 7 as shown in Scheme 1. The first steps from FPP to intermediate F were computed before by Uchiyama and co-workers for the biosynthesis of 6 (marked in red in Figure 2).<sup>15</sup> The



**Figure 2.** Computed energy profile for the cyclization mechanism from FPP to 7 as shown in Scheme 1 (mPW1PW91/6-31G\*\*//B97D3/6-31G\*\*, cf. SI). Steps computed by Uchiyama and co-workers are shown in red,<sup>15</sup> and steps computed in this work are shown in black.

subsequent steps from F to 7 could all be realized computationally, albeit with fairly high transition state barriers; e.g. the highest barrier was found with +16.6 kcal/mol for the 1,3-hydride shift from the *trans* fused bicyclic intermediate I to J. However, 1,2-hydride migrations that may have lower transition state barriers cannot explain the formation of 7 from I. Previous DFT calculations have shown that the transition state barriers for similar 1,3-hydride shifts for *trans*-fused intermediates in the biosynthesis of guaiane sesquiterpenes are similarly high, and also in such cases these barriers can be overcome as verified experimentally for the guai-4(15)-en-11-ol synthase from poplar.<sup>21</sup> In contrast and as a result of steric constraints, for *cis*-fused systems the transition state barriers for 1,3-hydride shifts were generally found to be too high,<sup>21</sup> and therefore these systems preferentially react through a sequence of two 1,2-hydride shifts whose transition state barriers are often much lower than for the 1,3-hydride migrations at *trans*-annulated rings.<sup>21,22</sup> However, an alternative cyclization mechanism from FPP to 7 through a series of stereoisomeric intermediates avoiding a 1,3-hydride shift was not in line with the results from the labeling experiments and also showed a surprisingly high barrier for a terminal 1,2-hydride migration of a *cis*-fused bicyclic intermediate (Figure S22). Notably, the

calculations for the pathway of Scheme 1 also correctly predicted the observed stereochemical course for the ring opening from G to H based on isotopic labeling experiments (Scheme 2E). The overall process from A to J was highly exergonic with a free energy release of  $-36.2$  kcal/mol.

In summary, we have identified the sesquiterpene synthase KvKS from *Kitasatospora viridis* for kitaviridene, a compound with an additional methyl group equivalent as compared to regular sesquiterpenes. Isotopic labeling experiments in conjunction with DFT calculations revealed the intricate reaction mechanism for its formation. The absolute configuration of kitaviridene is opposite to that of the hypothetical parent hydrocarbon of the fungal clitocybols, suggesting that a terpene synthase for the enantiomer of kitaviridene or a structurally similar compound remains to be discovered.

## ■ ASSOCIATED CONTENT

### Data Availability Statement

The data underlying this study are available in the published article and its Supporting Information.

### SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.orglett.3c01211>.

Experimental and computational details, analytical data of kitaviridene (7), and results from labeling experiments (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Author

Jeroen S. Dickschat – *Kekulé-Institute for Organic Chemistry and Biochemistry, University of Bonn, 53121 Bonn, Germany*; [orcid.org/0000-0002-0102-0631](https://orcid.org/0000-0002-0102-0631);  
Email: [dickschat@uni-bonn.de](mailto:dickschat@uni-bonn.de)

### Authors

Houchao Xu – *Kekulé-Institute for Organic Chemistry and Biochemistry, University of Bonn, 53121 Bonn, Germany*  
Bernd Goldfuss – *Department for Chemistry, University of Cologne, 50939 Cologne, Germany*; [orcid.org/0000-0002-1814-8818](https://orcid.org/0000-0002-1814-8818)

Complete contact information is available at:  
<https://pubs.acs.org/doi/10.1021/acs.orglett.3c01211>

### Author Contributions

H.X. performed all experimental work and analyzed the data; B.G. performed DFT calculations; and J.S.D. analyzed the data and designed and coordinated the research. The manuscript was written by J.S.D. with contributions and approval by all authors.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

This work was funded by the Deutsche Forschungsgemeinschaft DFG (project number 469042295) and supported by the computing center of the University of Cologne (RRZK), providing CPU time on the DFG-funded supercomputer CHEOPS.

## ■ REFERENCES

- (1) Croteau, R.; Purkett, P. T. Geranyl pyrophosphate synthase: Characterization of the enzyme and evidence that this chain-length specific prenyltransferase is associated with monoterpene biosynthesis in sage (*Salvia officinalis*). *Arch. Biochem. Biophys.* **1989**, *271*, 524–535.
- (2) Lynen, F.; Eggerer, H.; Henning, U.; Kessel, I. Farnesyl-pyrophosphat und 3-Methyl- $\Delta^3$ -butenyl-1-pyrophosphat, die biologischen Vorstufen des Squalens. Zur Biosynthese der Terpene, III. *Angew. Chem.* **1958**, *70*, 738–742.
- (3) Nandi, D. L.; Porter, J. W. The enzymatic synthesis of geranyl geranyl pyrophosphate by enzymes of carrot root and pig liver. *Arch. Biochem. Biophys.* **1964**, *105*, 7–19.
- (4) Tachibana, A. A novel prenyltransferase, farnesylgeranyl diphosphate synthase, from the haloalkaliphilic archaeon, *Natronobacterium pharaonis*. *FEBS Lett.* **1994**, *341*, 291–294.
- (5) Tao, H.; Lauterbach, L.; Bian, G.; Chen, R.; Hou, A.; Mori, T.; Cheng, S.; Hu, B.; Lu, L.; Mu, X.; Li, M.; Adachi, N.; Kawasaki, M.; Moriya, T.; Senda, T.; Wang, X.; Deng, Z.; Abe, I.; Dickschat, J. S.; Liu, T. Discovery of non-squalene triterpenes. *Nature* **2022**, *606*, 414–419.
- (6) (a) Minami, A.; Ozaki, T.; Liu, C.; Oikawa, H. Cyclopentane-forming di/esterterpene synthases: widely distributed enzymes in bacteria, fungi, and plants. *Nat. Prod. Rep.* **2018**, *35*, 1330–1346. (b) Christianson, D. W. Structural and Chemical Biology of Terpenoid Cyclases. *Chem. Rev.* **2017**, *117*, 11570–11648. (c) Dickschat, J. S. Bacterial terpene cyclases. *Nat. Prod. Rep.* **2016**, *33*, 87–110.
- (7) Gerber, N. N. A volatile metabolite of actinomycetes, 2-methylisoborneol. *J. Antibiot.* **1969**, *22*, 508–509.
- (8) Wang, C.-M.; Cane, D. E. Biochemistry and Molecular Genetics of the Biosynthesis of the Earthy Odorant Methylisoborneol in *Streptomyces coelicolor*. *J. Am. Chem. Soc.* **2008**, *130*, 8908–8909.
- (9) von Reuss, S. H.; Kai, M.; Piechulla, B.; Francke, W. Octamethylbicyclo[3.2.1]octadienes from the Rhizobacterium *Serratia odorifera*. *Angew. Chem., Int. Ed.* **2010**, *49*, 2009–2010.
- (10) von Reuss, S.; Domik, D.; Lemfack, M. C.; Magnus, N.; Kai, M.; Weise, T.; Piechulla, B. Sodorifen Biosynthesis in the Rhizobacterium *Serratia plymuthica* Involves Methylation and Cyclization of MEP-Derived Farnesyl Pyrophosphate by a SAM-Dependent C-Methyltransferase. *J. Am. Chem. Soc.* **2018**, *140*, 11855–11862.
- (11) Hu, Z.-Y.; Li, Y.-Y.; Huang, Y.-J.; Su, W.-J.; Shen, Y.-M. Three New Sesquiterpenoids from *Xylaria* sp. NCY2. *Helv. Chim. Acta* **2008**, *91*, 46–52.
- (12) Magnusson, G.; Thoren, S.; Wickberg, B. Fungal Extractives I. Structure of a Sesquiterpene Dialdehyde from *Lactarius* by Computer Simulation of the NMR Spectrum. *Tetrahedron Lett.* **1972**, *13*, 1105–1108.
- (13) Ayer, W. A.; Shan, R.; Trifonov, L. S.; Hutchison, L. J. Sesquiterpenes from the Nematicidal Fungus *Clitocybula oculus*. *Phytochemistry* **1998**, *49*, 589–592.
- (14) Murai, K.; Lauterbach, L.; Teramoto, K.; Quan, Z.; Barra, L.; Yamamoto, T.; Nonaka, K.; Shiomi, K.; Nishiyama, M.; Kuzuyama, T.; Dickschat, J. S. An Unusual Skeletal Rearrangement in the Biosynthesis of the Sesquiterpene Trichobrasilenol from *Trichoderma*. *Angew. Chem., Int. Ed.* **2019**, *58*, 15046–15050.
- (15) Sato, H.; Hashishin, T.; Kanazawa, J.; Miyamoto, K.; Uchiyama, M. DFT Study of a Missing Piece in Brasilane-Type Structure Biosynthesis: An Unusual Skeletal Rearrangement. *J. Am. Chem. Soc.* **2020**, *142*, 19830–19834.
- (16) Starks, C. M.; Back, K.; Chappell, J.; Noel, J. P. Structural Basis for Cyclic Terpene Biosynthesis by Tobacco 5-Epi-Aristolochene Synthase. *Science* **1997**, *277*, 1815–1820.
- (17) Shishova, E. Y.; Di Costanzo, L.; Cane, D. E.; Christianson, D. W. X-ray Crystal Structure of Aristolochene Synthase from *Aspergillus terreus* and Evolution of Templates for the Cyclization of Farnesyl Diphosphate. *Biochemistry* **2007**, *46*, 1941–1951.



(18) Baer, P.; Rabe, P.; Fischer, K.; Citron, C. A.; Klapschinski, T. A.; Groll, M.; Dickschat, J. S. Induced Fit Mechanism in Class I Terpene Cyclases. *Angew. Chem., Int. Ed.* **2014**, *53*, 7652–7656.

(19) Cornforth, J. W.; Cornforth, R. H.; Popjak, G.; Yengoyan, L. Studies on the Biosynthesis of Cholesterol. XX. Steric course of decarboxylation of 5-pyrophosphomevalonate and of the carbon to carbon bond formation in the biosynthesis of farnesylpyrophosphate. *J. Biol. Chem.* **1966**, *241*, 3970–3987.

(20) Cornforth, J. W.; Cornforth, R. H.; Donninger, C.; Popjak, G. Studies on the biosynthesis of cholesterol. XIX. Steric course of hydrogen eliminations and of C-C bond formations in squalene biosynthesis. *Proc. R. Soc. London B* **1966**, *136*, 492–514.

(21) Xu, H.; Goldfuss, B.; Dickschat, J. S. 1,2- or 1,3-Hydride Shifts - What controls Guaiane Biosynthesis. *Chem. Eur. J.* **2021**, *27*, 9758–9762.

(22) Sato, H.; Teramoto, K.; Masumoto, Y.; Tezuka, N.; Sakai, K.; Ueda, S.; Totsuka, Y.; Shinada, T.; Nishiyama, M.; Wang, C.; Kuzuyama, T.; Uchiyama, M. Cation-Stitching Cascade": exquisite control of terpene cyclization in cyclooctatin biosynthesis. *Sci. Rep.* **2016**, *5*, 18471.

## Recommended by ACS

### Enantioselective Total Syntheses of Grayanane Diterpenoids and (+)-Kalmanol: Evolution of the Bridgehead Carbocation-Based Cyclization and Late-Stage Functiona...

Lingran Kong, Tuoping Luo, *et al.*

APRIL 24, 2023

THE JOURNAL OF ORGANIC CHEMISTRY

READ 

### Total Synthesis and Structural Revision of the 6,11-Epoxyisodaucane Natural Sesquiterpene Using an Anionic 8 $\pi$ Electrocyclic Reaction

Ranmaru Kato, Keiji Tanino, *et al.*

OCTOBER 21, 2022

ORGANIC LETTERS

READ 

### Chemical Emulation of the Biosynthetic Route to Anthrasteroids: Synthesis of Asperfloketal A

Mykhaylo Alekseychuk and Philipp Heretsch

NOVEMBER 21, 2022

JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

READ 

### Reconsideration of the Structures of Stemara-13(14)-en-18-ol and Related Diterpene Natural Products: Vinylic Hydrogen Chemical Shifts Are Key

Amy T. Merrill and Dean J. Tantillo

AUGUST 11, 2022

JOURNAL OF NATURAL PRODUCTS

READ 

Get More Suggestions >

## Appendix M

**Fragmentation and [4 + 3] cycloaddition in sodorifen biosynthesis**

*Nat. Chem.* **2023**, *15*, accepted

DOI: 10.1038/s41557-023-01223-z

# Fragmentation and [4 + 3] cycloaddition in sodorifen biosynthesis

Received: 4 November 2022

Accepted: 26 April 2023

Published online: 29 May 2023

 Check for updatesHouchao Xu<sup>1</sup>, Lukas Lauterbach<sup>1</sup>, Bernd Goldfuss<sup>2</sup>,  
Gregor Schnakenburg<sup>3</sup> & Jeroen S. Dickschat<sup>1</sup>✉

Terpenes constitute the largest class of natural products. Their skeletons are formed by terpene cyclases (TCs) from acyclic oligoprenyl diphosphates through sophisticated enzymatic conversions. These enzyme reactions start with substrate ionization through diphosphate abstraction, followed by a cascade reaction via cationic intermediates. Based on isotopic-labelling experiments in combination with a computational study, the cyclization mechanism for sodorifen, a highly methylated sesquiterpene from the soil bacterium *Serratia plymuthica*, was resolved. A peculiar problem in its biosynthesis lies in the formation of several methyl groups from chain methylene carbons. The underlying mechanism involves a methyltransferase-mediated cyclization and unprecedented ring contraction with carbon extrusion from the chain to form a methyl group. A terpene cyclase subsequently catalyses a fragmentation into two reactive intermediates, followed by hydrogen transfers between them and recombination of the fragments by [4 + 3] cycloaddition. This study solves the intricate mechanistic problem of extra methyl group formation in sodorifen biosynthesis.

Soil microorganisms release a large diversity of volatile organic compounds with important biological activities such as plant-growth-promoting effects<sup>1</sup> or increased bacterial resistance towards antibiotics<sup>2</sup>. Recently, the volatiles produced by rhizobacteria from the genus *Serratia* have come into the focus because they interfere with bacterial cell-to-cell communication systems<sup>3</sup> and inhibit the growth of *Arabidopsis thaliana* and of plant pathogenic fungi<sup>4</sup>. The active constituent in the bouquet of *Serratia* is not known, but research resulted in the discovery of the polymethylated hydrocarbon sodorifen (**1**) from *Serratia plymuthica* 4Rx13 (Fig. 1)<sup>5</sup>. Feeding experiments with [2-<sup>13</sup>C] acetate<sup>5</sup> and gene knockouts of a putative methyltransferase (MT) and a terpene cyclase (TC) gene encoded in a small biosynthetic gene cluster for **1** suggested its terpenoid origin<sup>6,7</sup>. This was confirmed by the in vitro conversion of farnesyl diphosphate (FPP) and *S*-adenosylmethionine (SAM) with recombinant MT and TC into **1** (ref. 8) and by heterologous gene expression in *Escherichia coli*<sup>9</sup>. The MT initiates a cyclization reaction by the addition of CH<sub>3</sub><sup>+</sup> to FPP that ultimately leads to presodorifen

diphosphate (**2**, Fig. 1a). This intermediate is further converted into **1** by the TC in a second cyclization cascade<sup>8</sup>. Based on isotopic-labelling experiments, a mechanistic hypothesis for sodorifen biosynthesis has been suggested<sup>8</sup>, but the published mechanism is doubtful because it proceeds through primary cations. Here we report on an experimentally and computationally refined model for the biosynthesis of **1** that includes several unprecedented mechanistic steps.

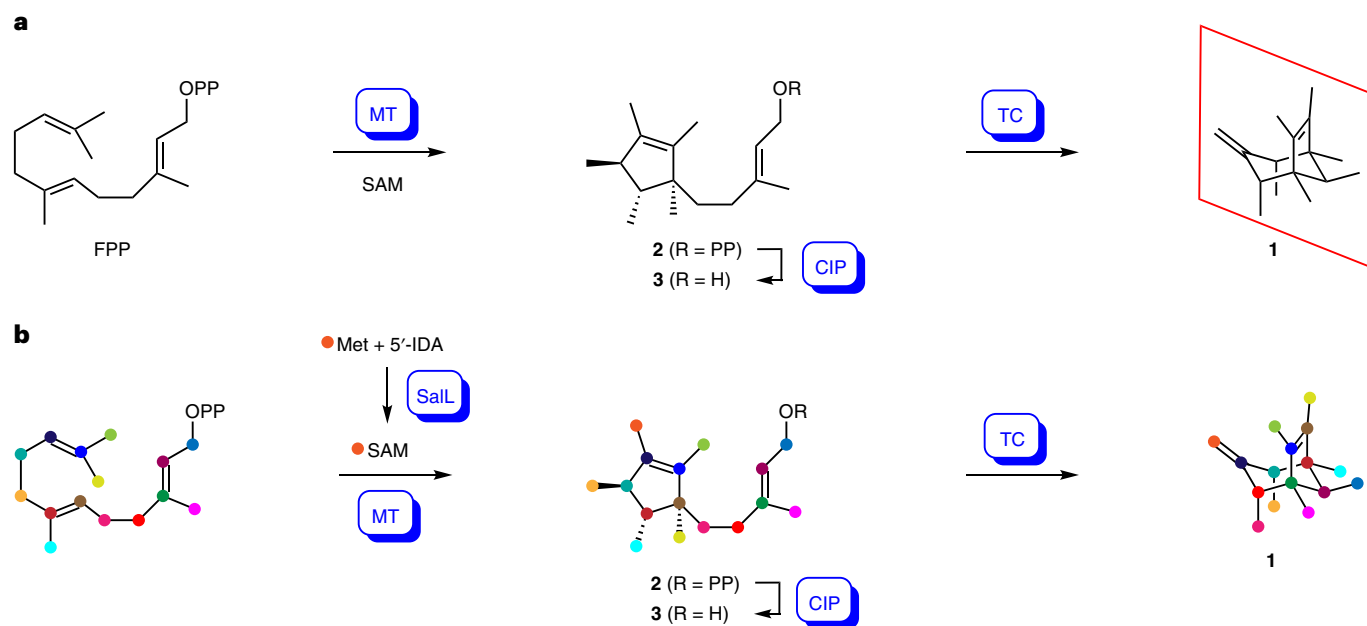
## Results and discussion

### Cyclization to presodorifen diphosphate

To study the biosynthesis of **1**, the genes for the MT and TC from *S. plymuthica* PRI-2C were cloned and expressed in *E. coli*. The purified enzymes converted FPP and SAM into **1**, while usage of only the MT resulted in the formation of **2**. Compound **2** is difficult to isolate, but after dephosphorylation with calf intestinal phosphatase (CIP) the corresponding alcohol presodorifen (**3**) can be obtained (Fig. 1a, Supplementary Figs. 1–18 and Supplementary Tables 1 and 2).

<sup>1</sup>Kekulé-Institut für Organische Chemie und Biochemie, Rheinische Friedrich-Wilhelms-Universität Bonn, Bonn, Germany. <sup>2</sup>Institut für Organische Chemie, Universität zu Köln, Köln, Germany. <sup>3</sup>Institut für Anorganische Chemie, Rheinische Friedrich-Wilhelms-Universität Bonn, Bonn, Germany.

✉e-mail: [dickschat@uni-bonn.de](mailto:dickschat@uni-bonn.de)



**Fig. 1 Biosynthesis of sodorifen (1).** **a**, The cyclization of FPP to presodorifen diphosphate (**2**) is initiated by methylation which is followed by terpene cyclization to **1**, a compound with an internal plane of symmetry (red).

Dephosphorylation of **2** with CIP allows the isolation of presodorifen (**3**). **b**,  $^{13}\text{C}$ -labelling experiments reveal the biosynthetic origin of all carbons in **2** and in **1** from FPP and SAM.

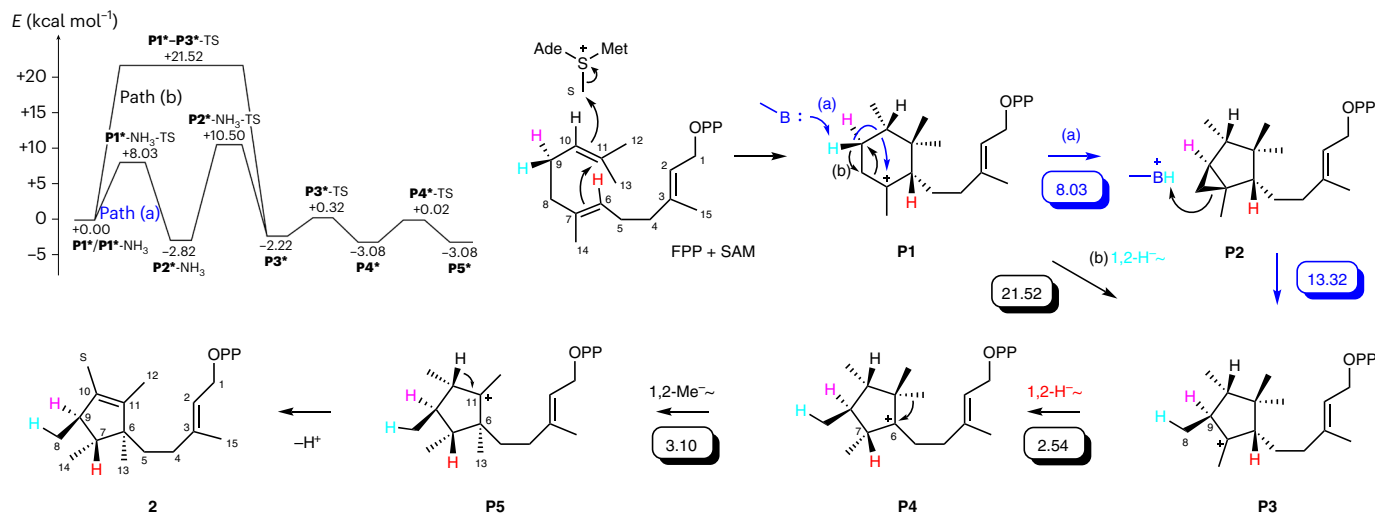
The biosynthesis of **1** was studied by isotopic-labelling experiments with the aim of addressing every single carbon and hydrogen atom of the substrates FPP and SAM, with the exception of methyl group hydrogens (Supplementary Table 3). For this purpose, first the biosynthesis of **2** was investigated by the conversion of all 15 isotopomers of ( $^{13}\text{C}$ )FPP<sup>10</sup> and of (*methyl*- $^{13}\text{C}$ )SAM with MT (Fig. 1b and Supplementary Fig. 19). Here, (*methyl*- $^{13}\text{C}$ )SAM was enzymatically generated from (*methyl*- $^{13}\text{C}$ )-L-methionine and 5'-iodo-5'-deoxyadenosine (5'-IDA) using the SAM chlorinase SalL from salinosporamide A biosynthesis<sup>11</sup>. In particular, these experiments established the methylation of FPP at C-10, the extrusion of the methylene group C-8 that becomes a methyl group in **3**, and the specific migration of one of the geminal methyl groups (C-13) from C-11 to C-6. The carbon derived from SAM is designated C-S.

Specific labelling experiments were required to investigate the C-8 extrusion in this process. The substrate (4,8,10- $^{13}\text{C}_3$ ,9,9- $^2\text{H}_2$ )FPP was enzymatically prepared from (2- $^{13}\text{C}$ ,1,1- $^2\text{H}_2$ )dimethylallyl diphosphate<sup>12</sup> (DMAPP) and (4- $^{13}\text{C}$ )isopentenyl diphosphate (IPP)<sup>13</sup> using the FPP synthase (FPPS) from *Streptomyces coelicolor* A3(2)<sup>12</sup> (Supplementary Fig. 20). Further conversion into **3** yielded a product that exhibited an upfield shifted triplet for C-8 in the  $^{13}\text{C}$  NMR spectrum, indicating a direct  $^{13}\text{C}$ - $^2\text{H}$  bond and confirming the origin of one hydrogen at C-8 in **2** from C-9 of FPP. The ring contraction was further studied for its stereochemical course using (*R*)- and (*S*)-(1- $^{13}\text{C}$ ,1- $^2\text{H}$ )IPP<sup>14</sup>. These substrates were converted with the isopentenyl diphosphate isomerase (IDI) from *S. plymuthica* PRI-2C<sup>15</sup>, FPPS (with inversion of configuration at C-1 of each allyl pyrophosphate intermediate<sup>16</sup>), MT and CIP into **3** (Supplementary Fig. 21), revealing a selective shift of the 9-*pro-R* hydrogen of FPP towards C-8, while the 9-*pro-S* hydrogen stayed at C-9. The same experiment also showed that both hydrogens at C-1 and C-5 remained bound (Supplementary Figs. 22 and 23). The incubation of (3- $^{13}\text{C}$ ,2- $^2\text{H}$ )GPP<sup>15</sup> and IPP with FPPS to obtain (7- $^{13}\text{C}$ ,6- $^2\text{H}$ )FPP, followed by enzymatic conversion into labelled **3** demonstrated that the C-7 hydrogen in **2** originates from the H-6 of FPP (Supplementary Fig. 24). Finally, the transformation from FPP to **3** was shown by gas chromatography–mass spectrometry (GC–MS) analysis to proceed without incorporation of deuterium from  $\text{D}_2\text{O}$ , and with loss of a deuterium from C-10 by incubation of (10- $^2\text{H}$ )FPP, enzymatically prepared from (2- $^2\text{H}$ )DMAPP<sup>17</sup> and IPP

with FPPS (Supplementary Fig. 25). Taken together, these experiments support a cyclization mechanism catalysed by the MT that starts with the transfer of  $\text{CH}_3^+$  from SAM to C-10 of FPP (Fig. 2). This induces a 6,11-cyclization to **P1**, followed by a dyotropic rearrangement<sup>18–20</sup> with specific migration of the 9-*pro-R* hydrogen (cyan) and extrusion of C-8 to form an additional methyl group in **P3**. Alternatively, this step could be substituted by a deprotonation with cyclopropanation to **P2** by an active site base, followed by reprotonation with transfer of the same proton and opening of the cyclopropane ring to **P3**. A 1,2-hydride migration from C-6 to C-7 (red) leads to **P4**, and 1,2-methyl group migration of Me-13 from C-11 to C-6 results in **P5** that yields **2** upon deprotonation from C-10.

### Cyclization to sodorifen

The mechanism of the formation of **1** was studied using MT and TC in conjunction with all 15 isotopomers of ( $^{13}\text{C}$ )FPP and (*methyl*- $^{13}\text{C}$ )SAM, showing specific incorporation into **1** in all cases (Supplementary Fig. 26). The plane of symmetry in **1** only allowed for unambiguous conclusions regarding the carbons within this plane (C-1, C-2, C-10 and C-S), while for all other cases two equal positions had to be considered. To resolve these ambiguities, the carbon connections were investigated with multiply  $^{13}\text{C}$ -labelled probes. The enzymatic conversion of (1,2,3,4,15- $^{13}\text{C}_5$ )FPP (Supplementary Scheme 1) yielded a product for which a contiguous spin system from C-4 and C-15 to C-1 was observed, establishing the direct connections between these carbons and their location within the plane of symmetry (C-1, C-2) and one hemisphere of **1** (C-3, C-4, C-15; Supplementary Fig. 27). In conclusion, C-7, C-9 and C-14 were assigned to the opposite side of **1**. The conversion of (4,5- $^{13}\text{C}_2$ )FPP, prepared from (1- $^{13}\text{C}$ )GPP<sup>21</sup> and (4- $^{13}\text{C}$ )IPP<sup>13</sup> with FPPS, gave **1** exhibiting a  $^1J$  coupling between the labelled carbons (Supplementary Fig. 28), which placed C-5 next to C-4 and, in conclusion, C-8 at the opposite half of **1**. Analogously, the conversion of (3,11- $^{13}\text{C}_2$ )FPP, made from (7- $^{13}\text{C}$ )GPP<sup>22</sup> and (3- $^{13}\text{C}$ )IPP<sup>13</sup> with FPPS, placed C-11 next to C-3 and C-6 in the opposite hemisphere of **1** (Supplementary Fig. 29). Finally, the enzymatic transformation of (11,13- $^{13}\text{C}_2$ )FPP (Supplementary Scheme 2) yielded a product with a small  $^2J$  coupling between the labelled carbons, indicating that C-11 and C-13 are not directly connected, that is, C-12 must be



**Fig. 2 | Cyclization mechanism from FPP to presodorifen diphosphate (2).** All explicitly drawn hydrogen atoms were followed by deuterium-labelling experiments. Computed reaction barriers (Gibbs energies at 298 K in kcal mol<sup>-1</sup>) are shown in boxes. Carbon numbering in **2** follows that of FPP to indicate the

origin of each carbon by the same number. The carbon derived from SAM is labelled 'S'. The computed energy profile is shown in the box (for an enlarged version, see Supplementary Fig. 42).

bound to C-11 (Supplementary Fig. 30). In summary, one hemisphere of **1** is constructed from carbons 15-3(-11-12)-4-5 and the other hemisphere is built from carbons 14-7(-6-13)-9-8, allowing for a reinterpretation of the single labelling experiments with the 15 isotopomers of (<sup>13</sup>C)FPP and (*methyl*-<sup>13</sup>C)SAM (Fig. 1b and Supplementary Fig. 31).

Additional deuterium-labelling experiments were performed to investigate the biosynthetic origin for the hydrogens of **1**. The enzymatic conversion of (*R*)- and (*S*)-(1-<sup>13</sup>C,1-<sup>2</sup>H)IPP with IDI, FPPS, MT and TC showed that the hydrogens at the C-1 and C-5 of FPP remain at their original carbons in **1** (Supplementary Figs. 32 and 33). Also, the 9-*pro-S* hydrogen did not migrate, while a specific shift of the 9-*pro-R* hydrogen into a neighbouring position (C-8) was indicated by a small upfield shift for C-9 (Supplementary Fig. 34). Direct evidence for its destination was obtained with (4,8,10-<sup>13</sup>C<sub>3</sub>,9,9-<sup>2</sup>H<sub>2</sub>)FPP, made from (2-<sup>13</sup>C,1,1-<sup>2</sup>H<sub>2</sub>)DMAPP and (4-<sup>13</sup>C)IPP with FPPS, yielding a triplet for C-8 (Supplementary Fig. 35). This reflects the situation for C-8 and C-9 observed for **2**, suggesting that the conversion of **2** into **1** does not further affect these positions. GPP and (2-<sup>13</sup>C,2,2-<sup>2</sup>H<sub>2</sub>)IPP (Supplementary Scheme 3) were converted with FPPS into (2-<sup>13</sup>C,2-<sup>2</sup>H)FPP, followed by transformation with MT and TC, revealing that the hydrogen at C-2 remains bound to this carbon in **1** (Supplementary Fig. 36). The MT- and TC-catalysed conversion of (1,5-<sup>13</sup>C<sub>2</sub>,2,6-<sup>2</sup>H<sub>2</sub>)FPP, generated with FPPS from DMAPP and (1-<sup>13</sup>C,2,2-<sup>2</sup>H<sub>2</sub>)IPP (Supplementary Scheme 4), established the movement of H-6 of FPP (H-6) to C-5 (Supplementary Fig. 37). A stereoselective deuterium labelling at C-4 and C-8 of FPP can be introduced using (*Z*)- or (*E*)-(4-<sup>13</sup>C,4-<sup>2</sup>H)IPP that react in the elongations of allyl diphosphates with specific attack from the *Si* face<sup>23</sup>. To investigate the fate of the hydrogens connected to C-4, GPP was reacted with (*Z*)- or (*E*)-(4-<sup>13</sup>C,4-<sup>2</sup>H)IPP under FPPS catalysis, followed by cyclization to labelled **1**, showing that the 4-*pro-R* hydrogen stays at C-4, whereas the 4-*pro-S* hydrogen migrates away to a distal place without noticeable influence on the chemical shift of C-4 (Supplementary Fig. 38). A similar experiment with DMAPP and (*Z*)- or (*E*)-(4-<sup>13</sup>C,4-<sup>2</sup>H)IPP demonstrated that both hydrogens at C-8 of FPP do not shift (Supplementary Fig. 39). The target atom of the migrating 4-*pro-S* hydrogen cannot be C-3, C-6, C-7, C-10 or C-11 (these are quaternary), C-2 (occupied by H-2), C-5 (occupied by 2 × H-5 and H-6), C-8 (occupied by 2 × H-8 and H-9-*pro-R*) or C-9 (occupied by H-9-*pro-S*). Carbons C-12, C-13, C-14, C-15 and C-S were considered unlikely, as they arise from the methyl groups of FPP and SAM. Both hydrogens at C-1 stay connected to this

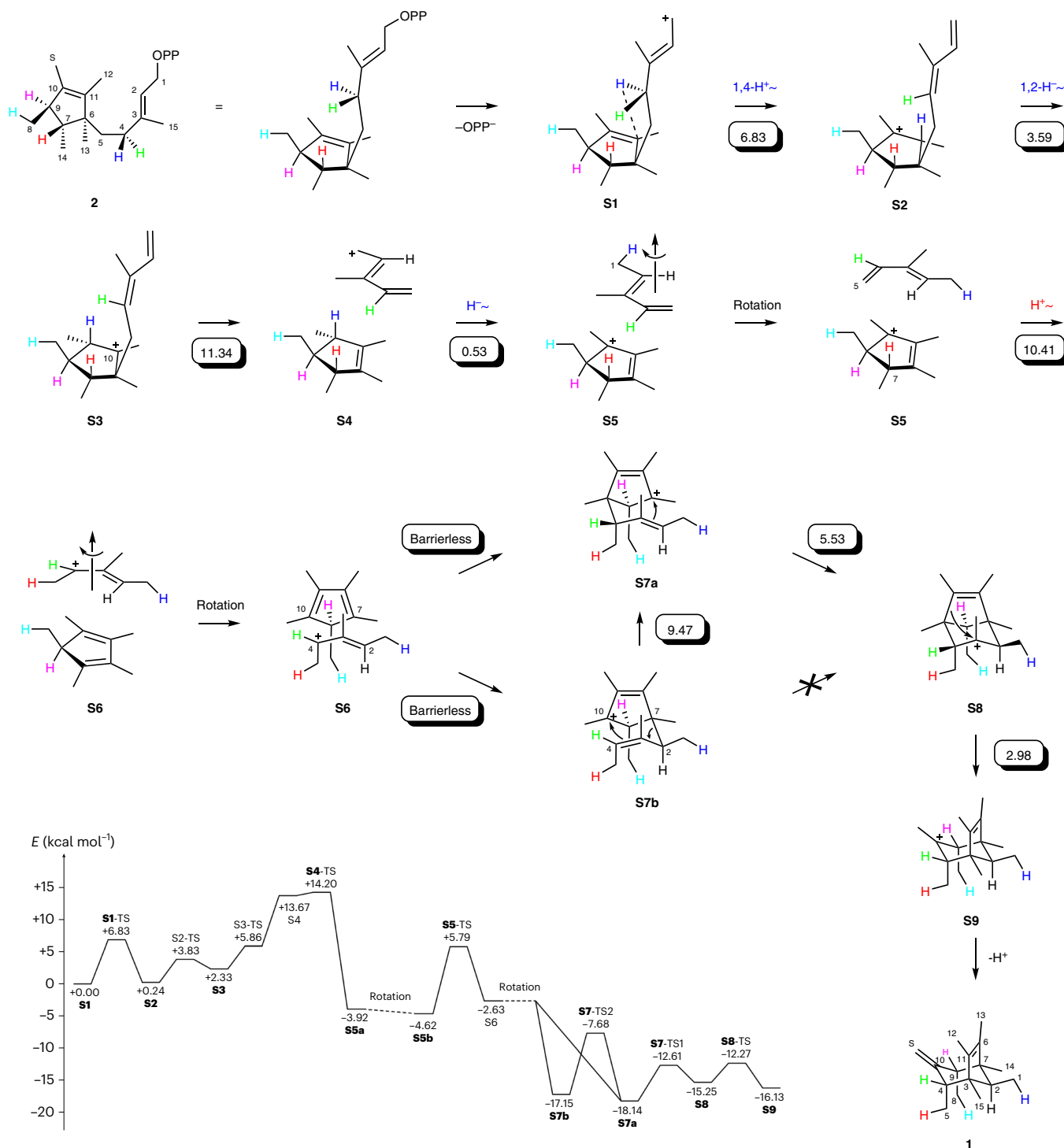
carbon, but the origin of the third hydrogen is so far unidentified. However, it does not stem from bulk water as shown by incubation in D<sub>2</sub>O (Supplementary Fig. 40), suggesting that the 4-*pro-S* hydrogen may migrate to C-1. Evidence for this assumption was obtained with (1-<sup>13</sup>C,4,4,4,5,5,5-<sup>2</sup>H<sub>6</sub>)DMAPP (Supplementary Scheme 5) that was used to elongate GPP by action of IDI and FPPS to yield (1-<sup>13</sup>C,4,4,15,15,15-<sup>2</sup>H<sub>5</sub>)FPP, followed by conversion into labelled **1** that showed a triplet signal for C-1 in the <sup>13</sup>C NMR (Supplementary Fig. 41).

Based on these results, a mechanistic model for the biosynthesis of **1** was developed (Fig. 3). The ionization of **2** by abstraction of diphosphate leads to the presodorifenyl cation (**S1**), which can undergo a 1,4-proton shift of the 4-*pro-S* hydrogen (blue) to **S2**. The same hydrogen is then transferred through a 1,2-hydride shift to C-10 in **S3** and—after fragmentation into a pentadienyl cation and a cyclopentene portion (**S4**)—to C-1 in **S5**, explaining its observed net migration from C-4 to C-1. Rotation of the pentadiene fragment brings C-5 into proximity of H-7 (red, originating from H-6 of FPP, Fig. 2), which allows for a proton transfer to yield **S6**. After further rotation, the two fragments are recombined in a [4 + 3] cycloaddition with attack of C-10 to C-4 and attack of C-2 to C-7, either concertedly or stepwise via **S7a** or **S7b**, to yield **S8**. Wagner–Meerwein rearrangement to **S9** and deprotonation lead to **1**.

### Density functional theory computations

For a deeper understanding of the mechanisms for the MT- and TC-catalysed reactions towards **1**, density functional theory computations were performed using the mPW1PW91/6-311 + G(d,p)//B97D3/6-31g(d,p) level of theory<sup>24,25</sup>, which proved of value in previous terpene biosynthesis studies<sup>26,27</sup>. The conversion of FPP to **2** was studied with the conformationally flexible side chain that does not directly participate in the reactions removed in all intermediates and transition states (TSs). The highest TS barrier was found for the dyotropic rearrangement (21.52 kcal mol<sup>-1</sup>; Fig. 2 and Supplementary Fig. 42). Substitution of this step by a base-catalysed cyclopropanation for which NH<sub>3</sub> was used in the calculations, and subsequent reprotonation induced opening of the cyclopropane ring in reverse order proceeded through much lower TS barriers (8.03 and 13.32 kcal mol<sup>-1</sup>, respectively). For the subsequent hydride and methyl shifts, low TS barriers (2.54 and 3.10 kcal mol<sup>-1</sup>, respectively) were obtained.

For the transformations from **2** to **1**, the calculations revealed low TS barriers for all steps along the cationic cascade, with the highest



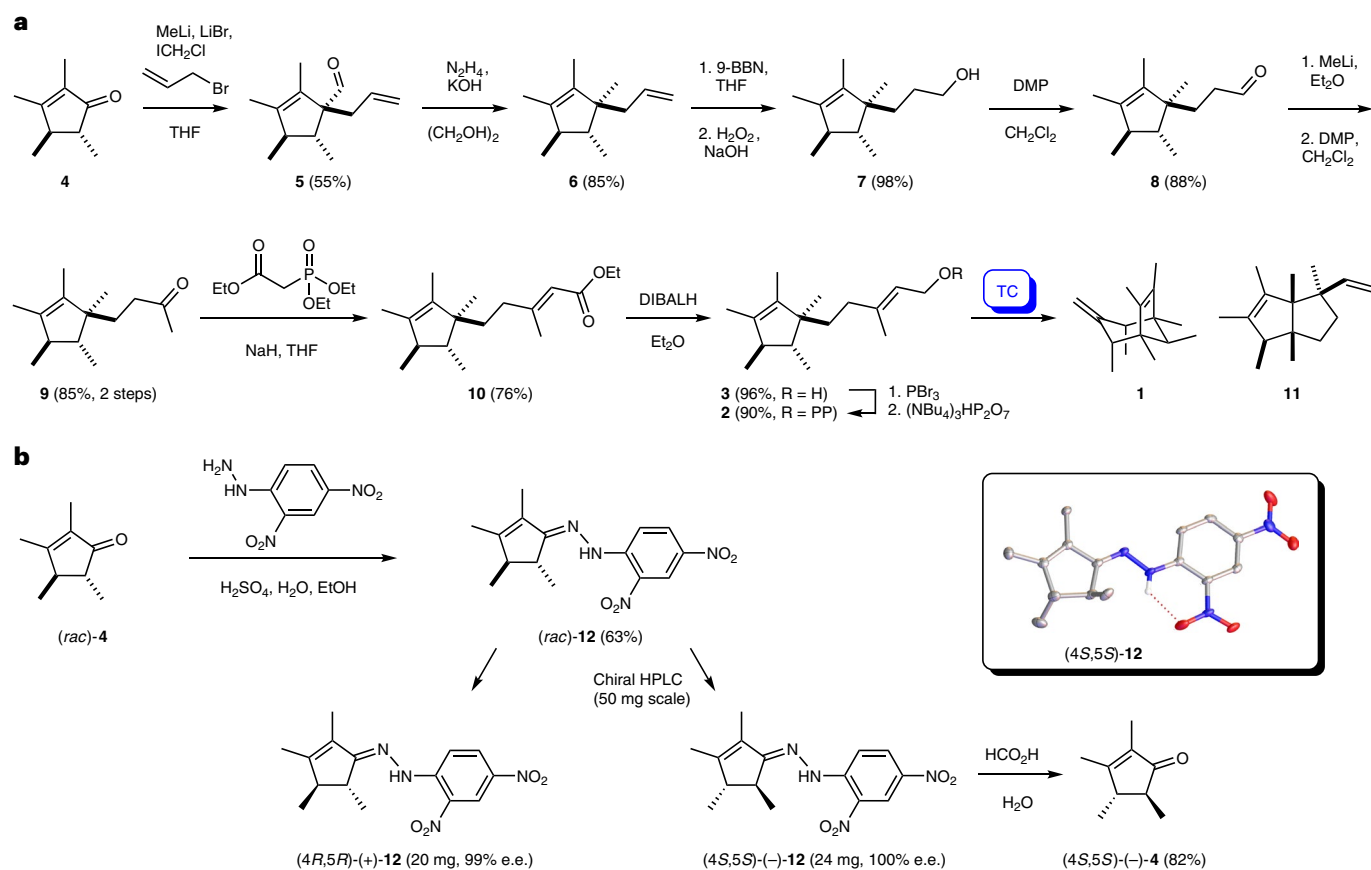
**Fig. 3 | Cyclization mechanism from presodorifen diphosphate (**2**) to sodorifen (**1**).** The explicitly shown hydrogen atoms were followed by deuterium-labelling experiments. Computed reaction barriers (Gibbs energies at 298 K in kcal mol<sup>-1</sup>) are shown in boxes. The computed energy profile is shown in the box (for an enlarged version, see Supplementary Fig. 43).

barrier of only 11.34 kcal mol<sup>-1</sup> for the fragmentation of **S3** (Fig. 3 and Supplementary Fig. 43). Notably, the [4 + 3] cycloaddition could not be realized computationally as a concerted reaction. Instead, a stepwise process is possible in which the first bond formation between C-4 and C-10 to **S7a** or between C-2 and C-7 to **S7b** is barrierless. Completion of the [4 + 3] cycloaddition to **S8** turned out to be only possible from **S7a**, but not from **S7b** which upon C-4/C-10 ring closure reacts with reopening of the bond on the other side (C-2/C-7) to **S7a**, a phenomenon that can be described as  $\pi$ -allyl rocking. The TS barrier for the unusual

Wagner–Meerwein rearrangement from **S8** to **S9** turned out to be very low (2.98 kcal mol<sup>-1</sup>). The whole cascade from **S1** to **S9** is exergonic ( $\Delta G = -16.13$  kcal mol<sup>-1</sup>).

#### Absolute configuration of presodorifen diphosphate

The absolute configuration of **2** was recently reported as (6*R*,7*S*,9*S*)-**2** (ref. 28) based on a comparison of measured to calculated electronic circular dichroism spectra. In contrast, our mechanisms for the biosynthesis of **1** make use of the enantiomer (6*S*,7*R*,9*R*)-**2**, for which the absolute



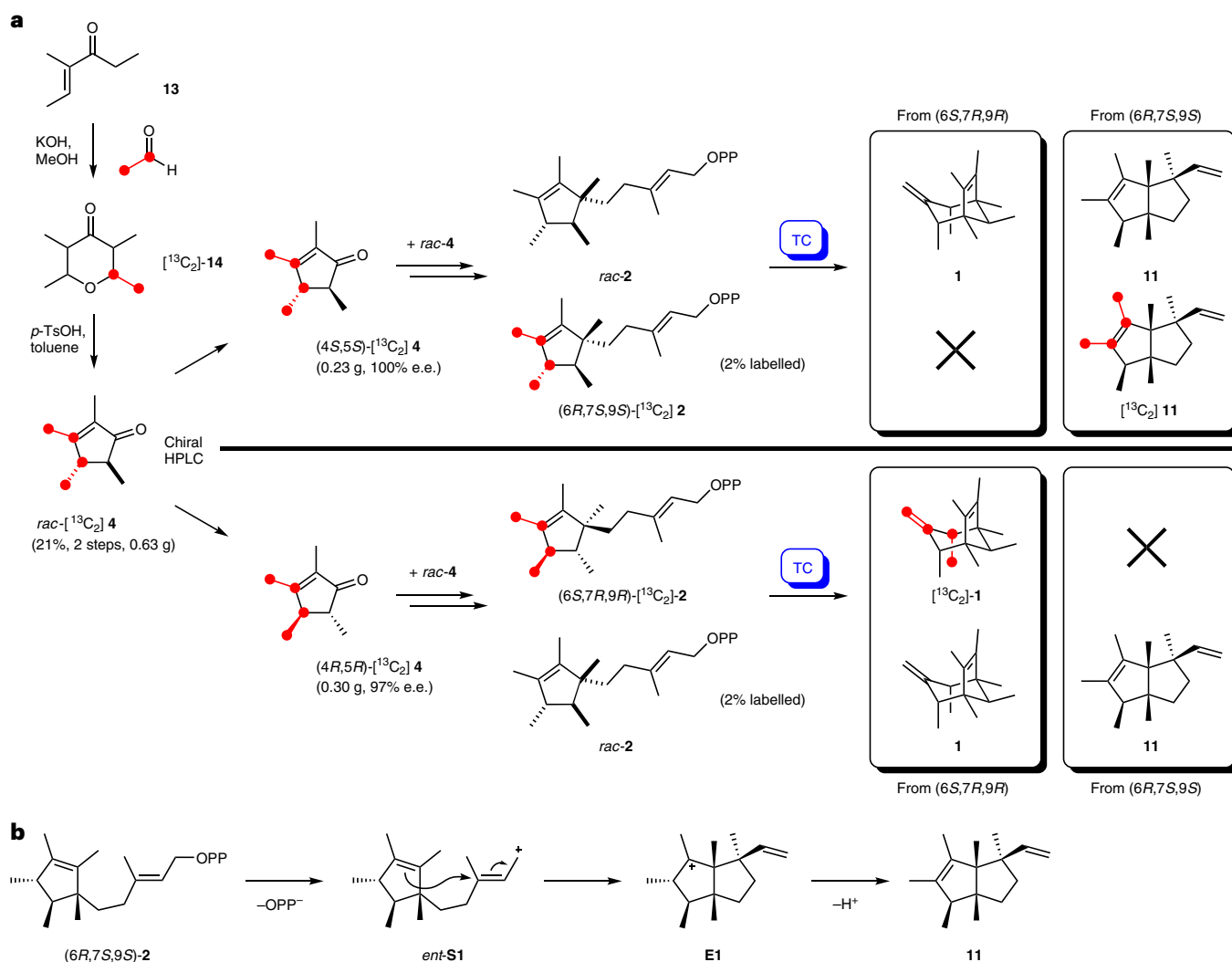
**Fig. 4 | The absolute configuration of **2**.** **a**, Synthesis of (*rac*)-**2** and enzymatic conversion into **1** and enantiofen (**11**). 9-BBN, 9-borabicyclo[3.3.1]nonane. **b**, Conversion of (*rac*)-**4** into dinitrophenylhydrazone **12**, separation of its enantiomers by chiral HPLC and hydrolysis of enantiomerically pure (4*S*,5*S*)-(-)-**12** to (4*S*,5*S*)-(-)-**4**. Box: Cu K $\alpha$  X-ray crystal structure of (4*S*,5*S*)-**12**.

configuration was concluded from the observed stereospecific hydrogen migrations from **P1** to **P2** (Fig. 2) and from **S1** to **S2** (Fig. 3). Alternative mechanisms for the MT and the TC reaction involving (6*R*,7*S*,9*S*)-**2** as intermediate would need to be in line with the isotopic-labelling experiments and computationally plausible, but such mechanisms could not be developed, causing doubt about the reported absolute configuration of **2**. Therefore, the absolute configuration of **2** was reinvestigated through a purely experimental approach (Fig. 4). For this purpose, (*rac*)-**2** was synthesized from known ketone **4** (ref. 29) (Fig. 4a) which was converted through a recently reported method<sup>30</sup> with carbenoid LiCH<sub>2</sub>Cl and allyl bromide into **5**. Wolff-Kishner reduction to **6** was followed by hydroboration and oxidative workup to yield alcohol **7**, which was further transformed by oxidation with Dess–Martin periodinane (DMP)<sup>31</sup> to **8**, addition of MeLi, and DMP oxidation to **9**. A Horner–Wadsworth–Emmons reaction<sup>32,33</sup> to **10** and reduction with diisobutylaluminium hydride (DIBALH) resulted in **3** which was converted into **2** by bromination and nucleophilic substitution. This synthesis made (*rac*)-**2** available with 23% yield over ten steps. Surprisingly, the incubation of (*rac*)-**2** with TC did not only yield **1**, giving direct evidence for **2** being the intermediate towards **1**, but also another compound **11** not observed in the reactions of FPP and SAM with MT and TC (Supplementary Fig. 44). Compound **11** was isolated and its structure, including the relative configuration, was elucidated by NMR spectroscopy (Supplementary Figs. 45–52 and Supplementary Table 4). Because **11** was probably formed from the non-natural enantiomer of **2**, this compound was named enantiofen.

Unfortunately, neither the enantiomers of **3** nor of any intermediate along the synthetic route were separable on various chiral stationary HPLC phases. However, after conversion of ketone **4** into the

dinitrophenylhydrazone **12** the enantiomers were separable by chiral HPLC (Supplementary Fig. 53) and the anomalous dispersion (Cu K $\alpha$ ) X-ray crystal structure of (4*S*,5*S*)-**12** was obtained (Fig. 4b, Supplementary Table 5 and Supplementary Fig. 54). Hydrolysis of the enantiomerically pure hydrazone (4*S*,5*S*)-(-)-**12** gave access to (4*S*,5*S*)-(-)-**4**.

The enantiomers of **4** were also separable by chiral HPLC, but their separation was difficult so that it was not possible to gain access to sufficient quantities of enantiomerically pure material to repeat the whole synthesis towards **2**. Therefore, an isotopic-labelling strategy was adopted in which ketone **13** was converted with (<sup>13</sup>C<sub>2</sub>)acetaldehyde into (<sup>13</sup>C<sub>2</sub>)-**14** (red dots, <sup>13</sup>C-labelled carbons; Fig. 5a). As a result of the symmetry of **14**, treatment with *p*-TsOH in boiling toluene resulted in (*rac*)-[<sup>13</sup>C<sub>2</sub>]-**4** in which the two <sup>13</sup>C-labelled carbons were distributed over two alternative positions indicated by the red bonds. A small amount of (*rac*)-[<sup>13</sup>C<sub>2</sub>]-**4** was separated by chiral HPLC, yielding (4*S*,5*S*)-[<sup>13</sup>C<sub>2</sub>]-**4** and (4*R*,5*R*)-[<sup>13</sup>C<sub>2</sub>]-**4** in high enantiomeric purity (100% and 97% e.e., respectively; Supplementary Fig. 55). The pure enantiomers were then mixed with unlabelled (*rac*)-**4** so that the resulting material was ~2% enriched in <sup>13</sup>C-labelling (1% in each of the two alternative positions). Both samples were then converted chemically into **2**, followed by incubation with TC and product isolation. With both samples unlabelled (*rac*)-**2** yielded a mixture of **1** and **11**, while the labelled material (6*R*,7*S*,9*S*)-[<sup>13</sup>C<sub>2</sub>]-**2** with the reported absolute configuration of **2** present in the first sample caused a <sup>13</sup>C enrichment in **11**, but not in **1** (Supplementary Figs. 56 and 57). Conclusively, (6*R*,7*S*,9*S*)-**2** is the precursor for **11**, but not for **1**, and the absolute configurations of natural presodorifen diphosphate (**2**) and presodorifen (**3**) must be revised as (6*S*,7*R*,9*R*). The circular dichroism spectrum of isolated **3** is very



**Fig. 5 | The absolute configuration of 2. a**, Labelling strategy to investigate the absolute configuration of natural 2. Red dots represent  $^{13}\text{C}$ -labelled carbons. For *rac*- $^{13}\text{C}_2$ -4 and materials derived therefrom, the  $^{13}\text{C}$ -labelling was distributed over the two  $\text{C}_2$  units connected by red bonds. **b**, Cyclization mechanism from (6*R*,7*S*,9*S*)-2 to 11. OPP, diphosphate.

similar to the previously published circular dichroism spectrum (Supplementary Fig. 58), demonstrating that (6*S*,7*R*,9*R*)-3 occurs also in the different strain *S. plymuthica* 4Rx13 (ref. 28). The absolute configuration of 2 was confirmed by the occurrence of labelling from (6*S*,7*R*,9*R*)- $^{13}\text{C}_2$ -2 in the second sample in 1, but not in 11. The cyclization mechanism to 11 can proceed by ionization of (6*R*,7*S*,9*S*)-2 through diphosphate abstraction to *ent*-S1, cyclisation to E1 and deprotonation (Fig. 5b).

## Conclusions

As shown in this study, sodorifen and its biosynthesis are highly unusual in many aspects. First, because of its internal symmetry plane sodorifen is one of only a few known achiral terpenes, similar to the case of 1,8-cineol<sup>34,35</sup>. Its biosynthesis proceeds with formation of two methyl groups from methylene carbons of FPP. The first one is formed in a methylation-induced cyclization catalysed by a methyltransferase, and the second one is formed upon substrate ionization by a terpene synthase, leading to a fragmentation with several proton and hydride shifts that ultimately result in an allyl cation and a pentamethylcyclopentadiene fragment. Their recombination proceeds through an asynchronous [4 + 3] cycloaddition, for which the enzyme is probably not needed to lower the energy barrier. A requirement for the active

site of the sodorifen synthase (TC) is to provide sufficient space for the reorientation of the two fragments before the [4 + 3] cycloaddition, and water needs to be rigorously excluded from the hydrophobic cavity to avoid premature cation quenching. Ultimately, sodorifen biosynthesis involves another unusual skeletal rearrangement. Several [4 + 2] cycloadditions and even a [6 + 4] cycloaddition catalysed by pericyclases are known from biosynthetic processes<sup>36,37</sup>, and [4 + 3] cycloadditions of allyl cations to butadienes have often been used in the synthesis of seven-membered rings<sup>38</sup>. Sodorifen biosynthesis involves an enzymatic [4 + 3] cycloaddition. Finally, the terpene synthase is not only able to accept the natural enantiomer of its substrate, but also the non-natural enantiomer that is ionized and cyclized into another polymethylated terpene, entiofen. Taken together, all these aspects make sodorifen synthase one of the most unusual biocatalysts known to mankind.

## Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41557-023-01223-z>.



## References

- Ryu, C.-M. et al. Bacterial volatiles promote growth in arabidopsis. *Proc. Natl Acad. Sci. USA* **100**, 4927–4932 (2003).
- Lee, H. H., Molla, M. N., Cantor, C. R. & Collins, J. J. Bacterial charity work leads to population-wide resistance. *Nature* **467**, 82–86 (2010).
- Chernin, L. et al. Quorum-sensing quenching by rhizobacterial volatiles. *Environ. Microbiol. Rep.* **3**, 698–704 (2011).
- Vespermann, A., Kai, M. & Piechulla, B. Rhizobacterial volatiles affect the growth of fungi and *Arabidopsis thaliana*. *Appl. Environ. Microbiol.* **73**, 5639–5641 (2007).
- von Reuss, S. H., Kai, M., Piechulla, B. & Francke, W. Octamethylbicyclo[3.2.1]octadienes from the rhizobacterium *Serratia odorifera*. *Angew. Chem. Int. Ed.* **49**, 2009–2010 (2010).
- Domik, D., Magnus, N. & Piechulla, B. Analysis of a new cluster of genes involved in the synthesis of the unique volatile organic compound sodorifen of *Serratia plymuthica* 4Rx13. *FEMS Microbiol. Lett.* **363**, fnw139 (2016).
- Domik, D. et al. A terpene synthase is involved in the synthesis of the volatile organic compound sodorifen of *Serratia plymuthica* 4Rx13. *Front. Microbiol.* **7**, 737 (2016).
- von Reuss, S. et al. Sodorifen biosynthesis in the rhizobacterium *Serratia plymuthica* involves methylation and cyclization of MEP-derived farnesyl pyrophosphate by a SAM-dependent C-methyltransferase. *J. Am. Chem. Soc.* **140**, 11855–11862 (2018).
- Duell, E. R. et al. Direct pathway cloning of the sodorifen biosynthetic gene cluster and recombinant generation of its product in *E. coli*. *Microb. Cell. Fact.* **18**, 32 (2019).
- Rabe, P. et al. Conformational analysis, thermal rearrangement and EI-MS-fragmentation mechanism of (1(10)E,4E,6S,7R)-germacradien-6-ol by <sup>13</sup>C-labeling experiments. *Angew. Chem. Int. Ed.* **54**, 13448–13451 (2015).
- Eustaquio, A. S., Pojer, F., Noel, J. P. & Moore, B. S. Discovery and characterization of a marine bacterial SAM-dependent chlorinase. *Nat. Chem. Biol.* **4**, 69–74 (2008).
- Rabe, P. et al. Terpene cyclases from social amoebae. *Angew. Chem. Int. Ed.* **55**, 15420–15423 (2016).
- Rabe, P. et al. Mechanistic investigations on two bacterial diterpene cyclases: spiroviolene synthase and tsukubadiene synthase. *Angew. Chem. Int. Ed.* **56**, 2776–2779 (2017).
- Rinkel, J. & Dickschat, J. S. Addressing the chemistry of germacrene A by isotope labeling experiments. *Org. Lett.* **21**, 2426–2429 (2019).
- Rinkel, J., Lauterbach, L., Rabe, P. & Dickschat, J. S. Two diterpene synthases for spiroalbatene and cembrene A from *Allokutzneria albata*. *Angew. Chem. Int. Ed.* **57**, 3238–3241 (2018).
- Cornforth, J. W., Cornforth, R. H., Donninger, C. & Popjak, G. Studies on the biosynthesis of cholesterol XIX. Steric course of hydrogen eliminations and of C–C bond formations in squalene biosynthesis. *Proc. R. Soc. London, Ser. B* **163**, 492–514 (1966).
- Rinkel, J. et al. Mechanisms of the diterpene cyclases β-pinacene synthase from *Dictyostelium discoideum* and hydroxyrene synthase from *Streptomyces clavuligerus*. *Chem. Eur. J.* **23**, 10501–10505 (2017).
- Reetz, M. T. Dytropic rearrangements, a new class of orbital-symmetry controlled reactions. Type I. *Angew. Chem. Int. Ed.* **11**, 129–130 (1972).
- Hugelshofer, C. L. & Magauer, T. Dytropic rearrangements in natural product total synthesis and biosynthesis. *Nat. Prod. Rep.* **34**, 228–234 (2017).
- Gutierrez, O. & Tantillo, D. J. Analogies between synthetic and biosynthetic reactions in which [1,2]-alkyl shifts are combined with other events: dytropic, Schmidt, and carbocation rearrangements. *J. Org. Chem.* **77**, 8845–8850 (2012).
- Li, H. & Dickschat, J. S. Isotopic labelling experiments and enzymatic preparation of iso-casbenes with casbene synthase from *Ricinus communis*. *Org. Chem. Front.* **9**, 795–801 (2022).
- Mitsuhashi, T., Rinkel, J., Okada, M., Abe, I. & Dickschat, J. S. Mechanistic characterization of two chimeric sesterterpene synthases from *Penicillium*. *Chem. Eur. J.* **23**, 10053–10057 (2017).
- Cornforth, J. W., Cornforth, R. H., Popjak, G. & Yengoyan, L. Studies on the biosynthesis of cholesterol XX. Steric course of decarboxylation of 5-pyrophosphomevalonate and of the carbon to carbon bond formation in the biosynthesis of farnesyl pyrophosphate. *J. Biol. Chem.* **241**, 3970–3987 (1966).
- Adamo, C. & Barone, V. Exchange functionals with improved long-range behavior and adiabatic connection methods without adjustable parameters: the mPW and mPW1PW models. *J. Chem. Phys.* **108**, 664–675 (1998).
- Matsuda, S. P. T., Wilson, W. K. & Xiong, Q. Mechanistic insights into triterpene synthesis from quantum mechanical calculations. Detection of systematic errors in B3LYP cyclization energies. *Org. Biomol. Chem.* **4**, 530–543 (2006).
- Hong, Y. J. & Tantillo, D. J. A maze of dytropic rearrangements and triple shifts: carbocation rearrangements connecting stemarene, stemodene, betaerdene, aphidicolene, and scopadulanol. *J. Org. Chem.* **83**, 3780–3793 (2018).
- Quan, Z., Hou, A., Goldfuss, B. & Dickschat, J. S. Mechanism of the bifunctional multiple product sesterterpene synthase AcAS from *Aspergillus calidoustus*. *Angew. Chem. Int. Ed.* **61**, e202117273 (2022).
- Lemfack, M. C. et al. Reaction mechanism of the farnesyl pyrophosphate C-methyltransferase towards the biosynthesis of pre-sodorifen pyrophosphate by *Serratia plymuthica* 4Rx13. *Sci. Rep.* **11**, 3182 (2021).
- Burger, U., Delay, A. & Mazonod, F. 1,2,3,4,5-Pentamethyl-5-acetyl-cyclopentadien-1,3, ein ungewöhnliches Keton. *Helv. Chim. Acta* **57**, 2106–2111 (1974).
- Pace, V. et al. Efficient access to all-carbon quaternary and tertiary α-functionalized homoallyl-type aldehydes from ketones. *Angew. Chem. Int. Ed.* **56**, 12677–12682 (2017).
- Dess, D. B. & Martin, J. C. Readily accessible 12-l-5<sup>1</sup> oxidant for the conversion of primary and secondary alcohols to aldehydes and ketones. *J. Org. Chem.* **48**, 4155–4156 (1983).
- Horner, L., Hoffmann, H., Wippel, H. G. & Klahre, G. Phosphinoxyde als olefinierungsreagenzien. *Chem. Ber.* **92**, 2499–2505 (1959).
- Wadsworth, W. S. & Emmons, W. D. The utility of phosphonate carbanions in olefin synthesis. *J. Am. Chem. Soc.* **83**, 1733–1738 (1961).
- Wise, M. L., Savage, T. J., Katahira, E. & Croteau, R. Monoterpene synthases from common sage (*Salvia officinalis*): cDNA Isolation, characterization, and functional expression of (+)-sabinene synthase, 1,8-cineol synthase, and (+)-bornyl diphosphate synthase. *J. Biol. Chem.* **273**, 14891–14899 (1998).
- Nakano, C., Kim, H.-K. & Ohnishi, Y. Identification of the first bacterial monoterpene cyclase, a 1,8-cineole synthase, that catalyzes the direct conversion of geranyl diphosphate. *ChemBioChem* **12**, 1988–1991 (2011).
- Jamieson, C. S., Ohashi, M., Liu, F., Tang, Y. & Houk, K. N. The expanding world of biosynthetic pericyclases: cooperation of experiment and theory for discovery. *Nat. Prod. Rep.* **36**, 698–713 (2019).
- Zhang, B. et al. Enzyme-catalysed [6+4] cycloadditions in the biosynthesis of natural products. *Nature* **568**, 122–126 (2019).
- Harmata, M. The (4+3)-cycloaddition reaction: simple allylic cations as dienophiles. *Chem. Commun.* **46**, 8886–8903 (2010).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with

the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2023

## Methods

### Strains and culture conditions

*S. plymuthica* PRI-2C was cultivated in TSB medium (17.0 g peptone from casein, 3.0 g peptone from soymeal, 2.5 g D-(+)-glucose, 5.0 g NaCl, 2.5 g K<sub>2</sub>HPO<sub>4</sub>, 1.0 l H<sub>2</sub>O, pH 7.3) at 28 °C. For agar plate cultures 15.0 g l<sup>-1</sup> agar-agar was added. *Saccharomyces cerevisiae* FY834 was cultivated in liquid YPAD medium (10.0 g yeast extract, 20.0 g peptone, 20.0 g glucose, 400 mg adenine sulphate, 1.0 l H<sub>2</sub>O) or on SM-URA agar plates (1.7 g yeast nitrogen base, 5.0 g ammonium sulfate, 20.0 g glucose, 770 mg nutritional supplement minus uracil, 20.0 g agar-agar, 1.0 l H<sub>2</sub>O) at 28 °C. *E. coli* BL21(DE3) was grown in LB medium (10.0 g tryptone, 5.0 g yeast extract, 5.0 g NaCl, 1.0 l H<sub>2</sub>O) at 37 °C. For agar plate cultures 16 g agar-agar was added. Kanamycin was used at a concentration of 50 µg ml<sup>-1</sup>. All media were autoclaved at 121 °C for 20 min prior to use.

### Gene cloning

The desired genes for the TC (accession number [ANS42760](#)) and the MT ([ANS42761](#)) were obtained by polymerase chain reaction (PCR) using Q5-High-Fidelity DNA polymerase (New England Biolabs) and freshly isolated genomic DNA from *S. plymuthica* PRI-2C as a template. Primers are listed in Supplementary Table 6. PCR standard conditions were used (initial denaturation at 98 °C for 1 min, 35 cycles with denaturation at 98 °C for 15 s, annealing at 62 °C for 30 s and elongation at 72 °C for 35 s, final elongation step at 72 °C for 5 min). The PCR products together with the linearized pYE-Express shuttle vector were used for a yeast homologous recombination by a standard protocol using PEG, LiOAc and salmon sperm DNA<sup>39,40</sup>. The transformed *S. cerevisiae* cultures were plated on SM-URA plates and grown for 3 days at 28 °C. Colonies were collected from the plates and plasmid DNA was isolated using Zymoprep Yeast Plasmid Miniprep II (Zymo Research). The isolated plasmid DNA was used for electroporation of *E. coli* BL21 (DE3) electrocompetent cells. The transformed *E. coli* was grown overnight at 37 °C on LB agar plates containing kanamycin (50 µg ml<sup>-1</sup>). Single colonies were selected to inoculate 8 ml LB with kanamycin. The resulting cultures were grown for 24 h to isolate plasmid DNA, yielding plasmid pYE-TC and pYE-MT which were checked by analytical digest and by sequencing.

### Gene expression

*E. coli* transformants harbouring the plasmid pYE-MT and pYE-TC were grown with shaking overnight at 37 °C in LB medium containing kanamycin. Expression cultures were inoculated using 20 ml l<sup>-1</sup> of preculture, followed by cultivation at 37 °C with shaking until an absorbance (600 nm wavelength) of 0.4–0.6 was reached. The cultures were cooled to 18 °C and protein expression was induced by addition of isopropylthiogalactoside (400 mM in water, 1 ml l<sup>-1</sup>), followed by shaking overnight at 18 °C. The cultures were centrifuged at 3,600g (4 °C), the supernatant was discarded and the cell pellet was resuspended in binding buffer (10 ml l<sup>-1</sup> culture; 20 mM Na<sub>2</sub>HPO<sub>4</sub>, 500 mM NaCl, 20 mM imidazole, 1 mM MgCl<sub>2</sub>, pH 7.4, 4 °C). The suspension was subjected to ultrasonication for cell lysis. The cell debris was removed by centrifugation (14,610g, 15 min, 4 °C) and the supernatant was loaded onto a Ni<sup>2+</sup>-nitrilotriacetic acid column (Super Ni-NTA, Generson). The column was washed with binding buffer (2 × 10 ml l<sup>-1</sup> culture), and the target protein was eluted using elution buffer (20 mM Na<sub>2</sub>HPO<sub>4</sub>, 500 mM NaCl, 500 mM imidazole, 1 mM MgCl<sub>2</sub>, pH 7.4, 10 ml l<sup>-1</sup> culture, 4 °C). Fractions containing protein were analysed by SDS-PAGE (Supplementary Figure 1) and used for incubation experiments.

### Enzymatic preparation of presodorifen

FPP (64 mg) and SAM (128 mg) were dissolved in aqueous NH<sub>4</sub>HCO<sub>3</sub> solution (25 mM, 10 ml), followed by addition of incubation buffer (50 mM Tris, 10 mM MgCl<sub>2</sub>, 20% glycerine (v/v), pH 8.2, 40 ml). The reaction was started by addition of MT elution fractions (20 ml) and

incubated at 28 °C for 3 h. Calf intestinal alkaline phosphatase (1,250 U) was added, followed by the incubation at 37 °C for 1 h. The reaction mixture was extracted with hexane (3 × 50 ml). The organic layers were dried with MgSO<sub>4</sub> and concentrated under reduced pressure. Column chromatography on AgNO<sub>3</sub>-coated silica gel with pentane/diethyl ether (2:1) yielded presodorifen (2.2 mg) as a colourless oil. Optical rotation:  $[\alpha]_D^{25} = +4.2$  (c 0.14). NMR data are given in Supplementary Table 1 and Supplementary Figs. 4–10.

### Enzymatic conversion of presodorifen diphosphate into sodorifen

Presodorifen diphosphate (80 mg) was dissolved in aqueous NH<sub>4</sub>HCO<sub>3</sub> solution (25 mM, 50 ml), followed by addition of incubation buffer (50 mM Tris, 10 mM MgCl<sub>2</sub>, 20% glycerine (v/v), pH 8.2, 125 ml). The reaction was started by addition of TC elution fractions (25 ml) and incubated at 28 °C overnight. The reaction mixture was extracted with pentane (3 × 100 ml). The organic layers were dried with MgSO<sub>4</sub> and concentrated under reduced pressure. Column chromatography on silica gel with pentane, followed by HPLC purification, yielded sodorifen (4.5 mg) and enantiofen (3.0 mg) as colourless oils.

### Sodorifen (1)

NMR data are given in Supplementary Table 2 and Supplementary Figs. 12–18.

### Enantiofen (11)

TLC (pentane):  $R_f = 0.77$ . Optical rotation:  $[\alpha]_D^{25} = +20.5$  (c 0.22, CH<sub>2</sub>Cl<sub>2</sub>). High-resolution MS (electron ionization):  $[M]^+$  calculated for C<sub>16</sub>H<sub>26</sub><sup>+</sup>,  $m/z$  218.2035; found,  $m/z$  218.2031. GC (HP-5MS):  $I = 1,470$ . MS (electron ionization, 70 eV): EI-MS (70 eV):  $m/z$  (%) = 218 (17), 203 (3), 190 (1), 175 (2), 162 (7), 149 (64), 137 (100), 121 (50), 105 (28), 91 (23), 79 (13), 67 (14), 55 (10), 41 (16). Infrared spectroscopy (diamond, attenuated total reflectance):  $\tilde{\nu}$  (cm<sup>-1</sup>) = 2,961 (s), 2,927 (s), 2,871 (m), 2,856 (m), 1,634 (w), 1,449 (m), 1,378 (m), 1,261 (w), 1,098 (m), 1,014 (m), 799 (w). NMR data are given in Supplementary Table 5 and Supplementary Figs. 46–52.

### Labelling experiments

Isotopic-labelling experiments were conducted with substrates (~1 mg each) dissolved in aqueous NH<sub>4</sub>HCO<sub>3</sub> solution (25 mM, 1 ml). Incubation buffer (5 ml) and enzyme elution fractions (2 ml for each enzyme) were added. For combinations of substrates and enzymes, see Supplementary Table 3. After incubation with shaking at 28 °C overnight the products were extracted with C<sub>6</sub>D<sub>6</sub> twice (600 µl and 200 µl). The combined extracts were dried with MgSO<sub>4</sub> and analysed by NMR and/or GC–MS.

### NMR spectroscopy

NMR spectra were recorded on a Bruker Avance I 500 MHz spectrometer or a Bruker Avance III HD 700 MHz cryospectrometer. Chemical shifts were referenced to the residual proton signal of the solvent for <sup>1</sup>H NMR and the <sup>13</sup>C signal for <sup>13</sup>C NMR. Data collections were done using Bruker Topspin 4.1.1 software, data analyses were performed with Bruker Topspin 4.1.1 or Mestrenova 8.

### GC–MS analyses

GC–MS analyses were carried out on a 7890B/5977A series gas chromatography/mass selective detector (Agilent). The GC was equipped with an HP5-MS fused silica capillary column (30 m, 0.25 mm inner diameter, 0.50 µm film; Agilent) and operated using the following settings: (1) inlet pressure, 77.1 kPa, helium at 23.3 ml min<sup>-1</sup>; (2) injection volume, 1 µl; (3) temperature programme, 5 min at 50 °C then increasing 10 °C min<sup>-1</sup> to 320 °C; (4) splitless or split ratio 50:1, 60 s valve time; and (5) carrier gas, helium at 1 ml min<sup>-1</sup>. The MS was operated with the following settings: (1) source, 230 °C; (2) transfer line, 250 °C; (3) quadrupole, 150 °C; and (4) electron energy, 70 eV. Data

collection and analysis were performed using MSD ChemStation D.02.00.237 software.

### Computational methods

All computed structures are geometry optimized without restrictions and are characterized as minima or as TS structures by frequency analyses, also providing Gibbs corrections, using the B97D3/6-31g(d,p) method with the density fitting approximation for s and p functions, including Grimme's empirical D3 dispersion correction<sup>41</sup> in Gaussian16 Revision B01<sup>42</sup>. For improved single-point energies, the mPW1PW91 functional is applied with the 6-311+G(d,p) basis set without density fitting and the ultrafine integration grid, as this method has been shown to be very reliable for examining carbocation cyclization and rearrangement reactions<sup>24–26,43</sup>. The Gibbs corrections include an entropic quasiharmonic treatment with a frequency cut-off value of 100.0 wavenumbers, according to Grimme, using a mixture of RRHO and free-rotor vibrational entropies<sup>44,45</sup>. Data analysis was performed using GoodVibes v.3.0.1.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

The authors declare that the main data supporting the findings of this study are available within the Article, Supplementary Videos, Supplementary Data, and Supplementary Information. Crystallographic data for the structure reported in this Article have been deposited at the Cambridge Crystallographic Data Centre, under deposition number CCDC 2213524. Copies of the data can be obtained free of charge via <https://www.ccdc.cam.ac.uk/structures/>. Data and plasmids described in this study can be obtained from the corresponding author on reasonable request. Source data are provided with this paper.

### References

39. Giets, R. D. & Schiestl, R. H. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc.* **2**, 31–34 (2007).
40. Dickschat, J. S., Pahirulzaman, K. A. K., Rabe, P. & Klapschinski, T. A. An improved technique for the rapid chemical characterisation of bacterial terpene cyclases. *ChemBioChem* **15**, 810–814 (2014).
41. Grimme, S., Ehrlich, S. & Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comp. Chem.* **32**, 1456–1465 (2011).
42. Frisch, M. J. et al. Gaussian 16, revision B.01 (Gaussian, 2016).
43. Lauterbach, L., Goldfuss, B. & Dickschat, J. S. Two diterpene syntheses from *Chryseobacterium*: chryseodiene synthase and wanjudiene synthase. *Angew. Chem. Int. Ed.* **59**, 11943–11947 (2020).
44. Grimme, S. Supramolecular binding thermodynamics by dispersion-corrected density functional theory. *Chem. Eur. J.* **18**, 9955–9964 (2012).
45. Luchini, G., Alegre-Requena, J. V., Guan, Y., Funes-Ardoiz, I. & Paton, R. S. GoodVibes v.3.0.1 (2019); <https://patonlab.com/robert-paton/>

### Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft (DFG) (DI1536/11-1, project number 469042295) and supported by the computing centre of the University of Cologne (RRZK), providing CPU time on the DFG-funded supercomputer CHEOPS. We thank P. Garbeva (Wageningen) for strain *S. plymuthica* PRI-2C, B. Piechulla (Rostock) and S. von Reuss (Neuchatel) for discussions, and Andreas Schneider (Bonn) for HPLC separations.

### Author contributions

H.X. and L.L. performed syntheses and labelling experiments. B.G. performed density functional theory computations. G.S. solved the X-ray structure of (4S,5S)-**12**. J.S.D. designed and supervised research and wrote the manuscript with additions by all other authors.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41557-023-01223-z>.

**Correspondence and requests for materials** should be addressed to Jeroen S. Dickschat.

**Peer review information** *Nature Chemistry* thanks the anonymous reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

# Appendix N

## A Detailed View on Geosmin Biosynthesis

*ChemBioChem* **2023**, 24, e202300101

DOI: 10.1002/cbic.202300101

# A Detailed View on Geosmin Biosynthesis

Houchao Xu<sup>[a]</sup> and Jeroen S. Dickschat<sup>\*[a]</sup>

The bacterial geosmin synthase is a fascinating bifunctional enzyme that has been discovered almost two decades ago. Several aspects of the cyclisation mechanism from FPP to geosmin are known, but a detailed picture of the stereochemical course of this reaction is unknown. This article reports on a deep investigation of the mechanism of geosmin synthase through isotopic labelling experiments. Furthermore, the effects

of divalent cations on geosmin synthase catalysis were investigated. The addition of cyclodextrin to enzymatic reactions, a molecule that can capture terpenes, suggests that the biosynthetic intermediate (1(10)*E*,5*E*)-germacradien-11-ol produced by the N-terminal domain is passed to the C-terminal domain not through a tunnel, but rather through release into the medium and uptake by the C-terminal domain.

## Introduction

Geosmin (**1**, Figure 1) is an earthy odorant that is responsible for the smell of freshly ploughed earth.<sup>[1]</sup> The compound is produced by soil bacteria and was first isolated from *Streptomyces griseus* by Gerber and Lechevalier in 1965,<sup>[2]</sup> followed by its structure elucidation shortly after.<sup>[3]</sup> Today the compound is known to be very widespread and was reported from actinobacteria,<sup>[4]</sup> cyanobacteria,<sup>[5]</sup> myxobacteria,<sup>[6]</sup> ascomycete fungi,<sup>[7]</sup> basidiomycota,<sup>[8]</sup> and amoebae.<sup>[9]</sup> Geosmin has also been reported from plants including the liverwort *Symphogyna bronngiartii*,<sup>[10]</sup> mosses,<sup>[11]</sup> buckwheat (*Fagopyrum tataricum*),<sup>[12]</sup> *Zea mays*,<sup>[13]</sup> and beetroot where it adds to its typical earthy flavour.<sup>[14]</sup> In rainbow trout,<sup>[15]</sup> shrimps,<sup>[16]</sup> and molluscs<sup>[17]</sup> it may be a contaminant of microbial origin which is a particular problem in aquaculture. Geosmin is also a constituent of the defence secretions of the polydesmid millipede *Niponia nodulosa*.<sup>[18]</sup> Several derivatives of geosmin have been reported from natural sources, including dehydrogeosmin (**2**) from the flower scent of the cactus *Ributia marsoneri*,<sup>[19]</sup> the epoxide **3** from the liverwort *Lophocolea bidentata*,<sup>[20]</sup> and oxygenated compounds such as **4** and **5** from an endophytic streptomycete.<sup>[21]</sup>

The biosynthesis of geosmin was a long standing problem. Initial speculations suggested that geosmin may be a degraded eudesmane sesquiterpene.<sup>[3]</sup> Subsequent feeding experiments with radioactively labelled precursors to *Streptomyces antibioticus* showed incorporations from [1-<sup>14</sup>C] and [2-<sup>14</sup>C]acetate, but not from [methyl-<sup>14</sup>C]methionine.<sup>[22]</sup> Based on the identification of the cometabolites (1(10)*E*,5*E*)-germacradien-11-ol (**6**) and dihydroagarofuran (**7**) from *Streptomyces citreus* a first detailed

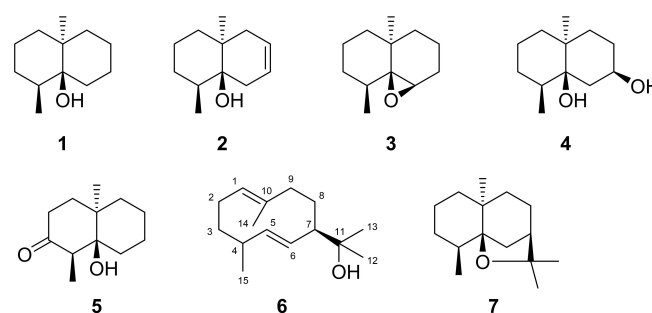


Figure 1. Structures of geosmin and its naturally occurring derivatives.

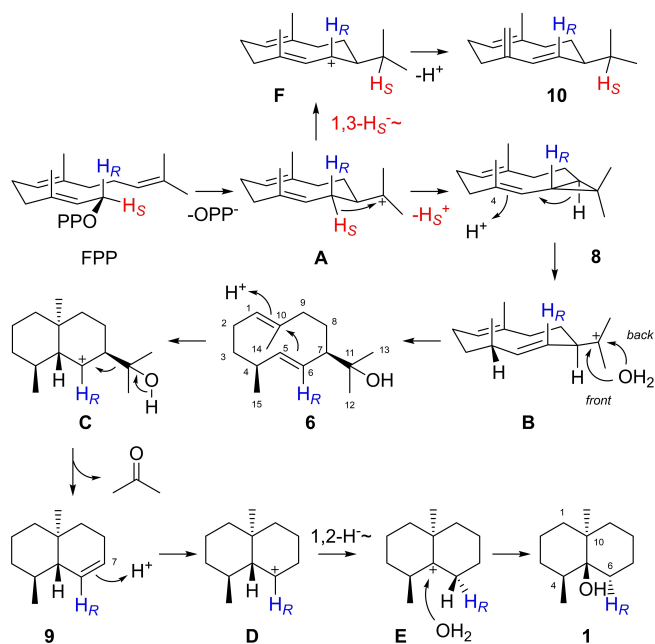
biosynthetic pathway through **6** and **7** as intermediates was proposed, but no satisfying mechanistic explanation was given.<sup>[23]</sup> At the time of the discovery of the geosmin synthase and its coding gene in 2003,<sup>[24]</sup> mechanistic proposals were raised that include oxidative degradation.<sup>[24b,25]</sup> Finally, based on feeding experiments with (<sup>2</sup>H<sub>10</sub>)leucine, (4,4,4,5,5,5-<sup>2</sup>H<sub>6</sub>)dimethylacrylate, and (4,4,6,6,6-<sup>2</sup>H<sub>5</sub>)mevalonic acid lactone a mechanism was developed that proceeds through the cyclisation of FPP to **6**, followed by reprotonation induced cyclisation to **C** that can undergo a retro-Prins reaction to the octalin **9** and acetone (Scheme 1); reprotonation to **D**, a 1,2-hydride shift to **E** and capture with water lead to **1**.<sup>[26]</sup> A deep investigation of the geosmin synthase from *Streptomyces coelicolor* (ScGS), an enzyme exhibiting two functional domains, through site-directed mutagenesis and in vitro experiments revealed that the N-terminal domain converts FPP into **6**, while the C-terminal domain further converts this intermediate into **1**.<sup>[27]</sup> Isotopic labelling experiments revealed that the deprotonation of **A** to **8** proceeds with specific loss of the 1-*pro-S* hydrogen of FPP. This is the same hydrogen that undergoes a 1,3-hydride shift from **A** to **F** on the pathway to the side product germacrene D (**10**).<sup>[28]</sup>

This mechanism was further supported by the rigorous structure elucidation of **9** through total synthesis<sup>[29]</sup> and the capture of acetone from an in vitro conversion of FPP.<sup>[30]</sup> Today also the crystal structure of the N-terminal domain of ScGS is known,<sup>[31]</sup> but full structural insights into the bifunctional enzyme have still not been obtained. Here we report on the

[a] H. Xu, Prof. Dr. J. S. Dickschat  
Kekulé Institute of Organic Chemistry and Biochemistry  
University of Bonn  
Gerhard-Domagk-Straße 1, 53121 Bonn (Germany)  
E-mail: dickschat@uni-bonn.de

Supporting information for this article is available on the WWW under <https://doi.org/10.1002/cbic.202300101>

© 2023 The Authors. ChemBioChem published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.



Scheme 1. Biosynthesis of geosmin (1).

functional characterisation of the geosmin synthase from *Allokutzneria albata* (AaGS) and investigations on the cyclisation mechanism by isotopic labelling experiments.

## Results and Discussion

According to the genome sequence information, geosmin synthase homologs are very widespread not only in the genus *Streptomyces*<sup>[32]</sup> but also in many other actinomycetes including the genera *Actinomadura*, *Amycolatopsis*, *Kitasatospora*, *Nocardia*, *Nocardioopsis* and others (Figure 2). A gene putatively coding for a geosmin synthase homolog from *Allokutzneria albata*, an organism that is known to encode terpene synthases for the diterpenes bonnadiene,<sup>[33]</sup> allokutznerene and phomopsene,<sup>[33]</sup> spiroalbatene,<sup>[34]</sup> and cembrene A,<sup>[34]</sup> was cloned into an expression vector and expressed in *Escherichia coli*. The purified recombinant enzyme (Figure S1) was incubated with farnesyl diphosphate (FPP), resulting in the identification of **1**, intermediates **6** and **9**, and the side product **10** by GC/MS analysis (Figures 3 and S2). The compounds **1**, **6** and **10** were isolated and their structures were confirmed by NMR spectroscopy (Tables S2–S4, Figures S3–S26). Based on the NOESY spectra (Figures S10, S18 and S26), full assignments for all diastereotopic hydrogens could be made and are reported here for the first time (Figures S3, S11 and S19; Tables S2–S4). The optical rotations confirmed the absolute configurations of **1** ( $[\alpha]_{\text{D}}^{25} = -14.0$ ,  $c$  0.2,  $\text{CH}_2\text{Cl}_2$ ; lit.<sup>[2]</sup>  $[\alpha]_{\text{D}}^{25} = -16.5$ , 0.5%,  $\text{CHCl}_3$ ), **6** ( $[\alpha]_{\text{D}}^{25} = -129.2$ ,  $c$  0.5,  $\text{CH}_2\text{Cl}_2$ , lit.<sup>[23]</sup>  $[\alpha]_{\text{D}}^{25} = -82$ ) and **10** ( $[\alpha]_{\text{D}}^{25} = -158.0$ ,  $c$  0.1,  $\text{CH}_2\text{Cl}_2$ , lit.<sup>[35]</sup>  $[\alpha]_{\text{D}} = -189$ ,  $c$  1.4,  $\text{CHCl}_3$ ).

The cyclisation mechanism to **1**, its intermediate **6** and the side product **10** was studied in detail through isotopic labelling

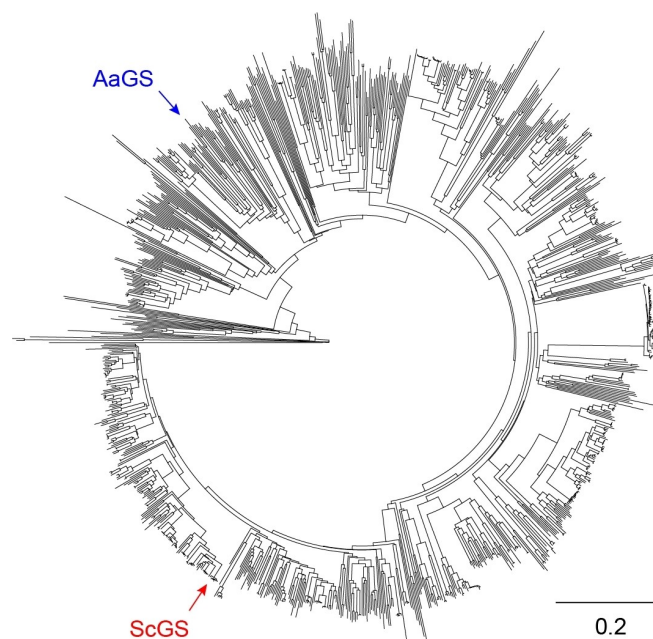


Figure 2. Phylogenetic tree constructed from the amino acid sequences of 1062 geosmin synthase homologs. The scale bar shows substitutions per site.

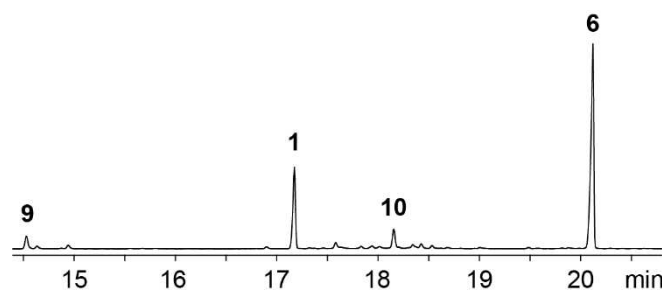


Figure 3. Total ion chromatogram of an extract of an enzyme incubation of FPP with AaGS.

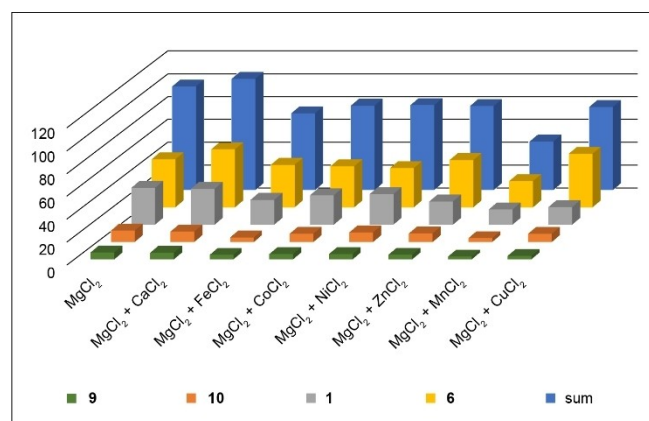
experiments (Table S5), with the aim to unravel the precise stereochemical course for every elementary step along the cyclisation cascade. The enzymatic conversion of (*R*)- and (*S*)-(1-<sup>2</sup>H)FPP<sup>[36]</sup> with AaGS and GC/MS analysis of the products revealed a specific loss of the 1-*pro-S* hydrogen in **6** and **1** and a specific migration of the same hydrogen into the *iPr* group of **10** (Figure S27), confirming earlier results with ScGS.<sup>[28]</sup> Furthermore, the usage of (*R*)- and (*S*)-(1-<sup>13</sup>C,1-<sup>2</sup>H)FPP<sup>[37]</sup> established that the 1-*pro-R* hydrogen remains bound to the original carbon in **1** and **6**, as indicated by a slightly upfield shifted triplet for C-6 resulting from <sup>13</sup>C-<sup>2</sup>H spin coupling (Figure S28). Moreover, HSQC analysis of the product mixture obtained from (*R*)-(1-<sup>13</sup>C,1-<sup>2</sup>H)FPP demonstrated the  $\alpha$  orientation of the deuterium atom in **1** (Figure S29). This is in line with a formation of **1** from **9** by reprotonation and a suprafacial 1,2-hydride shift, disfavoring a hypothetical role of **9** as a shunt product rather than intermediate. Interestingly, the <sup>13</sup>C NMR spectra for the enzymatic conversions of (12-<sup>13</sup>C)FPP and (13-<sup>13</sup>C)FPP (prepared

with ( $9\text{-}^{13}\text{C}$ )GPP, IPP and FPPS) with AaGS indicated that there was a minor exchange of  $^{13}\text{C}$  labellings between C-12 and C-13 in the biosynthesis towards **6**, while this phenomenon was not observed in the biosynthesis of the side product **10** which shares the same cationic intermediate **A** (Figure S30). This finding may be explained by a protonation induced opening of the three-membered ring in **8** to produce a cation at C-11 in **B** that is preferentially attacked from one side, but to a minor extent also from the other side. An alternative rotation of the *i*Pr group in **A** is less likely, because this should also lead to an exchange of labelling between C-12 and C-13 in **10**. The 1,2-hydride shift from **D** to **E** was also directly investigated using geranyl diphosphate (GPP) that was coupled with ( $1\text{-}^{13}\text{C}, 2, 2\text{-}^2\text{H}_2$ )IPP by *S. coelicolor* FPP synthase<sup>[38]</sup> to yield ( $1\text{-}^{13}\text{C}, 2\text{-}^2\text{H}$ )FPP. Its further conversion into labelled **1** resulted in a triplet signal for C-6, confirming the migration of deuterium to this labelled carbon (Figure S31).

The biosynthesis of **1** proceeds with a total number of three reprotonation steps: Intermediate **8** is reprotonated at C-4, **6** becomes reprotonated at C-1, and reprotonation of **9** happens at C-7. These reactions can be studied using a  $^{13}\text{C}$ -label in the substrate FPP at the reprotonated carbon in conjunction with incubation in a deuterium oxide buffer. Moreover, HSQC analysis of the labelled product **1** can show the stereochemical course of the reprotonation step. Enzymatic conversions of ( $3\text{-}^{13}\text{C}$ )FPP, ( $6\text{-}^{13}\text{C}$ )FPP and ( $10\text{-}^{13}\text{C}$ )FPP<sup>[39]</sup> with AaGS in deuterium oxide confirmed all three reprotonation steps and established deuterium incorporation into H-4, H-1 $\alpha$  and H-7 $\beta$  (Figures S32–S34).

Dionigi reported that the production of **1** in *Streptomyces albidoflavus* is increased in the presence of copper sulfate.<sup>[40]</sup> On the contrary, Schrader and Blevins found depleted levels of **1** in *Streptomyces halstedii* grown in the presence of divalent zinc, iron or copper.<sup>[41]</sup> These effects could be indirect or could be direct results of the presence of these divalent ions on the activity of geosmin synthase. To investigate the effect of different divalent cations, AaGS was incubated with FPP without any metal ions and with  $\text{MgCl}_2$ ,  $\text{CaCl}_2$ ,  $\text{FeCl}_2$ ,  $\text{NiCl}_2$ ,  $\text{ZnCl}_2$ ,  $\text{MnCl}_2$  and  $\text{CuCl}_2$ . These experiments revealed a strict dependency on  $\text{Mg}^{2+}$  that can only be substituted with  $\text{Mn}^{2+}$ , resulting in a relative activity of  $17.1 \pm 3.3\%$  in comparison to  $\text{Mg}^{2+}$  (set to 100% for the sum of the four main products **1**, **6**, **9**, and **10** of AaGS, Table S6). No activity was observed with any of the other tested salts or without addition of a divalent cation.

The effects of combinations of  $\text{MgCl}_2$  (5 mM) and additional divalent cations (1 mM) on the relative production of compounds **1**, **6**, **9** and **10** were also investigated (Figure 4 and Table S7, the sum of all four compounds was set to 100% for the incubation with  $\text{MgCl}_2$  only). To visualise specific changes for individual compounds their relative production is given in Figure S35. Addition of  $\text{CaCl}_2$  resulted in a slightly enhanced production of all compounds ( $107.5 \pm 3.3\%$ ) and especially of **6** ( $120.8 \pm 3.7\%$ ), suggesting that the activity of the N-terminal domain is more strongly increased than the activity of the C-terminal domain, which results in the accumulation of **6**. In contrast, the addition of  $\text{FeCl}_2$  gave an overall reduced activity ( $74.1 \pm 3.6\%$ ), with a particularly strong decrease of **10** ( $40.4 \pm$

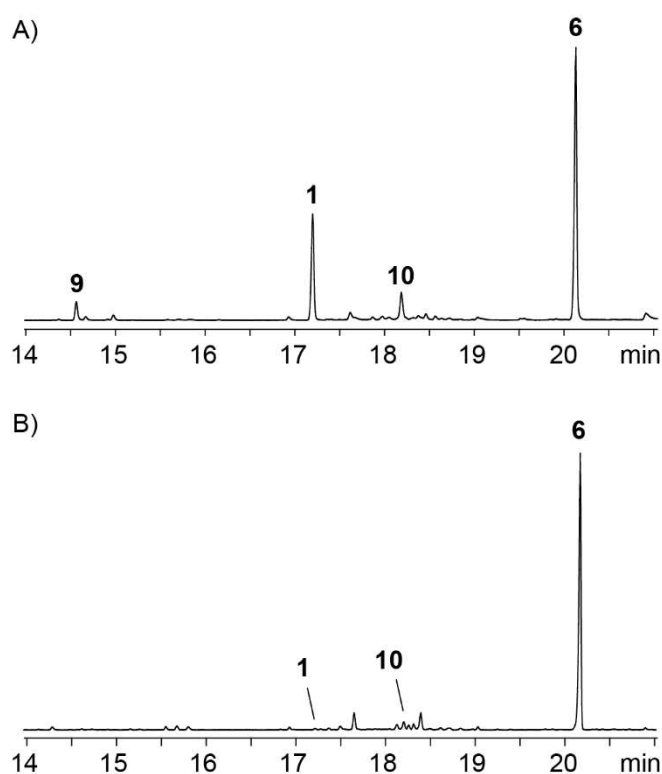


**Figure 4.** Relative activity of AaGS with different divalent cations. The sum of compounds **1**, **6**, **9** and **10** in the experiment with  $\text{MgCl}_2$  only was set to 100%. The bars show mean values from triplicates. For standard deviations cf. Table S7.

3.0%) and only a minor effect for **6** ( $88.4 \pm 4.8\%$ ). The presence of  $\text{Fe}^{2+}$  may shift the ratio for the reaction paths of the N-terminal domain towards the main product **6**. Administration of  $\text{CoCl}_2$  caused a slightly reduced activity ( $81.6 \pm 1.6\%$ ) with similar effects on all four compounds. Similar observations were made with  $\text{NiCl}_2$  ( $82.1 \pm 6.7\%$ ) and  $\text{ZnCl}_2$  ( $81.2 \pm 8.8\%$ ). The relative production with  $\text{MnCl}_2$  dropped to  $46.8 \pm 4.2\%$ . In first instance, this result seems unexpected, because this is the strongest decrease with any of the metal cations tested, although  $\text{Mn}^{2+}$  was the only divalent cation that could substitute for  $\text{Mg}^{2+}$ . An explanation for this finding may be that  $\text{Mn}^{2+}$  is besides  $\text{Mg}^{2+}$  the only divalent cation that can bind to the active site of AaGS, and the competing bindings between  $\text{Mn}^{2+}$  and  $\text{Mg}^{2+}$  may lead to a reduced activity. Finally, with  $\text{CuCl}_2$  a moderately reduced activity was found ( $80.3 \pm 6.0\%$ ). In particular, the production of **1** ( $47.8 \pm 4.8\%$ ) and **9** ( $50.8 \pm 3.1\%$ ) was suppressed, while the amounts of **6** were increased ( $111.4 \pm 7.4\%$ ), suggesting that  $\text{Cu}^{2+}$  specifically inhibits the C-terminal domain which consequently results in the accumulation of **6**.

A special feature of bifunctional enzymes or enzyme complexes are substrate tunnels that can be used to transport the product from the active site of one domain or subunit to the active site of the other domain or subunit. Substrate tunnels are e.g. known for the transport of indole between the  $\alpha$  and  $\beta$  subunits of tryptophan synthase,<sup>[42]</sup> or of ammonia in bifunctional enzymes in which one domain releases ammonia through hydrolysis of the amide function of Gln and another domain incorporates ammonia into a substrate molecule. Examples include glutamate synthase<sup>[43]</sup> or formylglycinamide ribonucleotide amidotransferase involved in the biosynthesis of purine nucleotides.<sup>[44]</sup> In contrast, cryo-EM studies on the bifunctional fusicoccadiene synthase containing a prenyltransferase domain for the biosynthesis of geranylgeranyl diphosphate (GGPP) and a terpene synthase domain for its cyclisation to fusicoccadiene revealed that in this enzyme substrate channeling does not proceed through a tunnel. Instead, substrate channeling is





**Figure 5.** Total ion chromatograms of extracts from incubations of FPP with AaGS A) without  $\beta$ -CD and B) with  $\beta$ -CD.

enabled by the proximity of multiple terpene synthase domains that surround an oligomeric core of prenyltransferase domains.<sup>[45]</sup> A similar situation was recently reported for farnesylfarnesyl diphosphate (FFPP) dependent bifunctional fungal triterpene synthases.<sup>[46]</sup> To investigate whether a substrate tunnel between the N- and the C-terminal domains of AaGS may be involved in channelling of **6** from one to the other active site, AaGS and FPP were incubated with or without the addition of  $\beta$ -cyclodextrin ( $\beta$ -CD). Terpenes can form host-guest complexes with  $\beta$ -CD,<sup>[47]</sup> and if **6** would be released into the medium by the N-terminal domain of AaGS, the addition of  $\beta$ -CD to the enzyme reaction could lead to a partial capturing of **6** that should consequently result in a lower production of **1**. This turned out to be the case: While the production of **6** and **10** upon addition of  $\beta$ -CD was similarly high as without  $\beta$ -CD, only traces of **1** and **9** were detected in the presence of  $\beta$ -CD (Figure 5).

## Conclusion

In summary, we have functionally characterised a new geosmin synthase from *Allokutzneria albata* (AaGS), and resolved the biosynthesis of geosmin (**1**), its intermediate (1(10)*E*,5*E*)-germacradien-11-ol (**6**) and its major side products **9** and **10** in all detail through isotopic labelling experiments. The results showed the sites and the stereochemical course of three reprotonation steps and the fate of the geminal methyl groups

of FPP that are slightly exchanged. Moreover, direct evidence for a 1,3-hydride migration and a late stage 1,2-hydride shift were obtained. The catalytic activity of AaGS was also shown to strictly depend on  $Mg^{2+}$  or  $Mn^{2+}$ , while other divalent cations cannot substitute for these metal ions. Reports in the literature mentioned an increased or decreased production of **1** during growth of geosmin producers in the presence of  $Cu^{2+}$ . These effects can have several reasons, but we showed here that  $Cu^{2+}$  has indeed a direct effect on AaGS and leads to decreased yields of **1**. Incubations of the bifunctional AaGS in the presence of  $\beta$ -CD, a molecule that can bind terpenes, resulted in a disrupted production of **1**, while **6** was not affected. The intermediate **6** is the product of the N-terminal domain that is passed on to the C-terminal domain for further conversion into **1**. The effect of  $\beta$ -CD suggests that **6** is not transferred from the N-terminal to the C-terminal domain through a tunnel, but rather through the medium where it becomes captured by the added  $\beta$ -CD. However, details of the interaction of the two domains of geosmin synthase remain unknown, also because still no structural data for the full bifunctional enzyme are available. This leaves room for additional future studies on this unique bacterial terpene synthase.

## Experimental Section

**Phylogenetic tree construction:** The phylogenetic tree of geosmin synthase homologs was constructed from 1062 amino acid sequences that were identified by BLAST search using the amino acid sequence of AaGS as a probe (accession number WP\_011030632). The tree was constructed using the tree builder function of Geneious (alignment type: global alignment with free end gaps, cost matrix: Blosum45, genetic distance model: Jukes-Cantor, tree build method: neighbor-joining, gap open penalty: 8, gap extension penalty: 2).

**Strains and culture conditions:** *Allokutzneria albata* NRRL B-24461 was obtained from the USDA-ARS Culture Collection (NRRL) (USA) and was cultivated in 65 GYM (4.0 g L<sup>-1</sup> glucose, 4.0 g L<sup>-1</sup> yeast extract, 10.0 g L<sup>-1</sup> malt extract, distilled water, pH 7.2) at 28 °C. For agar plate cultures 15.0 g agar-agar was added. *Saccharomyces cerevisiae* FY834 was cultivated in liquid YPAD medium (10.0 g yeast extract, 20.0 g peptone, 20.0 g glucose, 400 mg adenine sulphate, 1.0 L H<sub>2</sub>O) or on SM-URA agar plates (1.7 g yeast nitrogen base, 5.0 g ammonium sulphate, 20.0 g glucose, 770 mg nutritional supplement minus uracil, 20.0 g agar-agar, 1.0 L H<sub>2</sub>O) at 28 °C. *Escherichia coli* BL21 (DE3) was grown in LB medium (10.0 g tryptone, 5.0 g yeast extract, 5.0 g NaCl, 1.0 L H<sub>2</sub>O) at 37 °C. For agar plate cultures 16 g agar-agar was added. Kanamycin was used at a concentration of 50  $\mu$ g mL<sup>-1</sup>. All media were autoclaved at 121 °C for 20 min prior to use.

**Gene cloning and plasmid construction:** The desired gene was obtained from freshly isolated genomic DNA from *Allokutzneria albata* by PCR using Q5-High-Fidelity DNA polymerase (New England Biolabs, Ipswich, MA, USA) and the primers WP030430635\_Fw and WP030430635\_Rv (Table S1). PCR standard conditions were used (initial denaturation at 98 °C for 1 min, 35 cycles with denaturation at 98 °C for 15 sec, annealing at 62 °C for 30 sec and elongation at 72 °C for 68 seconds, final elongation step at 72 °C for 2 min). The PCR products together with the linearised pYE-Express shuttle vector were used for a yeast homologous recombination by a standard protocol using PEG, LiOAc and salmon sperm DNA.<sup>[48,49]</sup>

The transformed *Saccharomyces cerevisiae* cultures were plated on SM-URA plates and grown for 3 days at 28 °C. Colonies were collected from the plates and plasmid DNA was isolated using Zymoprep Yeast Plasmid Miniprep II (Zymo Research, Irvine, USA). The isolated plasmid DNA was used for electroporation of *E. coli* BL21 (DE3) electrocompetent cells. The transformed *E. coli* was grown overnight at 37 °C on LB agar plates containing kanamycin (50 µg mL<sup>-1</sup>). Single colonies were selected to inoculate LB medium with kanamycin (8 mL). The resulting cultures were grown for 24 h to isolate plasmid DNA, yielding plasmid pYE\_WP030430635 which was checked by analytical digest and by sequencing.

**Gene expression and enzyme purification:** For gene expression a preculture of the *E. coli* transformants harbouring the plasmid pYE\_WP030430635 was grown overnight at 37 °C in LB medium containing kanamycin. The expression cultures were then inoculated using 20 mL<sup>-1</sup> of preculture, followed by culturing at 37 °C with shaking until an OD<sub>600</sub> = 0.4–0.6 was reached. The cultures were cooled to 18 °C and protein expression was induced by addition of IPTG (400 mM in water, 1 mL<sup>-1</sup>, final concentration 0.4 mM). The expression cultures were shaken overnight at 18 °C and then centrifuged at 3.600 × g for 40 min. The supernatant was discarded and the cell pellet was resuspended in binding buffer (10 mL<sup>-1</sup> culture; 20 mM Na<sub>2</sub>HPO<sub>4</sub>, 500 mM NaCl, 20 mM imidazole, 1 mM MgCl<sub>2</sub>, pH = 7.4, 4 °C). The resulting suspension was subjected to ultra-sonication for cell lysis. The cell debris was removed by centrifugation (14.610 × g, 15 min) and the supernatant was loaded onto a Ni<sup>2+</sup>-NTA affinity chromatography column (Super Ni-NTA, Generson, Slough, UK). The column was washed with binding buffer (2 × 10 mL<sup>-1</sup> culture), and the desired His-tagged protein was eluted using elution buffer (10 mL<sup>-1</sup> culture, 20 mM Na<sub>2</sub>HPO<sub>4</sub>, 500 mM NaCl, 500 mM imidazole, 1 mM MgCl<sub>2</sub>, pH = 7.4, 4 °C). Fractions containing protein were pooled, analysed by Bradford assay to determine the protein concentration (0.6 mg mL<sup>-1</sup>)<sup>[50]</sup> and by SDS-PAGE (Figure S1), and used for analytical-scale incubation experiments.

**Incubation experiments and compound purification:** For a preparative scale incubation, FPP (80 mg) was dissolved in substrate buffer (50 mL), followed by addition of incubation buffer (125 mL, 50 mM Tris/HCl, 10 mM MgCl<sub>2</sub>, 20% glycerol, pH = 8.2). The reaction was started by addition of AaGS elution fraction (25 mL, ca. 2.0 mg mL<sup>-1</sup>) and incubated at 30 °C overnight. The reaction mixture was then extracted with Et<sub>2</sub>O (3 × 100 mL). The organic layers were dried with MgSO<sub>4</sub> and concentrated under reduced pressure. Column chromatography on silica gel with *n*-pentane/Et<sub>2</sub>O (1:0, 10:1 and 4:1) yielded compounds **10** (1.0 mg), **1** (2.0 mg), and **6** (5.0 mg) as colourless oils.

To investigate the metal ion dependency of AaGS, small scale incubations were performed with FPP (0.5 mg) in aqueous NH<sub>4</sub>HCO<sub>3</sub> solution (0.1 mL, 25 mM), AaGS elution fraction (0.4 mL, ca. 0.6 mg mL<sup>-1</sup>), and incubation buffer (0.5 mL, 50 mM Tris/HCl, 20% glycerol, pH = 8.2). Different salts were added (MgCl<sub>2</sub>, CaCl<sub>2</sub>, FeCl<sub>2</sub>, CoCl<sub>2</sub>, NiCl<sub>2</sub>, ZnCl<sub>2</sub>, MnCl<sub>2</sub> or CuCl<sub>2</sub>, 5 mM). After incubation with shaking at 30 °C overnight, the reaction mixtures were extracted with *n*-hexane (0.15 mL). The extracts were dried with MgSO<sub>4</sub> and analysed by GC-MS (Table S6).

To investigate the metal ion dependency in the presence of MgCl<sub>2</sub>, small scale incubations were performed with FPP (0.5 mg) in aqueous NH<sub>4</sub>HCO<sub>3</sub> solution (0.1 mL, 25 mM), AaGS elution fraction (0.4 mL, ca. 0.6 mg mL<sup>-1</sup>), incubation buffer (0.5 mL, 50 mM Tris/HCl, 20% glycerol, pH = 8.2), and MgCl<sub>2</sub> (5 mM). Other metal ions were added (CaCl<sub>2</sub>, FeCl<sub>2</sub>, CoCl<sub>2</sub>, NiCl<sub>2</sub>, ZnCl<sub>2</sub>, MnCl<sub>2</sub> or CuCl<sub>2</sub>, 1 mM). After incubation with shaking at 30 °C overnight, the reaction mixtures were extracted with *n*-hexane (0.15 mL). The extracts were dried with MgSO<sub>4</sub> and analysed by GC-MS (Table S7).

To investigate the effect of b-cyclodextrin, small scale incubations were performed with FPP (0.5 mg) in aqueous NH<sub>4</sub>HCO<sub>3</sub> solution (0.1 mL, 25 mM), AaGS elution fraction (0.4 mL, ca. 0.6 mg mL<sup>-1</sup>), incubation buffer (0.5 mL, 50 mM Tris/HCl, 20% glycerol, pH = 8.2) and MgCl<sub>2</sub> (5 mM). b-Cyclodextrin (10 mM) was added. After incubation with shaking at 30 °C overnight, the reaction mixtures were extracted with *n*-hexane (0.15 mL). The extracts were dried with MgSO<sub>4</sub> and analysed by GC-MS (Figure 5).

**GC/MS:** GC/MS analyses were performed on a 5977A GC/MSD system (Agilent, Santa Clara, CA, USA) with a 7890B GC and a 5977A mass selective detector. The GC was equipped with a HP5-MS fused silica capillary column (30 m, 0.25 mm i. d., 0.50 µm film). Specific GC settings were 1) inlet pressure: 77.1 kPa, He at 23.3 mL min<sup>-1</sup>, 2) injection volume: 2 µL, 3) temperature program: 5 min at 50 °C increasing at 5 °C min<sup>-1</sup> to 320 °C, 4) 60 s valve time, and 5) carrier gas: He at 1.2 mL min<sup>-1</sup>. MS settings were 1) source: 230 °C, 2) transfer line: 250 °C, 3) quadrupole: 150 °C and 4) electron energy: 70 eV. Retention indices (*I*) were determined from retention times in comparison to the retention times of *n*-alkanes (C<sub>7</sub>–C<sub>40</sub>).

**NMR spectroscopy:** NMR spectra were recorded on a Bruker (Billerica, MA, USA) Avance I (300 MHz), Avance I (400 MHz), Avance I (500 MHz), Avance III HD Prodigy (500 MHz) or an Avance III HD Cryo (700 MHz) NMR spectrometer. Spectra were measured in C<sub>6</sub>D<sub>6</sub> and referenced against solvent signals (<sup>1</sup>H-NMR, residual proton signal: δ = 7.16; <sup>13</sup>C-NMR: δ = 128.06).<sup>[51]</sup>

**IR spectroscopy and optical rotations:** IR spectra were recorded on an ALPHA II FTIR Spectrometer (Bruker Optics, MA, USA), and the scan range was set to 500 to 4000 cm<sup>-1</sup>. Optical rotations were recorded on a MCP 150 Modular Circular Polarimeter (Anton Paar GmbH, Graz, Austria).

Geosmin (**1**). TLC (pentane/Et<sub>2</sub>O = 10:1): R<sub>f</sub> = 0.37. Optical rotation: [α]<sub>D</sub><sup>25</sup> = -14.0 (c 0.2, CH<sub>2</sub>Cl<sub>2</sub>). GC (HP-5MS): *I* = 1425. MS (EI, 70 eV): Figure S2. IR (diamond ATR):  $\tilde{\nu}$ /cm<sup>-1</sup> = 2934 (s), 2864 (m), 1460 (w), 1447 (w), 1378 (w), 1178 (w), 948 (w). NMR data are given in Table S2 and Figures S3–S10.

(1(10)*E*,5*E*)-Germacradien-11-ol (**6**). TLC (pentane/Et<sub>2</sub>O = 4:1): R<sub>f</sub> = 0.19. Optical rotation: [α]<sub>D</sub><sup>20</sup> = -129.2 (c 0.5, CH<sub>2</sub>Cl<sub>2</sub>). GC (HP-5MS): *I* = 1661. MS (EI, 70 eV): Figure S2. IR (diamond ATR):  $\tilde{\nu}$ /cm<sup>-1</sup> = 3401 (m), 2967 (s), 2924 (s), 2850 (m), 1449 (w), 1382 (m), 1160 (w), 1112 (w), 985 (w), 867 (w). NMR data are given in Table S3 and Figures S11–S18.

Germacrene D (**10**). TLC (pentane): R<sub>f</sub> = 0.57. Optical rotation: [α]<sub>D</sub><sup>20</sup> = -158.0 (c 0.1, CH<sub>2</sub>Cl<sub>2</sub>). GC (HP-5MS): *I* = 1501. MS (EI, 70 eV): Figure S2. IR (diamond ATR):  $\tilde{\nu}$ /cm<sup>-1</sup> = 2955 (m), 2925 (s), 2854 (m), 1451 (w), 1259 (w), 1021 (w), 884 (w), 862 (w). NMR data are given in Table S4 and Figures S19–S26.

## Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft DFG (DI1536/7-2). We thank Patrick Rabe (Oxford) for gene cloning. Open Access funding enabled and organized by Projekt DEAL.

## Conflict of Interests

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:** biosynthesis · isotopes · odour · terpenes · volatiles

- [1] R. G. Buttery, J. A. Garibaldi, *J. Agric. Food Chem.* **1976**, *24*, 1246–1247.
- [2] N. N. Gerber, H. A. Lechevalier, *Appl. Microbiol.* **1965**, *13*, 935–938.
- [3] N. N. Gerber, *Tetrahedron Lett.* **1968**, *9*, 2971–2974.
- [4] a) N. N. Gerber, *Biotechnol. Bioeng.* **1967**, *9*, 321–327; b) C. P. Dionigi, D. F. Millie, A. M. Spanier, P. B. Johnsen, *J. Agric. Food Chem.* **1992**, *40*, 122–125; c) T. Rezanka, D. Libalova, J. Votruba, I. Viden, *Biotechnol. Lett.* **1993**, *16*, 75–78; d) C. E. G. Schoeller, H. Guertler, R. Pedersen, S. Molin, K. Wilkins, *J. Agric. Food Chem.* **2002**, *50*, 2615–2621; e) J. S. Dickschat, T. Martens, T. Brinkhoff, M. Simon, S. Schulz, *Chem. Biodiversity* **2005**, *2*, 837–865; f) K. Wilkins, C. Schöller, *Actinomycetologica* **2009**, *23*, 27–33; g) C. A. Citron, J. Gleitzmann, G. Laurenzano, R. Pukall, J. S. Dickschat, *ChemBioChem* **2012**, *13*, 202–214; h) P. Rabe, C. A. Citron, J. S. Dickschat, *ChemBioChem* **2013**, *14*, 2345–2354; i) C. A. Citron, L. Barra, J. Wink, J. S. Dickschat, *Org. Biomol. Chem.* **2015**, *13*, 2673–2683.
- [5] a) R. S. Safferman, A. A. Rosen, C. I. Mashni, M. E. Morris, *Environ. Sci. Technol.* **1967**, *1*, 25–26; b) T. Kikuchi, T. Mimura, K. Harimaya, H. Yano, T. Arimoto, Y. Masada, T. Inoue, *Chem. Pharm. Bull.* **1973**, *21*, 2342–2343; c) G. Izaguirre, C. J. Hwang, S. W. Krasner, M. J. McGuire, *Appl. Environ. Microbiol.* **1982**, *43*, 708–714; d) J. T. Wu, F. Jüttner, *Water Sci. Technol.* **1988**, *20*, 143–148; e) C. Höckelmann, P. G. Becher, S. H. von Reuss, F. Jüttner, *Z. Naturforsch.* **2009**, *64 C*, 49–55.
- [6] a) W. Trowitzsch, L. Witte, H. Reichenbach, *FEMS Microbiol. Lett.* **1981**, *12*, 257–260; b) S. Schulz, J. Fuhlendorff, H. Reichenbach, *Tetrahedron* **2004**, *60*, 3863–3872; c) J. S. Dickschat, S. C. Wenzel, H. B. Bode, R. Müller, S. Schulz, *ChemBioChem* **2004**, *5*, 778–787; d) J. S. Dickschat, H. Reichenbach, I. Wagner-Döbler, S. Schulz, *Eur. J. Org. Chem.* **2005**, 4141–4153; e) J. S. Dickschat, T. Nawrath, V. Thiel, B. Kunze, R. Müller, S. Schulz, *Angew. Chem. Int. Ed.* **2008**, *46*, 8287–8290.
- [7] a) T. Kikuchi, S. Kadota, H. Suehara, A. Nishi, K. Tsubaki, *Chem. Pharm. Bull.* **1981**, *29*, 1782–1784; b) J. P. Mattheis, R. G. Roberts, *Appl. Environ. Microbiol.* **1992**, *58*, 3170–3172; c) T. S. Boerjesson, U. M. Stoellman, J. L. Schnuerer, *J. Agric. Food Chem.* **1993**, *41*, 2104–2111; d) T. O. Larsen, J. C. Frisvad, *Mycol. Res.* **1995**, *99*, 1153–1166; e) S. La Guerche, C. Garcia, P. Darriet, D. Dubourdiou, J. Labarere, *Curr. Microbiol.* **2004**, *48*, 405–411; f) S. La Guerche, S. Chamont, D. Blancard, D. Dubourdiou, P. Darriet, *Antonie van Leeuwenhoek* **2005**, *88*, 131–139; g) K. Karlshoj, T. O. Larsen, *J. Agric. Food Chem.* **2005**, *53*, 708–715.
- [8] a) S. Breheret, T. Talou, S. Rapior, J.-M. Bessiere, *Mycologia* **1999**, *91*, 117–120; b) P. C. Schmidberger, P. Schieberle, *J. Agric. Food Chem.* **2017**, *65*, 9287–9296.
- [9] S. J. Hayes, K. P. Hayes, B. S. Robinson, *J. Protozool.* **1991**, *38*, 44–47.
- [10] J. Spoerle, H. Becker, N. S. Allen, M. P. Gupta, *Z. Naturforsch.* **1991**, *46 C*, 183–188.
- [11] Y. Saritas, M. M. Sonwa, H. Iznaguen, W. A. König, H. Muhle, R. Mues, *Phytochemistry* **2001**, *57*, 443–457.
- [12] J. Shi, G. Tong, Q. Yang, M. Huang, H. Ye, Y. Liu, J. Wu, J. Zhang, X. Sun, D. Zhao, *J. Agric. Food Chem.* **2021**, *69*, 11361–11371.
- [13] R. A. Flath, R. R. Forrey, J. O. John, B. G. Chan, *J. Agric. Food Chem.* **1978**, *26*, 1290–1293.
- [14] a) K. E. Murray, P. A. Bannister, R. G. Buttery, *Chem. Ind.* **1975**, 973–974; b) T. E. Acree, C. Y. Lee, R. M. Butts, J. Barnard, *J. Agric. Food Chem.* **1976**, *24*, 430–431.
- [15] P. E. Persson, *Water Res.* **1980**, *14*, 1113–1118.
- [16] R. T. Lovell, D. Broce, *Aquaculture* **1985**, *50*, 169–174.
- [17] T. C. Y. Hsieh, U. Tanchotikul, J. E. Matiella, *J. Food Sci.* **1988**, *53*, 1228–1229.
- [18] a) H. Omura, Y. Kuwahara, T. Tanabe, *J. Chem. Ecol.* **2002**, *28*, 2601–2612; b) Y. Kuwahara, Y. Ichiki, M. Morita, T. Tanabe, Y. Asano, *J. Chem. Ecol.* **2015**, *41*, 15–21.
- [19] R. Kaiser, C. Nussbaumer, *Helv. Chim. Acta* **1990**, *73*, 133–139.
- [20] A. Rieck, N. Bülow, W. A. König, *Phytochemistry* **1995**, *40*, 847–851.
- [21] L. Ding, C. Hertweck, *J. Nat. Prod.* **2020**, *83*, 2207–2211.
- [22] R. Bentley, R. Meganathan, *FEBS Lett.* **1981**, *125*, 220–222.
- [23] F. C. Pollak, R. G. Berger, *Appl. Environ. Microbiol.* **1996**, *62*, 1295–1299.
- [24] a) B. Gust, G. L. Challis, K. Fowler, T. Kieser, K. F. Chater, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 1541–1546; b) D. E. Cane, R. M. Watt, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 1547–1551.
- [25] D. Spittler, A. Jux, J. Piel, W. Boland, *Phytochemistry* **2002**, *61*, 827–834.
- [26] J. S. Dickschat, H. B. Bode, T. Mahmud, R. Müller, S. Schulz, *J. Org. Chem.* **2005**, *70*, 5174–5182.
- [27] J. Jiang, X. He, D. E. Cane, *Nat. Chem. Biol.* **2007**, *3*, 711–715.
- [28] a) X. He, D. E. Cane, *J. Am. Chem. Soc.* **2004**, *126*, 2678–2679; b) J. Jiang, X. He, D. E. Cane, *J. Am. Chem. Soc.* **2006**, *128*, 8128–8129.
- [29] T. Nawrath, J. S. Dickschat, R. Müller, J. Jiang, D. E. Cane, S. Schulz, *J. Am. Chem. Soc.* **2008**, *130*, 430–431.
- [30] J. Jiang, D. E. Cane, *J. Am. Chem. Soc.* **2008**, *130*, 428–429.
- [31] G. G. Harris, P. M. Lombardi, T. A. Pemberton, T. Matsui, T. M. Weiss, K. E. Cole, M. Köksal, F. V. Murphy, L. S. Vedula, W. K. W. Chou, D. E. Cane, D. W. Christianson, *Biochemistry* **2015**, *54*, 7142–7155.
- [32] L. Martin-Sanchez, K. S. Singh, M. Avalos, G. P. van Wezel, J. S. Dickschat, P. Garbeva, *Beilstein J. Org. Chem.* **2019**, *15*, 1181–1193.
- [33] L. Lauterbach, J. Rinkel, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2018**, *57*, 8280–8283; *Angew. Chem.* **2018**, *130*, 8412–8415.
- [34] J. Rinkel, L. Lauterbach, P. Rabe, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2018**, *57*, 3238–3241; *Angew. Chem.* **2018**, *130*, 3292–3296.
- [35] A. R. Hayman, R. T. Weavers, *Phytochemistry* **1990**, *29*, 3157–3162.
- [36] J. Rinkel, P. Rabe, P. Garbeva, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2016**, *55*, 13593–13596; *Angew. Chem.* **2016**, *128*, 13791–13794.
- [37] P. Rabe, J. Rinkel, E. Dolja, T. Schmitz, B. Nubbemeyer, T. H. Luu, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2017**, *56*, 2776–2779; *Angew. Chem.* **2017**, *129*, 2820–2823.
- [38] P. Rabe, J. Rinkel, B. Nubbemeyer, T. G. Köllner, F. Chen, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2016**, *55*, 15420–15423; *Angew. Chem.* **2016**, *128*, 15646–15649.
- [39] P. Rabe, L. Barra, J. Rinkel, R. Riclea, C. A. Citron, T. A. Klapschinski, A. Janusko, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2015**, *54*, 13448–13451; *Angew. Chem.* **2015**, *127*, 13649–13653.
- [40] C. P. Dionigi, *Water Sci. Technol.* **1995**, *31*, 135–138.
- [41] K. K. Schrader, W. T. Blevins, *J. Ind. Microbiol. Biotechnol.* **2001**, *26*, 241–247.
- [42] C. C. Hyde, S. A. Ahmed, E. A. Padlan, E. W. Miles, D. R. Davies, *J. Biol. Chem.* **1988**, *263*, 17857–17871.
- [43] C. Binda, R. T. Bossi, S. Wakatsuki, S. Arzt, A. Coda, B. Curti, M. A. Vanoni, A. Mattevi, *Structure* **2000**, *8*, 1299–1308.
- [44] R. Anand, A. A. Hoskins, J. Stubbe, S. E. Ealick, *Biochemistry* **2004**, *43*, 10328–10342.
- [45] J. L. Faylo, T. van Eeuwen, H. J. Kim, J. J. Gorbea Colon, B. A. Garcia, K. Murakami, D. W. Christianson, *Nat. Commun.* **2021**, *12*, 3487.
- [46] H. Tao, L. Lauterbach, G. Bian, R. Chen, A. Hou, T. Mori, S. Cheng, B. Hu, L. Lu, X. Mu, M. Li, N. Adachi, M. Kawasaki, T. Moriya, T. Senda, X. Wang, Z. Deng, I. Abe, J. S. Dickschat, T. Liu, *Nature* **2022**, *606*, 414–419.
- [47] N. Ajsjaka, K. Hara, K. Mikuni, K. Hara, H. Hashimoto, *Biosci. Biotechnol. Biochem.* **2000**, *64*, 731–734.
- [48] J. S. Dickschat, K. A. K. Pahirulzaman, P. Rabe, T. A. Klapschinski, *ChemBioChem* **2014**, *15*, 810–814.
- [49] R. D. Giets, R. H. Schiestl, *Nat. Protoc.* **2007**, *2*, 31–34.
- [50] M. M. Bradford, *Anal. Biochem.* **1976**, *72*, 248–254.
- [51] G. R. Fulmer, A. J. M. Miller, N. H. Sherden, H. E. Gottlieb, A. Nudelman, B. M. Stoltz, J. E. Bercaw, K. I. Goldberg, *Organometallics* **2010**, *29*, 2176–2179.

Manuscript received: February 7, 2023

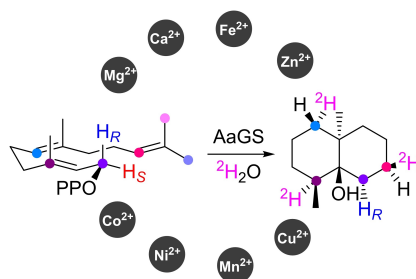
Revised manuscript received: February 21, 2023

Accepted manuscript online: February 21, 2023

Version of record online: ■ ■ ■

# RESEARCH ARTICLE

A new **geosmin synthase** from *Allokutzneria albata* was characterised for its detailed reaction mechanism through isotopic labelling experiments and its metal dependency. Incubations in the presence of  $\beta$ -cyclodextrin gave insights into the mode of interaction of the N-terminal and C-terminal domains of this bifunctional enzyme.



H. Xu, Prof. Dr. J. S. Dickschat\*

1 – 7

**A Detailed View on Geosmin Biosynthesis**



## Appendix O

### **Revision of the Cyclisation Mechanism for the Diterpene Spiroviolene and Investigations on its Mass Spectrometric Fragmentation**

*ChemBioChem* **2021**, 22, 850

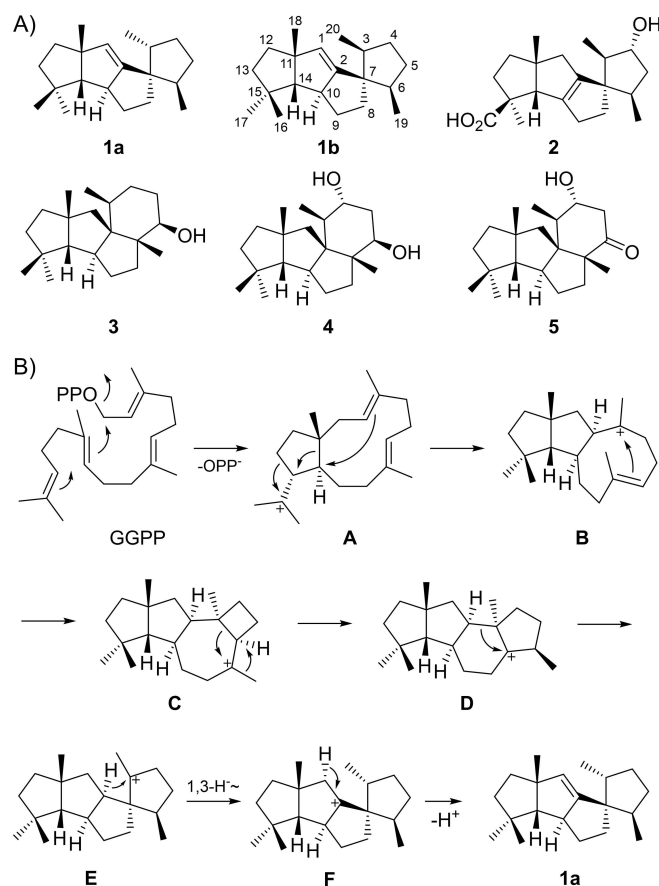
DOI: 10.1002/cbic.202000682

# Revision of the Cyclisation Mechanism for the Diterpene Spiroviolene and Investigations of Its Mass Spectrometric Fragmentation

Houchao Xu<sup>[a]</sup> and Jeroen S. Dickschat<sup>\*[a]</sup>

The diterpene spiroviolene, its diterpene synthase from *Streptomyces violens* and the experimentally determined terpene cyclisation mechanism were reported in 2017. Recently, the structure of spiroviolene was revised based on a total synthesis, with consequences for the cyclisation mechanism. Herein, a reinvestigation of the terpene cyclisation to spiroviolene and the mass spectrometric fragmentation mechanism investigated by <sup>13</sup>C-labelling experiments are presented.

Diterpenes are made by diterpene synthases (DTSs) that convert the acyclic and achiral precursor geranylgeranyl diphosphate (GGPP) in remarkable one-step enzymatic transformations into usually enantiomerically pure, often polycyclic, sometimes even cage-like molecules with multiple stereogenic centres. Several type I DTSs were recently discovered from bacteria which are monofunctional enzymes,<sup>[1]</sup> while fungal DTSs are usually bifunctional and exhibit a prenyltransferase (GGPP synthase, GGPPS) and a DTS domain.<sup>[2,3]</sup> Spiroviolene is a spirocyclic triquinane diterpene from *Streptomyces violens* for which we had originally reported the structure of **1a** (Scheme 1A). The compound is made by a diterpene synthase that has been deeply studied for its cyclisation mechanism through the use of several isotopically labelled substrates.<sup>[4]</sup> In particular, a double labelling experiment with (3-<sup>13</sup>C,2-<sup>2</sup>H)GGPP resulted in an upfield shifted triplet for C3 in the <sup>13</sup>C NMR spectrum, indicating a direct <sup>13</sup>C-<sup>2</sup>H bond in labelled **1** obtained with spiroviolene synthase (SvS) from this substrate. This result was interpreted by the cyclisation mechanism shown in Scheme 1B in which the cationic intermediate **E** reacts by a 1,3-hydride migration to **F**, establishing the stereogenic centre at C3. Notably, the structure of **1a** differs with respect to the configuration of the stereogenic centre at this carbon from those of structurally similar molecules, including the fungal cyclopiane-type diterpenes conidiogenol (**4**) and conidiogenone (**5**) from *Penicillium cyclopium*,<sup>[5]</sup> several derivatives from other *Penicillium* spp.,<sup>[6,7]</sup>



**Scheme 1.** A) Originally reported (**1a**) and revised (**1b**) structures of spiroviolene from *S. violens* and related compounds from *Penicillium*. B) Originally proposed biosynthetic hypothesis for the cyclisation of GGPP to **1a**.

and spirograterpene **A** (**2**) from *Penicillium granulatum* that has the same skeleton as **1**.<sup>[8]</sup> Also the bifunctional cyclopiane-type diterpene synthase from *Penicillium chrysogenum* has been reported that is responsible for the biosynthesis of **3** as the proposed precursor to **4**, **5** and other cyclopiane type diterpenes. In this study, the structure of compound **3** was established by X-ray analysis.<sup>[9]</sup> Recently, the structure of the bacterial compound **1a** was revised by Snyder and co-workers to that of **1b** based on a total synthesis of both stereoisomers, showing that the stereogenic centre at C3 of **1** has the same configuration as for the fungal compounds.<sup>[10]</sup> This finding has consequences on the cyclisation mechanism, because the proposed 1,3-hydride shift from **E** to **F** can only proceed with the facial selectivity to explain

[a] H. Xu, Prof. Dr. J. S. Dickschat  
Kekulé Institute for Organic Chemistry and Biochemistry  
University of Bonn  
Gerhard-Domagk-Straße 1, 53121 Bonn (Germany)  
E-mail: dickschat@uni-bonn.de

Supporting information and the ORCID identification numbers for the authors of this article can be found under <https://dx.doi.org/10.1002/cbic.202000682>.

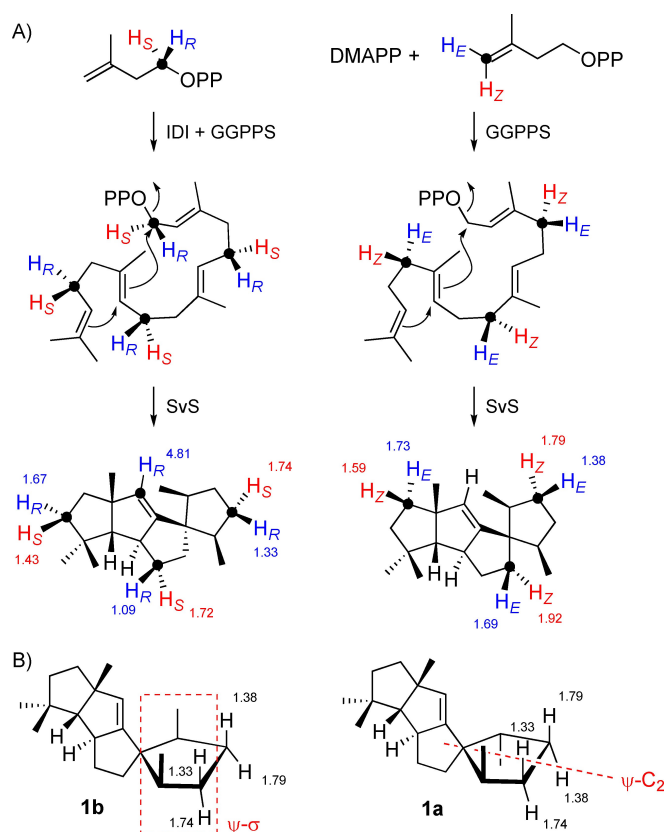
© 2020 The Authors. ChemBioChem published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

the formation of **1a**, but cannot explain the revised stereochemistry at C3 in **1b**. Herein, we report additional labelling experiments that further support the revised structure of **1b** for spiroviolene, a new biosynthetic hypothesis that is in line with all previous labelling experiments and explains the formation of **1b**, and the EI-MS fragmentation mechanism of spiroviolene based on  $^{13}\text{C}$ -labelling of each individual carbon.

During the original structure elucidation of **1**, the assignment of the relative configuration for its western half by interpretation of the NOESY spectrum proved to be clear, while the assignment for the relative configuration at C3 turned out to be difficult. This was also the case for the NOESY based assignment of the diastereotopic hydrogens  $\text{H}_\alpha$  and  $\text{H}_\beta$  at some of the methylene carbons that was undoubtedly established for C5 and C9 by enantioselective deuterations using the substrates (*R*)- and (*S*)-( $1\text{-}^{13}\text{C},1\text{-}^2\text{H}$ )farnesyl diphosphate (FPP) as well as (*R*)- and (*S*)-( $1\text{-}^{13}\text{C},1\text{-}^2\text{H}$ )geranyl diphosphate (GPP) in conjunction with isopentenyl diphosphate (IPP) and GGPPS from *Streptomyces cyaneofuscatus*.<sup>[4]</sup> Meanwhile, additional stereoselectively deuterated substrates have been made available by synthesis in our laboratory, including (*R*)- and (*S*)-( $1\text{-}^{13}\text{C},1\text{-}^2\text{H}$ )IPP that can be used with isopentenyl diphosphate isomerase (IDI) and GGPPS to introduce enantioselective deuterations at C1, C5, C9, and C13 of GGPP.<sup>[11]</sup> Furthermore, (*E*)- and (*Z*)-( $4\text{-}^{13}\text{C},4\text{-}^2\text{H}$ )IPP together with dimethylallyl diphosphate (DMAPP) and GGPPS give rise to enantioselectively deuterated GGPP at C4, C8 and C12.<sup>[12]</sup> The additional  $^{13}\text{C}$ -labelling at the deuterated carbons allow for a sensitive detection of the connected hydrogens by HSQC spectroscopy, while the signal for the hydrogen replaced by deuterium is extinguished. Data interpretation in these experiments is based on the fundamental work by Cornforth and co-workers on the stereochemical course of the prenyltransferase reaction.<sup>[13]</sup>

The enzymatic conversions of these probes with SvS and HSQC analysis of the obtained products (Table S1 and Figures S1–S4 in the Supporting Information) resulted in the assignments for the diastereotopic hydrogens as summarised in Scheme 2A, showing that the original assignments for the hydrogens at C4 and C12 require revision (Table 1). While the erroneous assignments for  $\text{H}_{4\alpha}$  and  $\text{H}_{4\beta}$  fitted better to the structure of **1a** with a pseudo- $\text{C}_2$  axis ( $\psi\text{-C}_2$ ) for the eastern cyclopentane ring (Scheme 2B), the corrections for  $\text{H}_{4\alpha}$  and  $\text{H}_{4\beta}$  are in line with Snyder's revised structure of **1b** with a pseudo-symmetry plane ( $\psi\text{-}\sigma$ ). This also demonstrates how useful the stereoselectively deuterated precursor probes are for the structure elucidation of terpenes.

After the structural revision of **1**, a modified biosynthetic proposal is required that can now be unified to a general biosynthetic hypothesis for bacterial spiroviolene and the fungal cyclopentane-type diterpenes (Scheme 3). Following a 1,11–10,14 cyclisation of GGPP to **A'** (=A in Scheme 1), the next two steps represent a modification of our earlier proposal and are similar to the initial steps in variediene biosynthesis suggested by Hong and Tantillo based on DFT calculations.<sup>[14]</sup> This includes expansion of the cyclopentane ring, followed by a transformation that was described as "highly asynchronous ring-opening/ring-closing process that accomplishes the same net result as a 1,2-alkyl shift"<sup>[14]</sup> from C10 to C14, as indicated in **B'**. This reaction, together with a



**Scheme 2.** A) Enantioselective deuteration of **1b** and resulting assignments of the NMR data for hydrogens at methylene groups. Data at hydrogens indicate  $^1\text{H}$  chemical shifts in ppm. B) Pseudo-symmetry plane ( $\psi\text{-}\sigma$ ) in **1b** with corrected NMR assignments for  $\text{H}_{4\alpha}$  and  $\text{H}_{4\beta}$ , and pseudo- $\text{C}_2$  axis ( $\psi\text{-C}_2$ ) in the eastern cyclopentane ring of **1a** with original erroneous assignments.

2,10-cyclisation, leads to **C'** (=B in Scheme 1, note that the stereochemistry of **C'** at C2 is different to the corresponding intermediate proposed for variediene biosynthesis<sup>[14]</sup>). Cation **C'** then reacts by a 1,2-hydride shift from C2 to C3 to yield **D'**, which substitutes for the 1,3-hydride migration that we had established experimentally in our previous study,<sup>[4]</sup> but is now reinterpreted to explain the corrected stereochemistry at C3 of **1**.<sup>[10]</sup> The following 2,7-cyclisation then leads to the secondary cation **E'** that can be trapped by water to yield **3**,<sup>[9]</sup> a compound that may be oxidised via **4** to **5**,<sup>[5]</sup> for example, by cytochrome P450 oxygenases. The configuration of the stereocentre at the carbinol carbon in **3** is in line with a concerted **D'**-to-**3** conversion with *anti* addition to the double bond between C6 and C7 in **D'**. Alternatively, in the absence of water **E'** may transiently react through the nonclassical cation **F'** to **G'** with Me19 now being shifted from C7 to C6, which represents a skeletal rearrangement that was established experimentally in our previous study by  $^{13}\text{C}$ -labelling of each individual carbon and substitutes for the **C**-to-**D** transformation in Scheme 1.<sup>[4]</sup> The spirocentre is installed by ring contraction of **G'** to **H'**, substituting for the ring contraction from **D** to **E** in Scheme 1. Cation **H'** can react by alternative deprotonations to spiroviolene (**1**) or the hypothetical natural product **6** that likely serves as the precursor to spirograterpene **A** (**2**).<sup>[8]</sup> Notably, in the final deprotonations to **1** and **6** the proton is abstracted from the

Table 1. Revised NMR data for spiroviolene (1) in C <sub>6</sub> D <sub>6</sub> (700 MHz).			
C <sup>[a]</sup>	type	<sup>13</sup> C <sup>[b]</sup>	<sup>1</sup> H <sup>[b]</sup>
1	CH	128.9	4.81 (d, <i>J</i> = 2.9)
2	C <sub>q</sub>	148.9	–
3	CH	44.7	1.60 (m)
4	CH <sub>2</sub>	31.3	1.79 (m, H <sub>α</sub> )* 1.38 (m, H <sub>β</sub> )*
5	CH <sub>2</sub>	30.7	1.74 (m, H <sub>α</sub> ) 1.33 (m, H <sub>β</sub> )
6	CH	46.6	1.81 (m)
7	C <sub>q</sub>	53.8	–
8	CH <sub>2</sub>	39.5	1.92 (ddd, <i>J</i> = 12.7, 6.9, 6.9, H <sub>α</sub> ) 1.69 (m, H <sub>β</sub> )
9	CH <sub>2</sub>	33.1	1.72 (m, H <sub>α</sub> ) 1.09 (dddd, <i>J</i> = 12.2, 12.2, 11.3, 7.6, H <sub>β</sub> )
10	CH	59.4	2.77 (dddd, <i>J</i> = 12.5, 6.4, 6.4, 2.9)
11	C <sub>q</sub>	63.7	–
12	CH <sub>2</sub>	38.6	1.73 (m, H <sub>α</sub> )* 1.59 (m, H <sub>β</sub> )*
13	CH <sub>2</sub>	40.8	1.67 (m, H <sub>α</sub> ) 1.43 (dddd, <i>J</i> = 11.8, 6.6, 1.5, 1.5, H <sub>β</sub> )
14	CH	66.0	1.58 (m)
15	C <sub>q</sub>	41.3	–
16	CH <sub>3</sub>	29.1	1.04 (s)
17	CH <sub>3</sub>	26.1	1.03 (s)
18	CH <sub>3</sub>	32.4	1.34 (s)
19	CH <sub>3</sub>	15.2	0.97 (d, <i>J</i> = 6.7)
20	CH <sub>3</sub>	15.1	0.94 (d, <i>J</i> = 6.7)

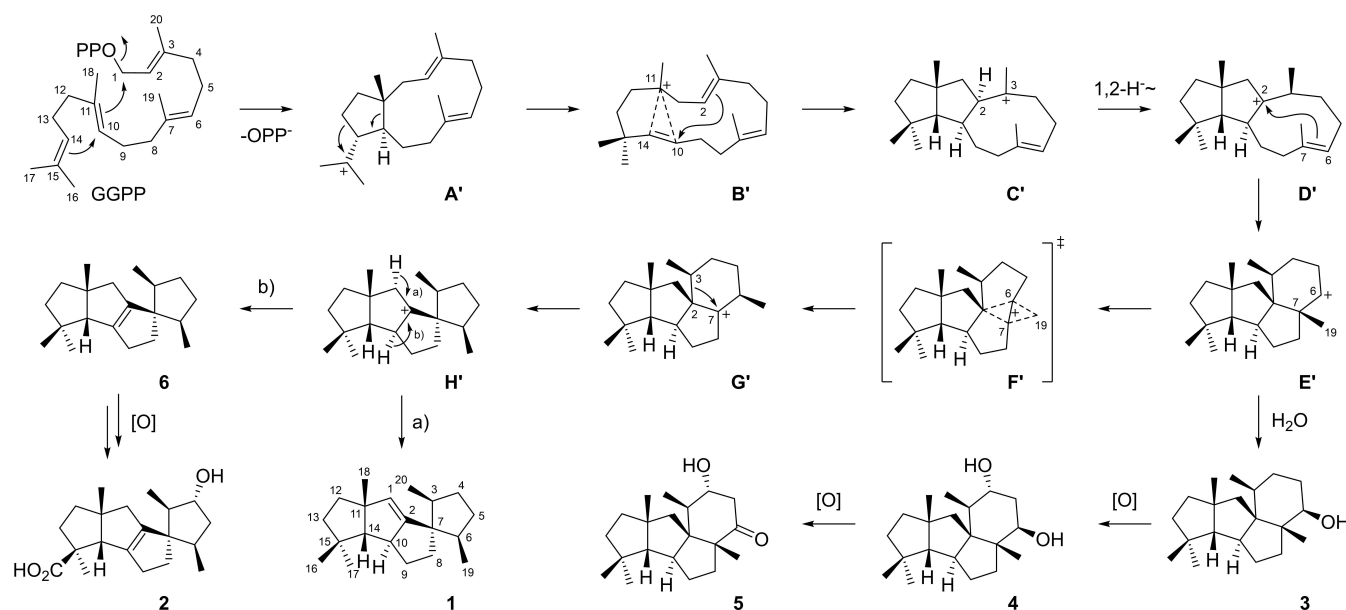
[a] Carbon numbering as shown in Scheme 1. [b] Chemical shifts δ in ppm, multiplicity (s = singlet, d = doublet, m = multiplet), and coupling constants *J* in Hz.

same hemisphere of H', for which the selectivity was supported by stereoselective deuteration at C1 for 1.<sup>[4]</sup>

Besides for biosynthetic investigations, isotopically labelled compounds are also very useful to study mass spectrometric fragmentation mechanisms (for important mass spectrometric fragmentation reactions such as  $\sigma$ -bond cleavage,  $\alpha$ -fragmenta-

tion, inductive cleavage, McLafferty rearrangement and retro-Diels-Alder fragmentation, cf. ref. [15]). The required isotopic labelling can be introduced into well-defined positions by feeding of correspondingly labelled precursors to cultures of the producing organisms, but if the incorporation rates are low, mixtures of different isotopomers and/or of the labelled and the unlabelled compound will be obtained, which can significantly hamper data interpretation. Synthetic or semisynthetic approaches can give access to compounds with well-defined isotopic substitutions, but are very laborious and might require different synthetic strategies for each target position of the natural product. Deuterium has often been used successfully to study EI-MS fragmentation mechanisms of terpenes,<sup>[16–20]</sup> but sometimes gave unclear results as a consequence of unspecific scrambling.<sup>[21]</sup> In previous studies on EI-MS fragmentation mechanisms we have enzymatically prepared <sup>13</sup>C-labelled terpenes from the corresponding synthetic <sup>13</sup>C-labelled terpene precursors, which allowed to label each individual carbon.<sup>[22–26]</sup> Using the same approach all 20 isotopomers of (<sup>13</sup>C)-1 were enzymatically prepared from labelled terpene precursors with SvS (Table S1). Based on a comparison of their mass spectra (Figure S5) to the mass spectrum of unlabelled 1 (Figure 1), a position-specific mass shift analysis (PMA<sub>*m/z*</sub>) indicates for a studied fragment ion (*m/z*) which carbons contribute to its formation (Figure 2).

The PMA for *m/z* 243 ([*M*–C<sub>2</sub>H<sub>5</sub>]<sup>+</sup>) reveals the specific formation of this fragment ion by cleavage of C12 and C13. After electron impact ionisation of 1 to the radical cation 1<sup>•+</sup>, this is explainable by a sequence of  $\alpha$ -cleavage to a<sup>•+</sup>, hydrogen rearrangement to the conjugated butadienyl cation b<sup>•+</sup>, and another  $\alpha$ -fragmentation to c<sup>+</sup> (Scheme 4A). The PMAs for *m/z* 230 ([*M*–C<sub>3</sub>H<sub>6</sub>]<sup>+</sup>) and 229 ([*M*–C<sub>3</sub>H<sub>7</sub>]<sup>+</sup>) demonstrate their formation from the same portion of 1, with partial loss of three of the carbons from the C20–C3–C4–C5–C6–C19 unit. Most likely, the lost carbons extrude as an intact portion C20–C3–C4 or C5–C6–C19.



Scheme 3. Revised cyclisation mechanism from GGPP to 1 and biosynthetic links to related fungal compounds.



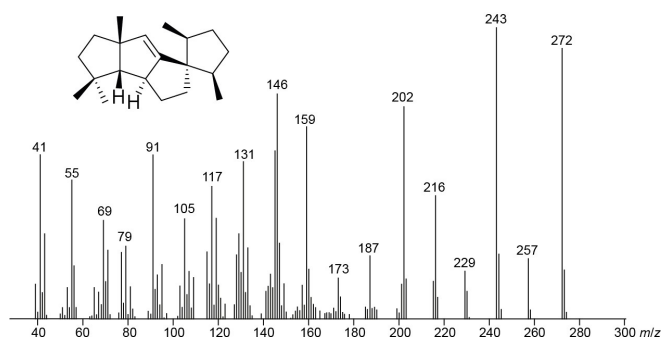


Figure 1. EI mass spectrum of 1.

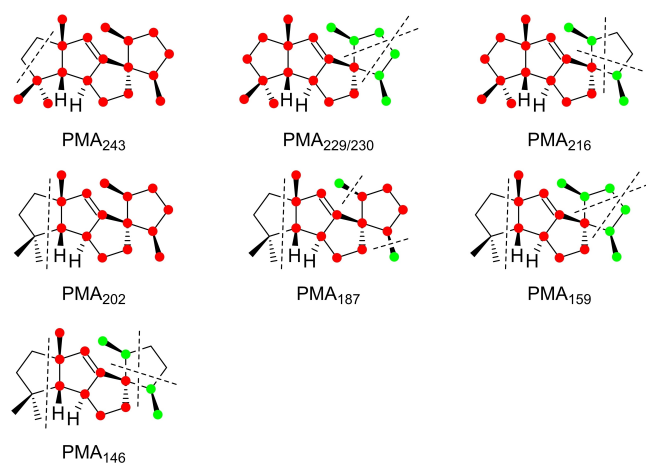
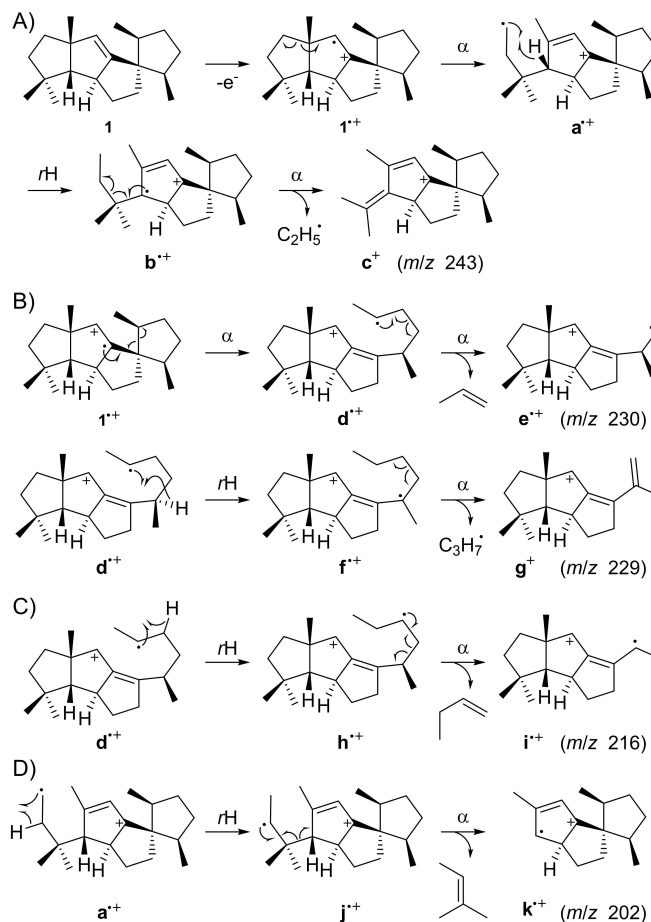


Figure 2. Position-specific mass-shift analysis (PMA<sub>*m/z*</sub>) for main EI fragment ions *m/z* of 1. Red dots indicate carbons that contribute fully, green dots indicate carbons that contribute partially to the formation of a fragment ion. Dotted lines indicate relevant carbon-carbon bond cleavages.

Starting from  $1^{+\bullet}$  two sequential  $\alpha$ -cleavages can proceed via  $d^{+\bullet}$  to  $e^{+\bullet}$ , or by alternative hydrogen rearrangement of  $d^{+\bullet}$  to  $f^{+\bullet}$  and  $\alpha$ -fragmentation to the pentadienyl cation  $g^+$ . Both reactions are shown in Scheme 4B for the loss of the C20-C3-C4 unit, while analogous reactions can explain cleavage of carbons C5-C6-C19. The fragmentation mechanism towards *m/z* 216 ( $[M-C_4H_8]^+$ ) with cleavage of C4-C5 and either C20-C3 or C6-C19 can similarly be understood by loss of an intact unit C20-C3-C4-C5 or C4-C5-C6-C19. Scheme 4C shows possible fragmentation reactions for the first case starting from  $d^{+\bullet}$  by hydrogen rearrangement to  $h^{+\bullet}$  and  $\alpha$ -fragmentation with loss of butene to  $i^+$ . The PMA for *m/z* 202 ( $[M-C_5H_{10}]^+$ ) indicates the formation of this fragment ion by specific cleavage of C12-C13-C15(-C17)-C16. This finding can be explained by hydrogen rearrangement from  $a^{+\bullet}$  to  $j^{+\bullet}$  and  $\alpha$ -fragmentation with fragmentation of 2-methylbut-2-ene to  $k^+$  (Scheme 4D).

The PMAs for *m/z* 187, 159 and 146 (Figure 2) reveal that these fragments arise by a two-step process with loss of the C12-C13-C15(-C17)-C16 portion in all cases, in addition to cleavage of a methyl group (C19 or C20), or of a  $C_3H_7$  or a  $C_4H_8$  fragment as in the formation of  $g^+$  or  $h^{+\bullet}$ , respectively. Starting from  $k^+$ , an  $\alpha$ -cleavage to  $l^+$ , hydrogen rearrangement to  $m^{+\bullet}$ , and another  $\alpha$ -

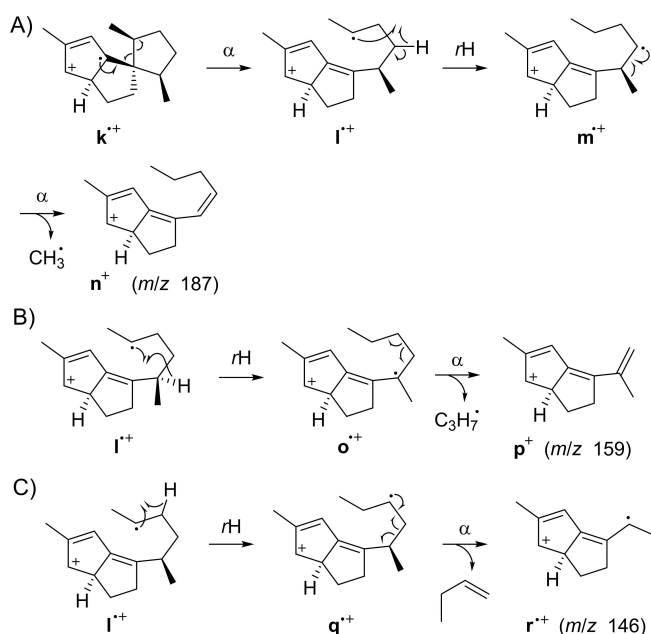


Scheme 4. Fragmentation mechanisms for fragment ions *m/z* 243, 230, 229, 216 and 202 of 1.  $\alpha$ :  $\alpha$ -cleavage, rH: hydrogen rearrangement.

fragmentation yield  $n^+$  with a conjugated heptatrienyl cation system to explain *m/z* 187 ( $[M-C_5H_{10}-CH_3]^+$ ) (Scheme 5A, shown for the loss of C19, the loss of C20 can proceed analogously). A similar sequence through  $l^+$ , hydrogen rearrangement to  $o^{+\bullet}$ , and  $\alpha$ -fragmentation gives  $p^+$  to explain *m/z* 159, again hypothetically best represented by a conjugated heptatrienyl cation ( $[M-C_5H_{10}-C_3H_7]^+$ , Scheme 5B). The fragment ion at *m/z* 146 ( $[M-C_5H_{10}-C_4H_8]^+$ ) can be formed from  $l^+$  by hydrogen rearrangement to  $q^{+\bullet}$  and subsequent  $\alpha$ -cleavage to the hexatrienyl cation  $r^+$  (Scheme 5C).

## Conclusion

The structural revision of spiroviolene (1) as recently promoted based on a total synthesis by Snyder and co-workers<sup>[10]</sup> has prompted us to reinvestigate the complete NMR data assignment through the use of stereoselectively deuterated precursors. As discussed here, the now rigorously assigned data are in line with and further support the structural revision of 1. Moreover, a revised biosynthetic hypothesis for the terpene cyclisation of GGPP to 1 has been developed. The previously suggested biosynthetic hypothesis for the initially assigned structure of 1 was



**Scheme 5.** Fragmentation mechanisms for fragment ions  $m/z$  187, 159 and 146 of **1**.  $\alpha$ :  $\alpha$ -cleavage,  $rH$ : hydrogen rearrangement.

fully supported by extensive isotopic labelling experiments, as is the revised mechanism for the corrected structure of **1**, demonstrating that enzyme mechanistic models, like reaction mechanisms for any chemical reaction, can only be supported by experimental data, while absolute proof for chemical mechanistic models is in principle impossible to obtain. The revised structure of **1** now allows for a unified hypothesis of a biosynthetic mechanism towards this compound and several structurally related diterpenes from fungi. The DTC domain of the cyclopentane-type diterpene synthase from *P. chrysogenum*<sup>[9]</sup> has only 15% amino acid sequence identity to SvS,<sup>[4]</sup> demonstrating that similar functions have evolved independently in fungi and bacteria. A similar finding was made previously for the fungal and bacterial diterpene synthases for phomopsene,<sup>[27,28]</sup> while the fungal and bacterial sesquiterpene synthases for corvol ethers have a common evolutionary origin, suggesting cross-kingdom horizontal gene transfer.<sup>[29,30]</sup> Besides biosynthetic and enzyme mechanistic investigations, isotopically labelled terpene precursors can give valuable insights into EI-MS fragmentation mechanisms of terpenes, as the enzymatic access of specifically labelled terpenes from these precursors, after their individual chemical synthesis, is straight forward and superior to a chemical synthesis of all positional singly <sup>13</sup>C-labelled isotopomers of terpenes.

## Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft (DI1536/7-2). We thank Professor Scott A. Snyder (University of

Chicago) for fruitful discussions on the structural revision of spiroviolene. Open access funding enabled and organized by Projekt DEAL.

## Conflict of Interest

The authors declare no conflict of interest.

**Keywords:** biosynthesis · enzyme mechanisms · isotopes · mass spectrometry · terpenes

- [1] J. S. Dickschat, *Angew. Chem. Int. Ed.* **2019**, *58*, 15964.
- [2] A. Minami, T. Ozaki, C. Liu, H. Oikawa, *Nat. Prod. Rep.* **2018**, *35*, 1330.
- [3] T. Mitsuhashi, I. Abe, *ChemBioChem* **2018**, *19*, 1106.
- [4] P. Rabe, J. Rinkel, E. Dolja, T. Schmitz, B. Nubbemeyer, T. H. Luu, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2017**, *56*, 2776.
- [5] T. Roncal, S. Cordobés, U. Ugalde, Y. He, O. Sterner, *Tetrahedron Lett.* **2002**, *43*, 6799.
- [6] L. Du, D. Li, T. Zhu, S. Cai, F. Wang, X. Xiao, Q. Gu, *Tetrahedron* **2009**, *65*, 1033.
- [7] S. Niu, Z. Fan, X. Tang, Q. Liu, Z. Shao, G. Liu, X.-W. Yang, *Tetrahedron Lett.* **2018**, *59*, 375.
- [8] S. Niu, Z.-W. Fan, C.-L. Xie, Q. Liu, Z.-H. Luo, G. Liu, X.-W. Yang, *J. Nat. Prod.* **2017**, *80*, 2174.
- [9] T. Mitsuhashi, T. Kikuchi, S. Hoshino, M. Ozeki, T. Awakawa, S.-P. Shi, M. Fujita, I. Abe, *Org. Lett.* **2018**, *20*, 5606.
- [10] H. M. Chi, C. J. F. Cole, P. Hu, C. A. Taylor, S. A. Snyder, *Chem. Sci.* **2020**, *11*, 10939-10944.
- [11] J. Rinkel, J. S. Dickschat, *Org. Lett.* **2019**, *21*, 2426.
- [12] L. Lauterbach, J. Rinkel, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2018**, *57*, 8280.
- [13] J. W. Cornforth, R. H. Cornforth, G. Popják, L. Yengoyan, *J. Biol. Chem.* **1966**, *241*, 3970.
- [14] Y. J. Hong, D. J. Tantillo, *Aust. J. Chem.* **2017**, *70*, 362.
- [15] F. W. McLafferty, F. Turecek, *Interpretation of Mass Spectra*, 4th ed., University Science Books, Mill Valley, **1993**.
- [16] D. S. Weinberg, C. Djerassi, *J. Org. Chem.* **1966**, *31*, 115.
- [17] J. Karliner, C. Djerassi, *J. Org. Chem.* **1966**, *31*, 1945.
- [18] J. L. Holmes, D. McGillivray, *Org. Mass Spectrom.* **1971**, *5*, 1349.
- [19] R. R. Muccino, C. Djerassi, *J. Am. Chem. Soc.* **1973**, *95*, 8726.
- [20] H. D. Lee, L. S. Eichmeier, D. M. Piatak, *Org. Mass Spectrom.* **1985**, *20*, 247.
- [21] D. Harris, S. McKinnon, R. K. Boyd, *Org. Mass Spectrom.* **1979**, *14*, 265.
- [22] P. Rabe, L. Barra, J. Rinkel, R. Riclea, C. A. Citron, T. A. Klapschinski, A. Janusko, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2015**, *54*, 13448.
- [23] P. Rabe, T. A. Klapschinski, J. S. Dickschat, *ChemBioChem* **2016**, *17*, 1333.
- [24] P. Rabe, J. S. Dickschat, *Beilstein J. Org. Chem.* **2016**, *12*, 1380.
- [25] J. Rinkel, P. Rabe, J. S. Dickschat, *Eur. J. Org. Chem.* **2019**, 351.
- [26] J. Rinkel, J. S. Dickschat, *Org. Lett.* **2019**, *21*, 2426.
- [27] T. Toyomasu, A. Kaneko, T. Tokiwano, Y. Kanno, R. Niida, S. Miura, T. Nishioka, C. Ikeda, W. Mitsuhashi, T. Dairi, T. Kawano, H. Oikawa, N. Kato, T. Sassa, *J. Org. Chem.* **2009**, *74*, 1541.
- [28] J. Rinkel, S. T. Steiner, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2019**, *58*, 9230.
- [29] P. Rabe, K. A. K. Pahirulzaman, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2015**, *54*, 6041.
- [30] Q. Jia, X. Chen, T. G. Köllner, J. Rinkel, J. Fu, J. Labbé, W. Xiong, J. S. Dickschat, J. Gershenzon, F. Chen, *Sci. Rep.* **2019**, *9*, 9223.

Manuscript received: October 1, 2020

Revised manuscript received: October 21, 2020

Accepted manuscript online: October 21, 2020

Version of record online: November 10, 2020

## Appendix P

### **Mechanistic Characterisation of the Bacterial Sesterviridene Synthase from *Kitasatospora viridis***

*Angew. Chem. Int. Ed.* **2023**, *62*, e202306429

DOI: 10.1002/ange.202306429

*Angew. Chem.* **2023**, *135*, e202306429

DOI: 10.1002/anie.202306429

## Biosynthesis

# Mechanistic Characterisation of the Bacterial Sesterviridene Synthase from *Kitasatospora viridis*

Houchao Xu, Gregor Schnakenburg, Bernd Goldfuss, and Jeroen S. Dickschat\*

**Abstract:** A gene coding for a terpene synthase homolog from *Kitasatospora viridis* was cloned and expressed in *Escherichia coli*. The purified recombinant protein possessed sesterterpene synthase activity and efficiently converted geranylgeranyl diphosphate (GGPP) with 19% yield into the sesterterpene hydrocarbon sesterviridene A. Large scale enzymatic conversions also allowed for the isolation of two side products that are generated with very low yields of ca. 0.1%. Several derivatives of sesterviridene A were obtained by chemical transformations, securing the NMR-based structural assignments. The absolute configuration of sesterviridene A was determined by chemical correlation using stereoselectively deuterated precursors and by anomalous dispersion X-ray crystallography. The cyclisation mechanism from GGPP to sesterviridene A was extensively studied through isotopic labelling experiments and DFT calculations.

Terpenoids constitute with about 100000 known compounds the largest class of natural products. Despite this fact, all terpenoids originate from just two C<sub>5</sub> precursors, the electrophile dimethylallyl diphosphate (DMAPP) and the nucleophile isopentenyl diphosphate (IPP), that can be fused to yield higher oligoprenyl diphosphates through the action of prenyltransferases (PTs). These enzymes successively generate the monoterpene precursor geranyl diphosphate (GPP, C<sub>10</sub>),<sup>[1]</sup> the sesquiterpene precursor farnesyl

diphosphate (FPP, C<sub>15</sub>),<sup>[2]</sup> and then geranylgeranyl diphosphate (GGPP, C<sub>20</sub>) and geranylgeranyl diphosphate (GGPP, C<sub>20</sub>) and geranylgeranyl diphosphate (GGPP, C<sub>20</sub>) as precursors to di- and sesterterpenes.<sup>[3,4]</sup> Triterpenes usually arise from squalene, but it was recently demonstrated that also farnesylfarnesyl diphosphate (FFPP) can be formed and then converted into triterpenes by bifunctional fungal enzymes with a PT and a terpene synthase (TS) domain.<sup>[5]</sup> Also fungal sesterterpenes, a comparably small subclass of terpenoids, are generated by such PT-TSs. The characterised enzymes include AcOS from *Aspergillus clavatus* for ophiobolin F (**1**),<sup>[6]</sup> EvSS from *Emericella varicolor* for stellatriene (**2**),<sup>[7]</sup> NfSS from *Neosartorya fischeri* for sesterfisherol (**3**),<sup>[8]</sup> and several other examples (Figure 1).<sup>[9]</sup> In plants sesterterpene synthases are usually monofunctional enzymes that are genetically clustered with a gene for a GGPP synthase (GFPPS). Characterised enzymes include inter alia AtTPS19 from *Arabidopsis thaliana* for retigeranin B (**4**),<sup>[10]</sup> AtTPS30 from the same plant for astellatene (**5**),<sup>[11]</sup> Br580 from *Brassica rapa* for brarapadienes,<sup>[12]</sup> and LcTPS2 from *Leucosceptrum canum* that produces a mixture of macrocyclic sesterterpenes.<sup>[13]</sup> Also in bacteria sesterterpene biosynthesis is catalysed by two discrete enzymes, a GFPPS and a TS. Besides the UbiA related sesterterpene synthase for somaliensenes from *Streptomyces somaliensis*<sup>[14]</sup> and the type II terpene synthase involved in atylovene biosynthesis,<sup>[15]</sup> the only two characterised type I sesterterpene synthases are SmTS1 from

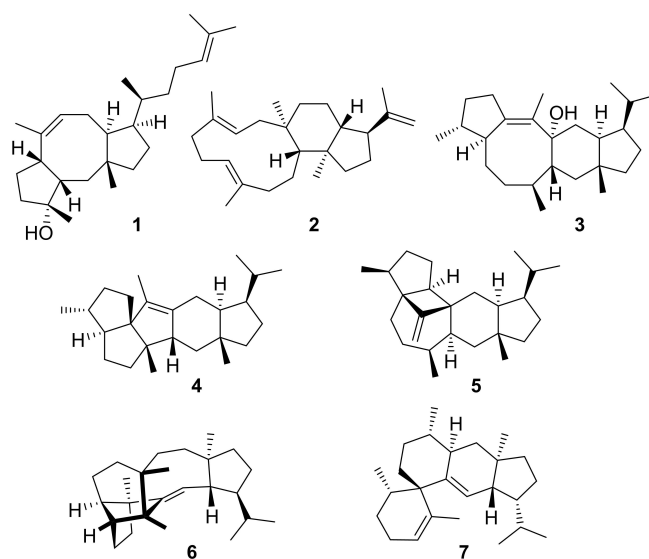


Figure 1. Products of characterised sesterterpene synthases.

[\*] H. Xu, Prof. Dr. J. S. Dickschat  
 Kekulé-Institute for Organic  
 Chemistry and Biochemistry,  
 University of Bonn  
 Gerhard-Domagk-Straße 1, 53121 Bonn (Germany)  
 E-mail: dickschat@uni-bonn.de

Dr. G. Schnakenburg  
 Institute for Inorganic Chemistry,  
 University of Bonn  
 Gerhard-Domagk-Straße 1, 53121 Bonn (Germany)

Prof. Dr. B. Goldfuss  
 Department for Chemistry,  
 University of Cologne  
 Greinstraße 4, 50939 Cologne (Germany)

© 2023 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

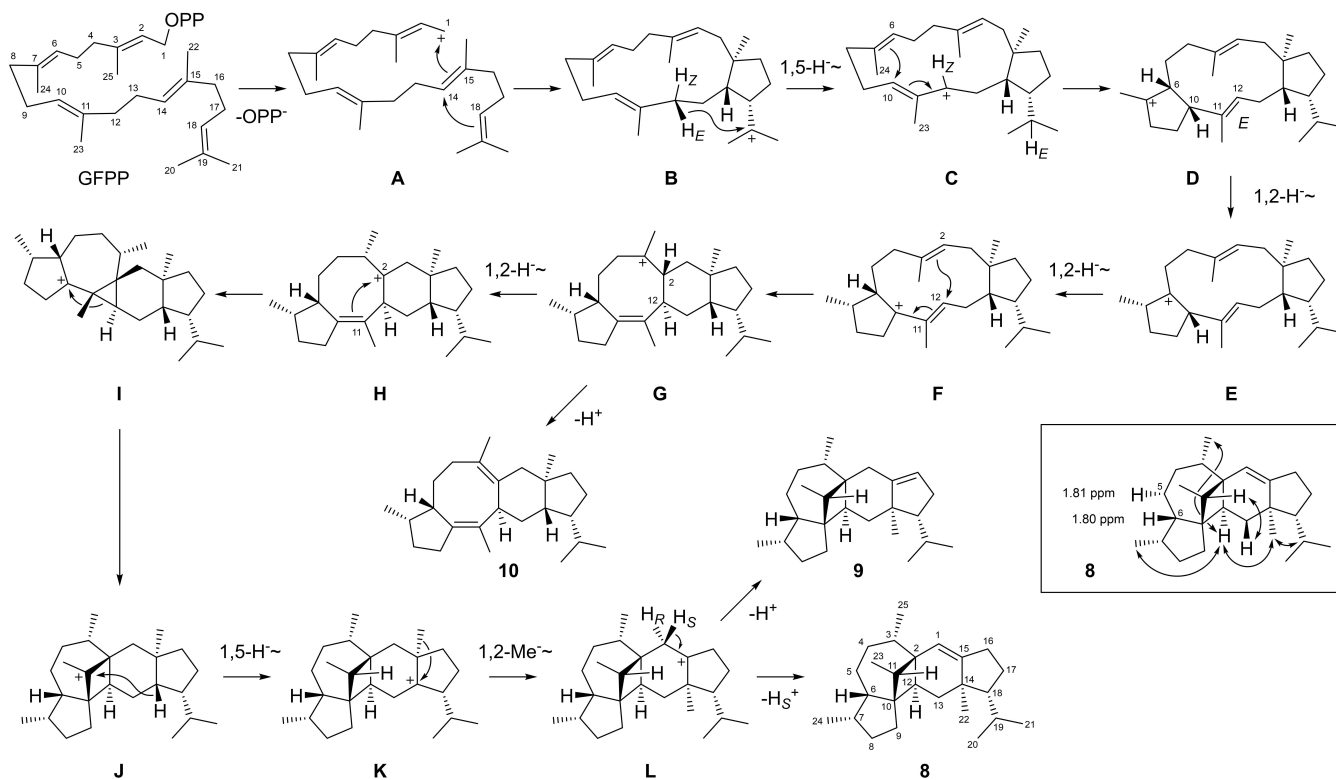
*Streptomyces mobaraensis* with the main product sesterbaraene A (**6**),<sup>[16]</sup> and the sesterviolene (**7**) synthase SvSS from *Streptomyces violens*.<sup>[17]</sup> After our recent discovery of the sesquiterpene synthase for kitaviridene from *Kitasatospora viridis*,<sup>[18]</sup> we now report on the identification of a sesterterpene synthase from the same organism and its mechanistic characterisation through isotopic labelling experiments and DFT calculations.

The gene of a terpene synthase from *Kitasatospora viridis* that does not cluster with other characterised enzymes in a phylogenetic tree constructed from 4018 terpene synthase sequences (Figure S1) was cloned and expressed in *Escherichia coli*. The predicted amino acid sequence shows the presence of all highly conserved motifs and residues required for functionality (Figure S2). The purified recombinant protein (Figure S3) was incubated with different oligoprenyl diphosphates, showing the efficient conversion of GFPP into one main and several minor sesterterpene hydrocarbons, while GGPP was converted with low efficiency into cembrene A, and GPP and FPP gave only elimination and hydrolysis products (Figures S4 and S5). The main product was isolated from an enzymatic transformation of 80.0 mg GFPP, yielding 10.0 mg (19%) of a sesterterpene hydrocarbon. Its structure was elucidated by NMR spectroscopy (Figures S6–S13, Table S2), revealing the structure of a pentacyclic compound for which we propose the name sesterviridene A (**8**). Repeated collection of material from enzymatic conversions of three batches of 80.0 mg GFPP also allowed for the isolation of two side products, sesterviridene B (**9**, 0.1 mg, 0.06%) and sestervir-

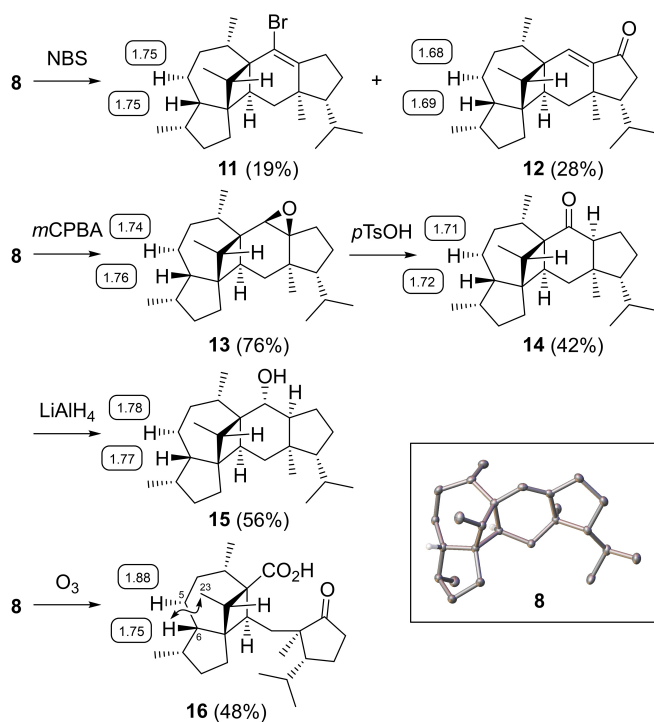
idene C (**10**, 0.2 mg, 0.13%) whose structures were also determined by NMR spectroscopy (Figures S14–S27, Tables S3 and S4).

While the observed NOESY correlations allowed for the unambiguous assignment of the relative configuration for most stereogenic centers of **8**, H-6 and one of the hydrogens at C-5 showed almost identical chemical shifts (box in Scheme 1), which resulted in difficulties to assign the relative configuration at C-6. In an attempt to resolve the situation, several derivatives of **8** were prepared (Scheme 2). Treatment of **8** with NBS yielded 1-bromosesterviridene (**11**) and sesterviridene (**12**) (Figures S28–S43, Tables S5 and S6). The oxidation with *m*CPBA resulted in the highly diastereoselective formation of sesterviridene epoxide (**13**) without noticeable formation of the other diastereoisomer (Figures S44–S51, Table S7). Further conversion through an acid catalysed epoxide-ketone rearrangement gave sesterviridone (**14**) that was reduced using LiAlH<sub>4</sub> with high diastereoselectivity to sesterviridol (**15**) (Figures S52–S67, Tables S8 and S9). Unfortunately, because of overlapping NMR signals for H-6 and one of the hydrogens at C-5 the assignment of the relative configuration at C-6 was similarly difficult for all these compounds.

Ozonolysis of **8** resulted in *seco*-sesterviridonic acid (**16**) (Figures S68–S75, Table S10) for which clearly resolved signals for H5<sub>α</sub> (δ=1.88 ppm) and H6 (δ=1.75 ppm) were obtained that allowed to conclude on the relative configuration of **16** based on a key NOESY correlation between H6 and Me23. This in turn finally resolved the relative configuration of **8**. Crystallisation of compound **8** gave



**Scheme 1.** Cyclisation mechanism from GFPP to **8**, **9** and **10** by StvirS.



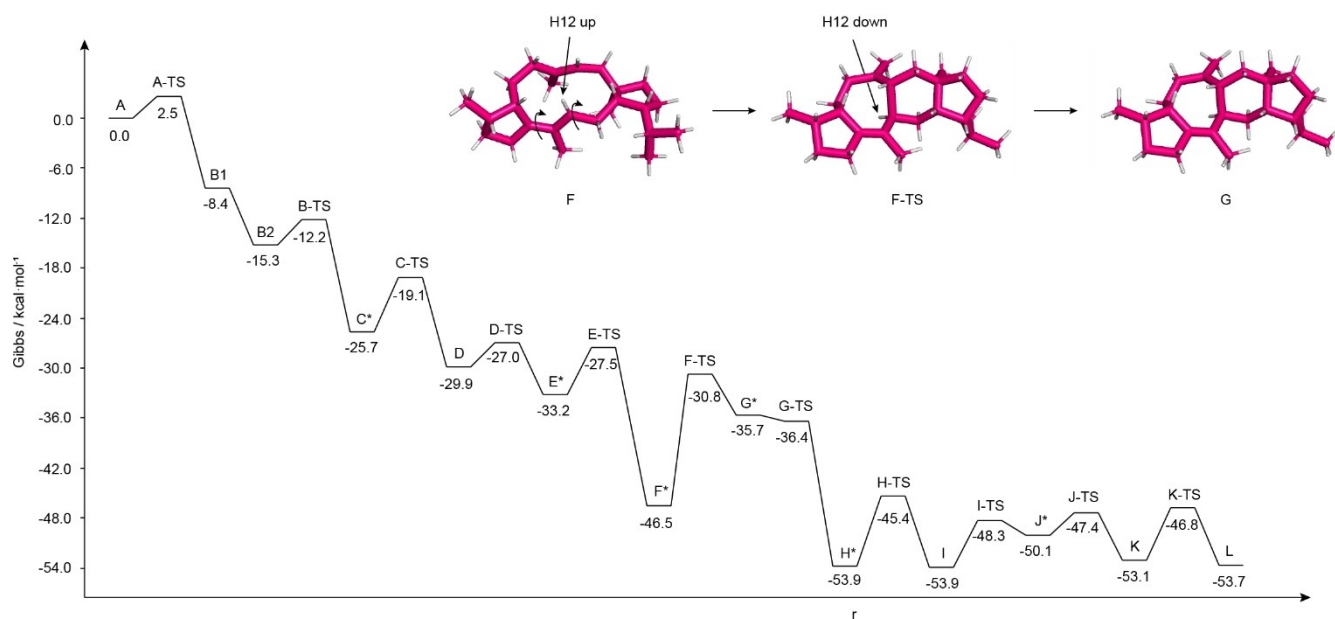
**Scheme 2.** Chemical transformations of **8**. Data in small boxes indicate  $^1\text{H}$  chemical shifts in ppm. The double headed arrow in **16** indicates a key NOESY correlation. Box: ORTEP representation of **8** by anomalous CuK $\alpha$  X-ray analysis (Flack parameter  $-0.17(14)$ ).

access to its absolute configuration by anomalous dispersion CuK $\alpha$  X-ray analysis as (*2R,3S,6S,7S,10R,11S,14R,18R*)-**8** (box in Scheme 2, Table S11, Figure S76).<sup>[19]</sup> The same absolute configuration was deduced through chemical correlation using the stereoselectively deuterated substrates (*R*)- and (*S*)-(1- $^{13}\text{C}$ ,1- $^2\text{H}$ )IPP<sup>[20]</sup> that were converted into **8** with *E. coli* isopentenyl diphosphate isomerase (IDI),<sup>[21]</sup> GFPPS from *S. mobaraensis*,<sup>[16]</sup> and StvirS (Figure S77, a summary of labelling experiments conducted in this study is given in Table S12). Also DMAPP and (*E*)- and (*Z*)-(4- $^{13}\text{C}$ ,4- $^2\text{H}$ )IPP<sup>[22]</sup> were converted into **8** with GFPPS and StvirS pointing to the same absolute configuration (Figure S78). The interpretation of the results from these experiments builds on the known stereochemical course of oligoprenyl diphosphate biosynthesis as established by Cornforth and co-workers in their fundamental investigations on terpene biosynthesis.<sup>[23]</sup>

The proposed cyclisation mechanism from GFPP to **8** is initiated by substrate ionisation through diphosphate abstraction to **A** (Scheme 1). A 1,15–14,18-cyclisation to **B** may be followed by a 1,5-hydride migration to **C** which can further react in a 6,10-cyclisation to **D**. Two sequential 1,2-hydride migrations lead via **E** to **F** that can undergo a 2,12-cyclisation to **G**, the precursor of **10** by deprotonation. Another 1,2-hydride shift to **H** may be followed by 2,11-cyclisation to **I**. Its rearrangement to **J**, a 1,5-hydride transfer to **K**, 1,2-methyl migration to **L** and final deprotonation explain the biosynthesis of **8** and **9**.

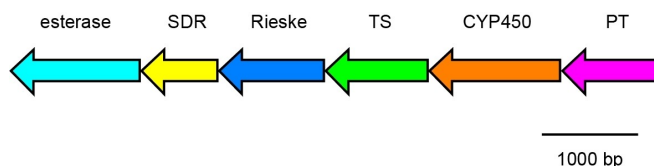
A series of labelling experiments was performed to support this biosynthetic hypothesis. The enzymatic conversion of all 25 isotopomers of ( $^{13}\text{C}$ )GFPP—these compounds were either obtained by chemical synthesis or enzymatically using GFPPS from  $^{13}\text{C}$ -labelled oligoprenyl diphosphate precursors according to Table S12—with StvirS proceeded with incorporation of the  $^{13}\text{C}$ -label into positions that are congruent with the mechanism of Scheme 1 in all cases (Figure S79). Specifically, the skeletal rearrangements from **I** to **J** and from **K** to **L** were supported by these experiments. The 1,5-hydride shift from **B** to **C** including its stereochemical course was investigated using (7- $^{13}\text{C}$ )GPP<sup>[9d]</sup> in conjunction with (*E*)- or (*Z*)-(4- $^{13}\text{C}$ ,4- $^2\text{H}$ )IPP, GFPPS and StvirS. Analysis of the products through  $^{13}\text{C}$  NMR showed a slightly upfield shifted triplet for C19 ( $\Delta\delta = -0.48$  ppm,  $^1J_{\text{C,D}} = 19.2$  Hz) from (*E*)-(4- $^{13}\text{C}$ ,4- $^2\text{H}$ )IPP as a result of  $^{13}\text{C}$ - $^2\text{H}$  spin coupling, while an unchanged singlet was obtained with (*Z*)-(4- $^{13}\text{C}$ ,4- $^2\text{H}$ )IPP, revealing the specific migration of a hydrogen originating from H $_4\text{E}$  of IPP (H $_E$ ) (Figure S80). The 1,2-hydride shift from **D** to **E** was evident from the enzymatic conversion of (3- $^{13}\text{C}$ ,2- $^2\text{H}$ )GGPP<sup>[24]</sup> and IPP with GFPPS and StvirS yielding a triplet for C7 ( $\Delta\delta = -0.50$  ppm,  $^1J_{\text{C,D}} = 19.8$  Hz, Figure S81). The subsequent 1,2-hydride migration from **E** to **F** was supported by the observed triplet ( $\Delta\delta = -0.53$  ppm,  $^1J_{\text{C,D}} = 19.2$  Hz) upon incubation of (2- $^2\text{H}$ )FPP<sup>[24]</sup> and (2- $^{13}\text{C}$ )IPP<sup>[25]</sup> with GFPPS and StvirS (Figure S82). Similarly, the 1,2-hydride transfer from **G** to **H** was studied through the enzymatic transformation of (3- $^{13}\text{C}$ ,2- $^2\text{H}$ )GFPP that was synthesised as shown in Scheme S5 with StvirS ( $\Delta\delta = -0.51$  ppm,  $^1J_{\text{C,D}} = 19.2$  Hz, Figure S83). Finally, the 1,5-hydride shift from **J** to **K** was shown by the incubation of (2- $^2\text{H}$ )GPP<sup>[26]</sup> and (3- $^{13}\text{C}$ )IPP<sup>[24]</sup> with GFPPS and StvirS, resulting in a triplet with  $\Delta\delta = -0.55$  ppm and  $^1J_{\text{C,D}} = 20.0$  Hz (Figure S84). The GC/MS analysis of the products obtained with IDI, GFPPS and StvirS from (*R*)- and (*S*)-(1- $^{13}\text{C}$ ,1- $^2\text{H}$ )IPP also demonstrated the specific loss of the 1-*pro-S* hydrogen (H $_S$ ) from C1 in the terminal deprotonation to **8** (Figure S85).

The cyclisation mechanism from GFPP to **8** was also investigated computationally (Figure 2, Table S13). The overall reaction from intermediate **A** to **L** is with  $-53.7$  kcal mol $^{-1}$  highly exergonic and all transition state barriers were low (the highest transition state barrier is with  $+15.7$  kcal mol $^{-1}$  found for the cyclisation from **F** to **G** in **F-TS**). Notably, the 1,15–14,18-cyclisation from **A** to **B** is an asynchronous process in which first the macrocyclic ring is formed (**B1** in Figure 2), which is followed by a barrierless closure of the cyclopentane ring (**B2**). The *trans* orientation of H6 and Me24 in **8** requires a *cis* orientation of H6 and H10 in **D** so that the two sequential 1,2-hydride shifts towards **F** result in the correct stereochemistry (Scheme 1). This in turn requires a specific conformation in **C** in which H6 and H10 point up and consequently Me23 and Me24 point down which leads to an 11*E* double bond in **D** with H12 pointing up. However, the *trans* fusion at C2 and C12 in **G** shows that after the ring closure from **F** to **G** H12 points down, which needs turning of the olefinic plane C11=C12 during the cyclisation process (structures in Figure 2), explaining the comparably high energy barrier for this step.



**Figure 2.** Computed energy profile (mPW1PW91/6-311+G(d,p)//B97D3/6-31G(d,p), 298.15 K) for the cationic cascade reaction from intermediate A to L as shown in Scheme 1 and structures of F, F-TS and G. Asterisks indicate intermediates for which minor conformational changes were required (only the energy of the more stable conformer is shown in these cases).

In summary, we have identified a bacterial sesterterpene synthase from the actinomycete *Kitasatospora viridis* for sesterterviridene A. The structure of this compound was elucidated by NMR spectroscopy and further secured by chemical derivatisation, while the absolute configuration was determined through stereoselective labelling and anomalous dispersion X-ray crystallography. This method requires the presence of a “heavy atom”, and as demonstrated here with the modern analytical equipment even carbon can reliably serve for this purpose. Sesterterviridene A has a unique skeleton, but is structurally similar to astellatene (**5**).<sup>[11]</sup> Also the enzyme is unique and no closely related homologs are found in other bacteria; the closest relative is with only 34% amino acid sequence identity an uncharacterised terpene synthase from *Streptomyces subrutilus* (Figure S1). As shown by isotopic labelling experiments in conjunction with computational chemistry, this sesterterpene hydrocarbon arises through a cyclisation cascade that involves complex rearrangement reactions and an unusual 1,5-hydride transfer, leading to the unusual skeleton of sesterterviridene A. The analysis of the volatiles emitted from a *K. viridis* agar plate culture using a closed-loop stripping apparatus (CLSA)<sup>[27]</sup> that is suitable for the extraction of sesterterpenes<sup>[16]</sup> did not show the production of any sesterterpenes. One possible explanation is that the gene coding for StvirS is not expressed under laboratory culture conditions. Alternatively, sesterterviridene may be rapidly converted into another oxidised terpenoid through the action of enzymes that are genetically clustered with the gene for StvirS (Figure 3). The oxidation product(s) of this gene cluster are currently unknown and will be investigated in our future research. Conclusively, this study gives another



**Figure 3.** Biosynthetic gene cluster containing the gene for the sesterterviridene synthase (TS, green). Other genes are for a prenyltransferase (PT, purple), a cytochrome P450 monooxygenase (CYP450, orange), a Rieske family protein (blue), a short chain dehydrogenase/reductase (SDR, yellow), and an esterase (cyan).

example for the discovery of unknown compounds through genome mining approaches.

### Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft DFG (project number 271721844) and supported by the computing center of the University of Cologne (RRZK), providing CPU time on the DFG-funded supercomputer CHEOPS. We thank Andreas Schneider for HPLC purifications. 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 Supporting Information of this article. Crystallographic data for **8** were deposited in the Cambridge Crystallographic Data Centre under accession no. CCDC 2261343.

**Keywords:** Biosynthesis · Enzyme Mechanisms · Hydrocarbons · Isotopes · Terpenoids

- [1] R. Croteau, P. T. Purkett, *Arch. Biochem. Biophys.* **1989**, *271*, 524–535.
- [2] F. Lynen, H. Eggerer, U. Henning, I. Kessel, *Angew. Chem.* **1958**, *70*, 738–742.
- [3] D. L. Nandi, J. W. Porter, *Arch. Biochem. Biophys.* **1964**, *105*, 7–19.
- [4] A. Tachibana, *FEBS Lett.* **1994**, *341*, 291–294.
- [5] H. Tao, L. Lauterbach, G. Bian, R. Chen, A. Hou, T. Mori, S. Cheng, B. Hu, L. Lu, X. Mu, M. Li, N. Adachi, M. Kawasaki, T. Moriya, T. Senda, X. Wang, Z. Deng, I. Abe, J. S. Dickschat, T. Liu, *Nature* **2022**, *606*, 414–419.
- [6] R. Chiba, A. Minami, K. Gomi, H. Oikawa, *Org. Lett.* **2013**, *15*, 594–597.
- [7] Y. Matsuda, T. Mitsunashi, Z. Quan, I. Abe, *Org. Lett.* **2015**, *17*, 4644–4647.
- [8] Y. Ye, A. Minami, A. Mandi, C. Liu, T. Taniguchi, T. Kuzuyama, K. Monde, K. Gomi, H. Oikawa, *J. Am. Chem. Soc.* **2015**, *137*, 11846–11853.
- [9] a) Y. Matsuda, T. Mitsunashi, S. Lee, M. Hoshino, T. Mori, M. Okada, H. Zhang, F. Hayashi, M. Fujita, I. Abe, *Angew. Chem. Int. Ed.* **2016**, *55*, 5785–5788; b) M. Okada, Y. Matsuda, T. Mitsunashi, S. Hoshino, T. Mori, K. Nakagawa, Z. Quan, B. Qin, H. Zhang, F. Hayashi, H. Kawaide, I. Abe, *J. Am. Chem. Soc.* **2016**, *138*, 10011–10018; c) K. Narita, H. Sato, A. Minami, K. Kudo, L. Gao, C. Liu, T. Ozaki, M. Kodama, X. Lei, T. Taniguchi, K. Monde, M. Yamazaki, M. Uchiyama, H. Oikawa, *Org. Lett.* **2017**, *19*, 6696–6699; d) T. Mitsunashi, J. Rinkel, M. Okada, I. Abe, J. S. Dickschat, *Chem. Eur. J.* **2017**, *23*, 10053–10057; e) G. Bian, Y. Han, A. Hou, Y. Yuan, X. Liu, Z. Deng, T. Liu, *Metab. Eng.* **2017**, *42*, 1–8; f) Z. Quan, J. S. Dickschat, *Org. Lett.* **2020**, *22*, 7552–7555; g) Z. Quan, J. S. Dickschat, *Org. Biomol. Chem.* **2020**, *18*, 6072–6076; h) J. Guo, Y. Cai, F. Cheng, C. Yang, W. Zhang, W. Yu, J. Yan, Z. Deng, K. Hong, *Org. Lett.* **2021**, *23*, 1525–1529; i) L. Jiang, X. Zhang, Y. Sato, G. Zhu, A. Minami, W. Zhang, T. Ozaki, B. Zhu, Z. Wang, X. Wang, K. Lv, J. Zhang, Y. Wang, S. Gao, C. Liu, T. Hsiang, L. Zhang, H. Oikawa, X. Liu, *Org. Lett.* **2021**, *23*, 4645–4650; j) Z. Quan, A. Hou, B. Goldfuss, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2022**, *61*, e202117273.
- [10] J. Shao, Q.-W. Chen, H.-J. Lv, J. He, Z.-F. Liu, Y.-N. Lu, H.-L. Liu, G.-D. Wang, Y. Wang, *Org. Lett.* **2017**, *19*, 1816–1819.
- [11] A. C. Huang, S. A. Kautsar, Y. J. Hong, M. H. Medema, A. D. Bond, D. J. Tantillo, A. Osbourn, *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E6005–E6014.
- [12] A. C. Huang, Y. J. Hong, A. D. Bond, D. J. Tantillo, A. Osbourn, *Angew. Chem. Int. Ed.* **2018**, *57*, 1291–1295.
- [13] Y.-G. Chen, D.-S. Li, Y. Ling, Y.-C. Liu, Z.-L. Zuo, L.-S. Gan, S.-H. Luo, J. Hua, D.-Y. Chen, F. Xu, M. Li, K. Guo, Y. Liu, J. Gershenzon, S.-H. Li, *Angew. Chem. Int. Ed.* **2021**, *60*, 25468–25476.
- [14] Y. Yang, Y. Zhang, S. Zhang, Q. Chen, K. Ma, L. Bao, Y. Tao, W. Yin, G. Wang, H. Liu, *J. Nat. Prod.* **2018**, *81*, 1089–1092.
- [15] S.-H. Kim, W. Lu, M. K. Ahmadi, D. Montiel, M. A. Ternei, S. F. Brady, *ACS Synth. Biol.* **2019**, *8*, 109–118.
- [16] A. Hou, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2020**, *59*, 19961–19965.
- [17] B. Gu, B. Goldfuss, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2023**, *62*, e202215688.
- [18] H. Xu, B. Goldfuss, J. S. Dickschat, *Org. Lett.* **2023**, *25*, 3330–3334.
- [19] Deposition Number 2261343 contains the supplementary crystallographic data for this paper. These data are provided free of charge by the joint Cambridge Crystallographic Data Centre and Fachinformationszentrum Karlsruhe Access Structures service.
- [20] J. Rinkel, J. S. Dickschat, *Org. Lett.* **2019**, *21*, 2426–2429.
- [21] F. M. Hahn, A. P. Hurlburt, C. D. Poulter, *J. Bacteriol.* **1999**, *181*, 4499–4504.
- [22] L. Lauterbach, J. Rinkel, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2018**, *57*, 8280–8283.
- [23] J. W. Cornforth, R. H. Cornforth, G. Popjak, L. Yengoyan, *J. Biol. Chem.* **1966**, *241*, 3970–3987.
- [24] P. Rabe, J. Rinkel, E. Dolja, T. Schmitz, B. Nubbemeyer, T. H. Luu, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2017**, *56*, 2776–2779.
- [25] J. Rinkel, L. Lauterbach, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2019**, *58*, 452–455.
- [26] G. Bian, J. Rinkel, Z. Wang, L. Lauterbach, A. Hou, Y. Yuan, Z. Deng, T. Liu, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2018**, *57*, 15887–15890.
- [27] K. Grob, F. Zürcher, *J. Chromatogr.* **1976**, *117*, 285–294.

Manuscript received: May 8, 2023

Accepted manuscript online: June 7, 2023

Version of record online: ■■, ■■

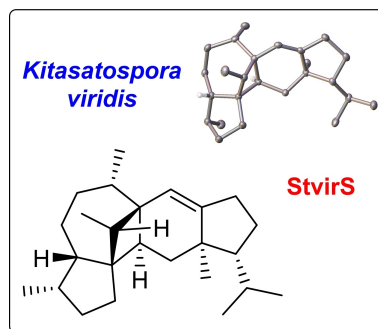


## Communications

## Biosynthesis

H. Xu, G. Schnakenburg, B. Goldfuss,  
J. S. Dickschat\* ————— e202306429

Mechanistic Characterisation of the Bacterial Sesterviridene Synthase from *Kitasatospora viridis*



A sesterterpene synthase has been discovered from *Kitasatospora viridis* that produces sesterviridene A in high yield. Several derivatives were synthesised, confirming the NMR based structure elucidation. The absolute configuration was determined through chemical correlation and anomalous dispersion X-ray. The cyclisation mechanism of the sesterviridene synthase *StvirS* was deeply studied in isotopic labelling experiments and DFT calculations.

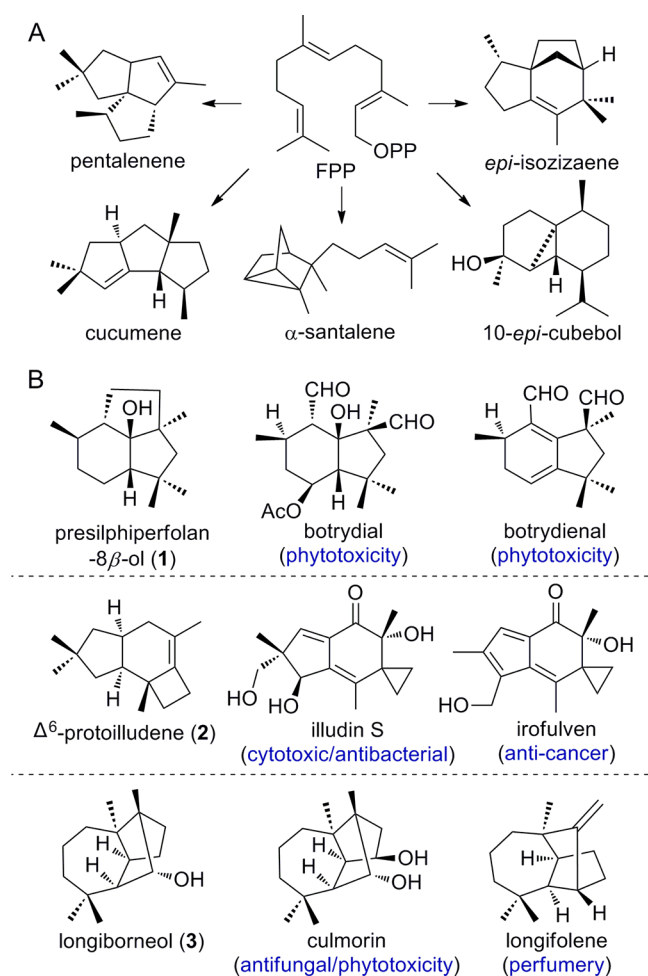
## Appendix Q

### **Structural Insights into Three Sesquiterpene Synthases for the Biosynthesis of Tricyclic Sesquiterpenes and Chemical Space Expansion by Structure-Based Mutagenesis**

*J. Am. Chem. Soc.* **2023**, *145*, 8474

DOI: [10.1021/jacs.3c00278](https://doi.org/10.1021/jacs.3c00278)





**Figure 1.** Tricyclic sesquiterpenes. (A) Five tricyclic sesquiterpenes for which the structures of their sesquiterpene synthases have been solved. (B) The three tricyclic sesquiterpenes (1–3) in this study and bioactive sesquiterpenoids (biological activities are shown in parentheses) derived from 1–3.

STSs are capable of converting FPP into polycyclic products, especially given that synthetic approaches for the construction of these complex frameworks are highly challenging. For this purpose, we selected three STSs involved in the biosynthesis of tricyclic sesquiterpenes for structural and mechanistic studies. This includes presilphiperfolan-8 $\beta$ -ol (1) synthase BcBOT2 from the fungus *Botrytis cinerea*<sup>17</sup> involved in the biosynthesis of botrydial and botrydienal (Figure 1B), the causal agents of the gray mold disease,<sup>17–23</sup>  $\Delta^6$ -protoilludene (2) synthase DbPROS from the fungus *Dendrothele bispora*<sup>24</sup> that makes the parent hydrocarbon of the antibacterial and cytotoxic illudin S<sup>25</sup> and the anticancer agent irofulven (Figure 1B),<sup>26</sup> and longiborneol (3) synthase CLM1 from the fungus *Fusarium graminearum*<sup>27a</sup> that generates the biosynthetic precursor of the antifungal and phytotoxic culmorin,<sup>27</sup> which can also be converted into the fragrant longifolene (Figure 1B).<sup>28</sup>

The three sesquiterpenes (1–3) represent distinct types of tricyclic skeletons, and no structures for their synthases have been reported. We here report on the crystal structures of apo-BcBOT2, BcBOT2 complexed with the benzyltriethylammonium cation (BTAC), diphosphate (PPi), and Mg<sup>2+</sup> (designated BcBOT2-BTAC-PPi-Mg<sup>2+</sup>), DbPROS complexed with BTAC and Mg<sup>2+</sup> (designated DbPROS-BTAC-Mg<sup>2+</sup>), and CLM1 complexed with BTAC, PPi, and Mg<sup>2+</sup> (designated

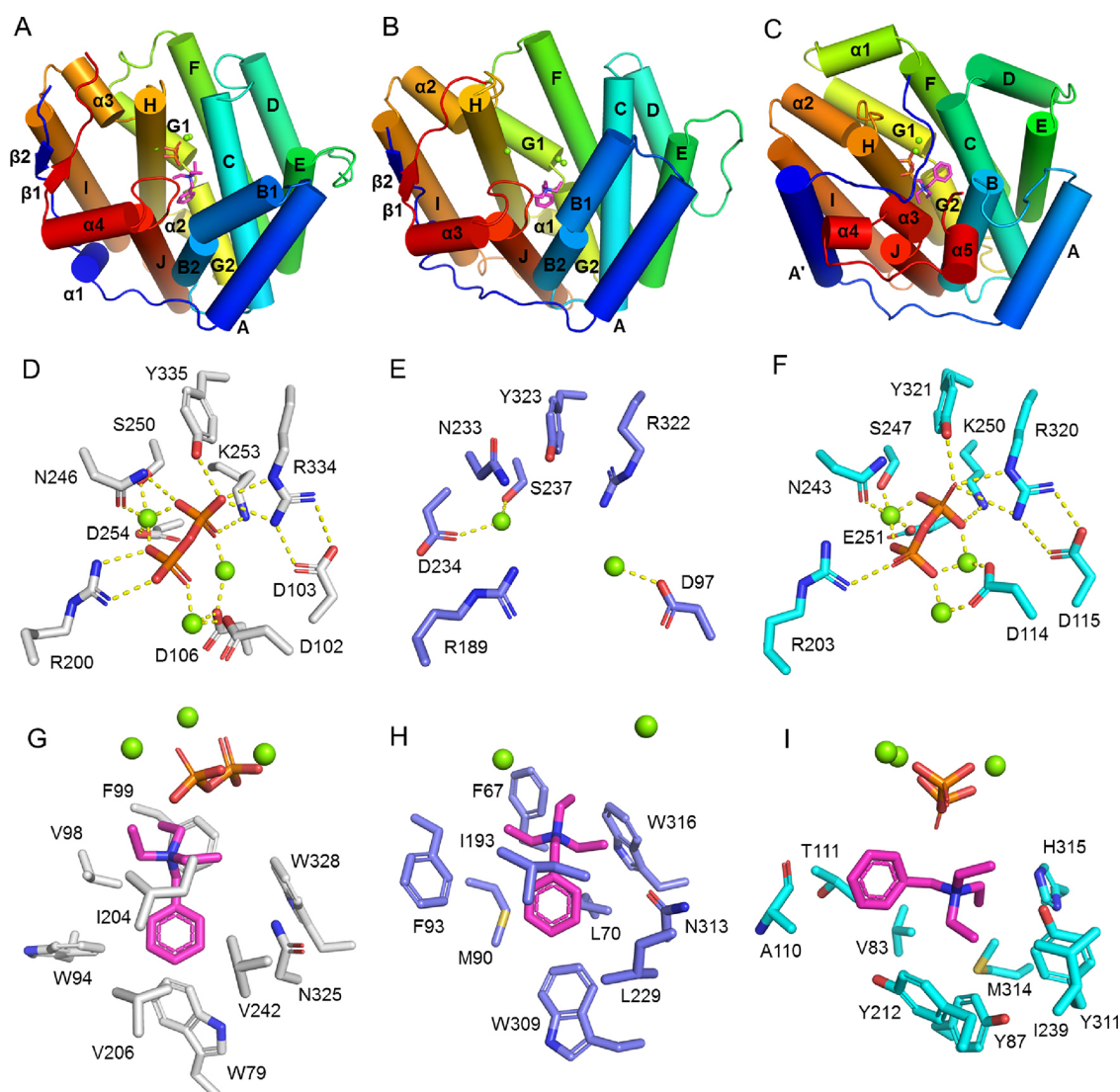
CLM1-BTAC-PPi-Mg<sup>2+</sup>). Quantum mechanics/molecular mechanics (QM/MM) studies were performed to investigate the interactions between all three synthases and the carbocationic intermediates during the cyclization cascades toward 1–3. This allowed for structure based site-directed mutagenesis experiments to change the product profile and expand the chemical space of polycyclic sesquiterpenes, in combination with a deep mechanistic investigation through isotopic labeling experiments.

## 2. RESULTS

**2.1. Crystal Structures of BcBOT2, DbPROS, and CLM1, and Structural Comparisons.** The codon-optimized *dbpros*, *bcbot2*, and *clm1* genes were expressed in *Escherichia coli* BL21 (DE3), and the purified recombinant proteins (Figure S2), with or without BTAC/PPi/Mg<sup>2+</sup>, were crystallized by the hanging drop vapor diffusion method (Supporting Information). The crystal structures of apo-BcBOT2, BcBOT2-BTAC-PPi-Mg<sup>2+</sup>, DbPROS-BTAC-Mg<sup>2+</sup>, and CLM1-BTAC-PPi-Mg<sup>2+</sup> were solved at 2.00, 2.78, 2.09, and 1.47 Å, respectively, by molecular replacement using the predicted models from ColabFold<sup>29</sup> or Tencent AI Lab (<https://drug.ai.tencent.com/cn>) as the templates (Supporting Information).

BcBOT2-BTAC-PPi-Mg<sup>2+</sup> possesses a typical class I terpene synthase fold consisting of 10 core  $\alpha$ -helices (designated A–J) and four short  $\alpha$ -helices ( $\alpha$ 1– $\alpha$ 4) (Figure 2A). The N- and C-termini interact with each other by two small  $\beta$  strands ( $\beta$ 1 and  $\beta$ 2) forming an antiparallel  $\beta$  sheet. Four molecules are found in one asymmetric unit of BcBOT2-BTAC-PPi-Mg<sup>2+</sup> (Figure S3A), which are highly identical to each other with a root mean square deviation (RMSD) of less than 0.19 Å for C $\alpha$  atoms (Figure S3B). The conserved metal binding motifs D<sup>102</sup>DQFD<sup>106</sup> (for Mg<sup>2+</sup><sub>A</sub> and Mg<sup>2+</sup><sub>C</sub> binding) and N<sup>246</sup>DVLS<sup>250</sup>YRKD<sup>254</sup> (for Mg<sup>2+</sup><sub>B</sub> binding) are located on the helices C and H, respectively (Figure 2D). R334 and Y335 form a conserved RY pair that is widely encountered in class I terpene cyclases<sup>30</sup> and serves in the recognition of PPi through hydrogen bonds (Figure 2D). The conserved effector triad associated with initiation of cyclization reaction<sup>31</sup> is found at R200-T203-I204 in BcBOT2-BTAC-PPi-Mg<sup>2+</sup>. Apo-BcBOT2 shows a similar structural fold and analogous secondary structures (Figure S3C,D), but the region around the active pocket entrance is different (Figure S3E). After binding of BTAC and PPi, most dramatic changes were observed for the positioning of the Mg<sup>2+</sup> binding motifs and the RY pair (Figure S3F), reflecting the dynamic conformational rearrangement of BcBOT2 from the open to the closed state upon substrate binding.

Also, DbPROS-BTAC-Mg<sup>2+</sup> and CLM1-BTAC-PPi-Mg<sup>2+</sup> both show the typical class I terpene synthase fold. Interestingly, DbPROS-BTAC-Mg<sup>2+</sup> possesses a very similar secondary structure as BcBOT2-BTAC-PPi-Mg<sup>2+</sup>, with 10 prominent  $\alpha$ -helices (A–J), in addition to three short  $\alpha$ -helices ( $\alpha$ 1– $\alpha$ 3) and two terminal  $\beta$ -strands ( $\beta$ 1 and  $\beta$ 2) (Figure 2B and Figure S4A). In contrast, CLM1-BTAC-PPi-Mg<sup>2+</sup> shows 11 main  $\alpha$ -helices (A' and A–J), besides five short  $\alpha$ -helices ( $\alpha$ 1– $\alpha$ 5) and an N-terminal loop covering the active site entrance (Figure 2C and Figure S4B). The conserved Mg<sup>2+</sup> binding motifs are also observed for DbPROS-BTAC-Mg<sup>2+</sup> (D<sup>97</sup>EHSD<sup>101</sup> and N<sup>233</sup>DLCS<sup>237</sup>YNVE<sup>241</sup>) and CLM1-BTAC-PPi-Mg<sup>2+</sup> (D<sup>114</sup>DAEA<sup>118</sup> and N<sup>243</sup>DVLS<sup>247</sup>FYKE<sup>251</sup>), interacting with the Mg<sup>2+</sup> ions in each structure (Figure 2E,F). The RY



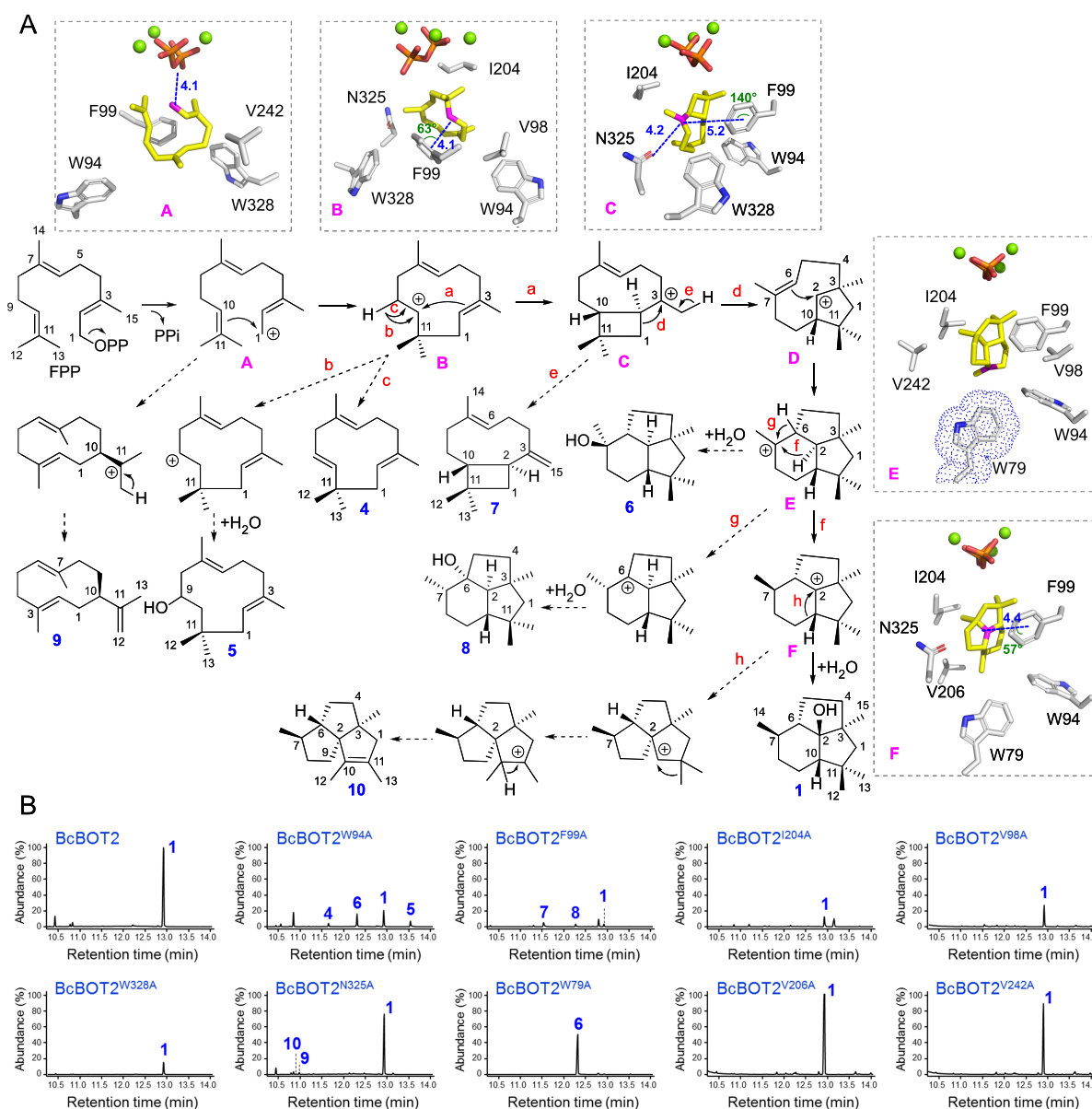
**Figure 2.** Crystal structures and active sites of BcBOT2-BTAC-PPi-Mg<sup>2+</sup>, DbPROS-BTAC-Mg<sup>2+</sup>, and CLM1-BTAC-PPi-Mg<sup>2+</sup>. The overall structures of (A) BcBOT2-BTAC-PPi-Mg<sup>2+</sup>, (B) DbPROS-BTAC-Mg<sup>2+</sup>, and (C) CLM1-BTAC-PPi-Mg<sup>2+</sup>. The metal binding motifs and RY pairs of (D) BcBOT2-BTAC-PPi-Mg<sup>2+</sup>, (E) DbPROS-BTAC-Mg<sup>2+</sup>, and (F) CLM1-BTAC-PPi-Mg<sup>2+</sup>. The BTAC binding pockets of (G) BcBOT2-BTAC-PPi-Mg<sup>2+</sup>, (H) DbPROS-BTAC-Mg<sup>2+</sup>, and (I) CLM1-BTAC-PPi-Mg<sup>2+</sup>. In (A)–(C), the structures are shown with coloring from blue at the N-terminus to red at the C-terminus. In (D)–(F), the interactions among Mg<sup>2+</sup> ions, phosphate, and active site residues are shown with yellow dashed lines. In (A)–(C) and (G)–(I), BTAC is shown in magenta. In (A)–(I), the Mg<sup>2+</sup> ions are shown as green spheres and diphosphate is shown in orange/red.

pairs and effector triads are defined in DbPROS-BTAC-Mg<sup>2+</sup> by R<sup>322</sup>Y<sup>323</sup> and R189-T192-I193 and in CLM1-BTAC-PPi-Mg<sup>2+</sup> by R<sup>320</sup>Y<sup>321</sup> and R203-D206-G207. Although BTAC is well defined in the active site pocket of DbPROS-BTAC-Mg<sup>2+</sup>, only two Mg<sup>2+</sup> ions and no PPi are found at the entrance of the active site, leading to less interactions with the Mg<sup>2+</sup> ions in DbPROS-BTAC-Mg<sup>2+</sup> (Figure 2E).

BTAC interacts with a series of residues in the active site of BcBOT2-BTAC-PPi-Mg<sup>2+</sup>, mainly nonpolar residues by hydrophobic interactions, including W79, W94, V98, F99, I204, V206, V242, N325, and W328 (Figure 2G). In DbPROS-BTAC-Mg<sup>2+</sup>, BTAC interacts with F67, L70, M90, F93, I193, L229, W309, N313, and W316 (Figure 2H). Herein, I193, N313, and W316 of DbPROS correspond to the residues I204, N325, and W328 in BcBOT2. These similar environments may cause a similar orientation of BTAC in BcBOT2-BTAC-PPi-Mg<sup>2+</sup> and DbPROS-BTAC-Mg<sup>2+</sup> (Figure 2G,H). In contrast,

BTAC binds to CLM1 in a distinct orientation by ~90° anticlockwise rotation compared to those observed in BcBOT2-BTAC-PPi-Mg<sup>2+</sup> and DbPROS-BTAC-Mg<sup>2+</sup> (Figure 2I). In this case, different residues in CLM1-BTAC-PPi-Mg<sup>2+</sup> are found to interact with BTAC including V83, Y87, A110, T111, Y212, I239, Y311, M314, and H315 (Figure 2I). Taken together, the interactions of BTAC observed in the structures of the three complexes BcBOT2-BTAC-PPi-Mg<sup>2+</sup>, DbPROS-BTAC-Mg<sup>2+</sup>, and CLM1-BTAC-PPi-Mg<sup>2+</sup> help define key active site residues, facilitating the QM/MM analyses toward their catalytic mechanisms.

**2.2. Catalytic Mechanisms and Sesquiterpene Variant Generation Based on QM/MM MD Analyses and Site-Directed Mutagenesis.** Based on the structures of the three complex, QM/MM MD studies were performed for the interactions between the three STSs and key carbocation intermediates during the cyclization cascades toward 1–3. The



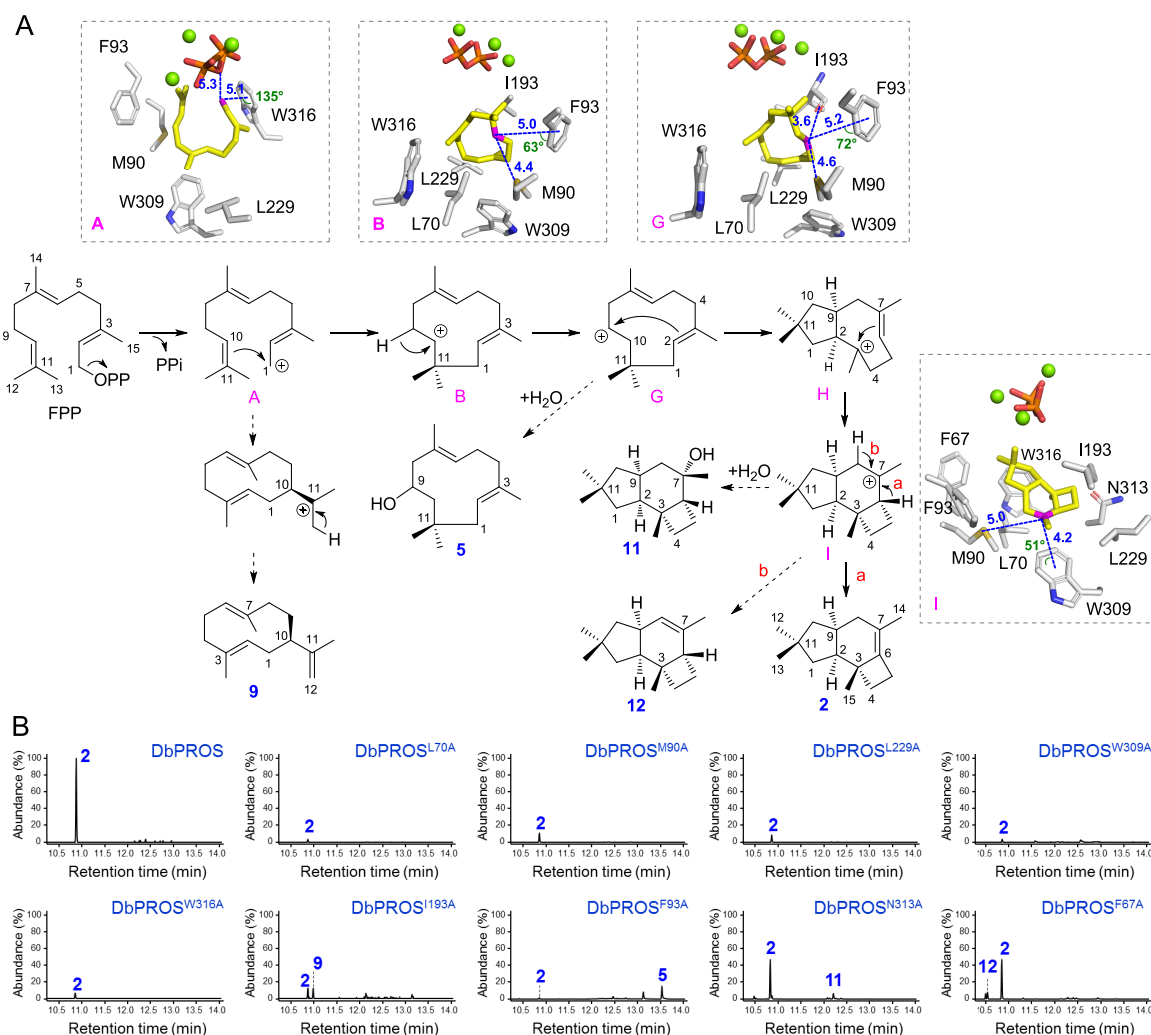
**Figure 3.** Biosynthesis of **1** by BcBOT2. (A) Cyclization of FPP to **1** (full arrows) and to shunt products obtained from enzyme variants (dashed arrows). QM/MM MD simulations of BcBOT2 with carbocation intermediates are shown in dashed boxes. Carbocation intermediates are shown in yellow sticks with the carbocation positions highlighted in magenta, and the interactions toward the carbocation positions are indicated with blue dashed lines. Interaction distances (Å) are shown in blue, and angles are shown in green. The state **D** is not an intermediate but a transition state with a very short life time as captured in the QM/MM MD trajectory (Figure S7A), consistent with the fact that no shunt products from **D** were observed. The electron density of W79 in **E** indicated by blue dots reveals its close proximity to the carbocation, which is important to maintain a hydrophobic environment. (B) GC–MS analyses of the products obtained from BcBOT2 and its enzyme variants.

intermediates in the biosynthesis of **1** and **2** have been proposed based on isotopic labeling studies before,<sup>32</sup> and the biosynthetic intermediates toward **3** were investigated through isotopic labeling experiments in this study (vide infra, Section 2.3). The enzyme-intermediate interactions resulting from QM/MM MD trajectories are shown in Figures 3–5 (cf. Supporting Information, section “Computational modeling” and Figure S5 for the detailed computational protocol and the definition of reaction coordinates). The kinetic and thermodynamic feasibility of the enzyme reactions is confirmed by the relative energy profiles from the QM/MM scans (Figure S8).

Following the results of these analyses, site-directed mutagenesis experiments were carried out. The compounds obtained from enzyme variants were made accessible using an

engineered *E. coli* strain that expressed multiple MVA pathway genes for production (Supporting Information and Figure S11). The mutagenesis results confirmed the roles of key active site residues and generated a series of sesquiterpenes (**4**–**20**) from different enzyme variants that were isolated from the large-scale fermentations. Their structures were elucidated based on NMR and MS analyses (Supporting Information).

The biosynthesis of **1** from FPP proceeds through a series of carbocation intermediates **A**–**F** (Figure 3A).<sup>32a</sup> QM/MM analyses of BcBOT2 with **A**–**F** showed that these intermediates interact with key active site residues. The residues W94 and F99 are involved in the interactions with most of the intermediates **A**–**F** (Figure 3A), and especially F99 forms  $\pi$ -cation interactions with carbocations in **B**, **C**, and **F**. Exchange



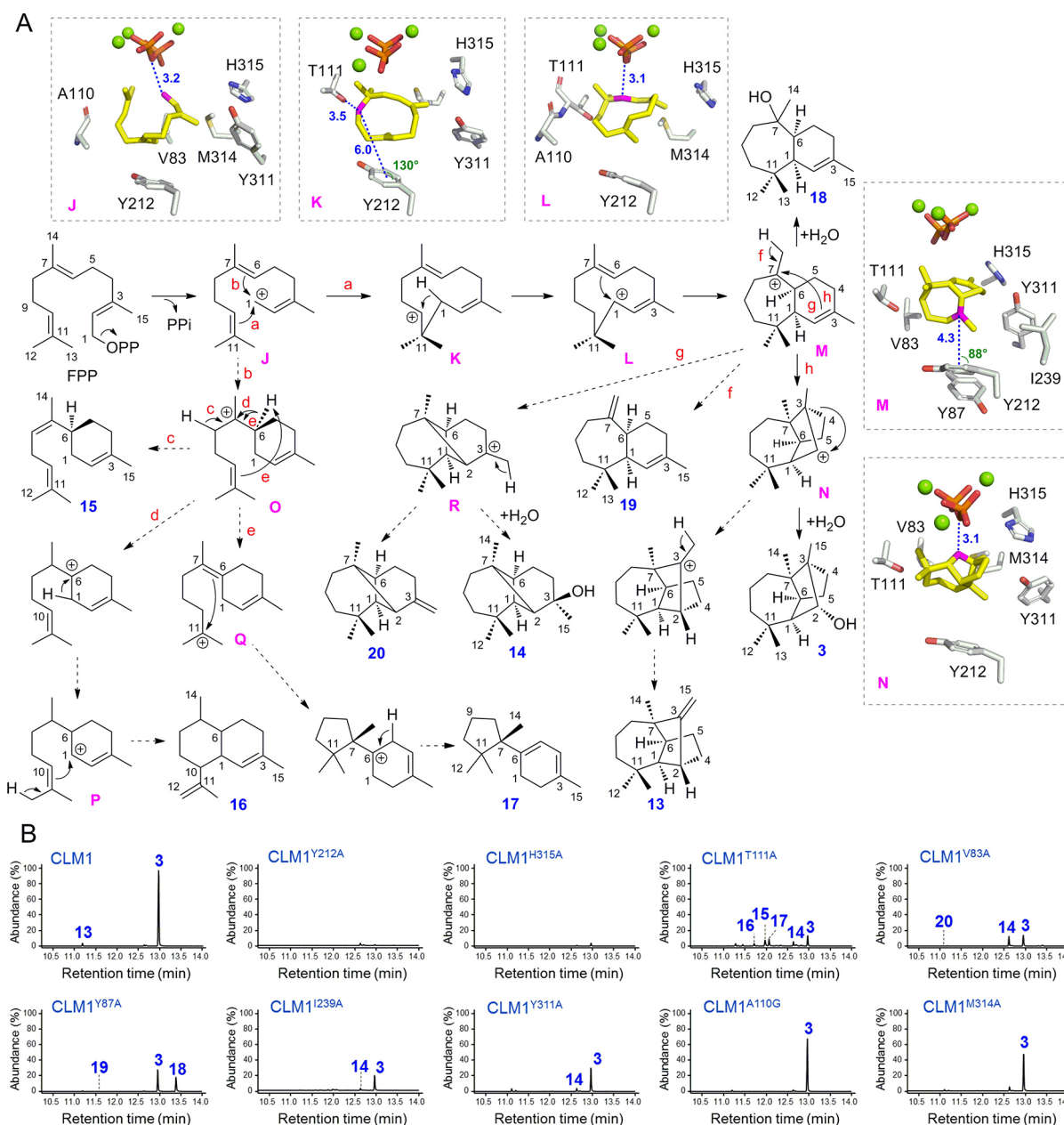
**Figure 4.** Biosynthesis of **2** by DbPROS. (A) Cyclization of FPP to **2** (full arrows) and to shunt products obtained from enzyme variants (dashed arrows). QM/MM MD simulations of DbPROS with carbocation intermediates are shown in dashed boxes. Carbocation intermediates are shown in yellow sticks with the carbocation positions highlighted in magenta, and the interactions toward the carbocation positions are indicated with blue dashed lines. Interaction distances (Å) are shown in blue, and angles are shown in green. The state **H** is not an intermediate but a transition state with a very short life time as captured in the QM/MM MD trajectory (Figure S7B), consistent with the fact that no shunt products from **H** were observed. (B) GC–MS analyses of the products obtained from DbPROS and its enzyme variants.

of these two residues to Ala (W94A or F99A) leads to strongly decreased yields of **1** (Figure 3B). Moreover, W94 provides strong hydrophobic interactions with **B** and **E** (Figure 3A), and widening of the active site cavity by the W94A exchange leads to the shunt products **4** and **5** from **B** and **6** from **E** (Figure 3B). The F99A variant has similar effects on the active site, generating the shunt products **7** and **8** from intermediates **C** and **E**, respectively. The F99Y variant abolishes the formation of shunt products **7** and **8**, but retains a low production of **1** (Figure S13), suggesting that Y99 can partially complement the role of F99 in shaping the active site and in stabilizing the intermediates through the  $\pi$ -cation interactions.

Also, V98, I204, and W328 each interact with more than one intermediate (Figure 3A), and the exchange of these residues to Ala leads to a substantially decreased production of **1** (Figure 3B). The I204V and W328F variants gave a higher production of **1** in comparison to I204A and W328A, respectively (Figure S13), suggesting that the exchanges to structurally similar residues have a lower impact on the enzyme-intermediate interaction. N325 interacts with **B** and **C**, especially forming a dipole–cation interaction with **C** (Figure

3A), and these intermediates may be less efficiently stabilized in the N325A variant, leading to the shunt product **9** from **A** (Figure 3B). The side chain of N325 points toward the carbocation in **F** and may assist in binding a water molecule for the hydroxylation to form **1** (Figure 3A). In agreement with this hypothesis, the N325A exchange results in the formation of the unhydroxylated product **10** (Figure 3B). Interestingly, for the W79A variant, the product spectrum is completely switched from **1** to **6** as a single product (Figure 3B). The exchange of V206 and V242 to Ala gives similar yields of **1** as for wild-type BcBOT2, suggesting that these two residues are not deeply involved in BcBOT2 catalysis.

The binding modes of the intermediates from the QM/MM MD are consistent with the stereoselective ring closure and alkyl and hydride shifts (Figure 3A): the *Re* face of C2 attacks the *Si* face of C10 in **B** to form the C2–C10 bond. In **C**, the C1–C2 bond is parallel to the empty orbital at C3 allowing for a smooth ring expansion to form the C1–C3 bond in **D**. In **E**, the C2 hydride shifts to the C7 carbocation from the *Si* face. In **F**, the water molecule attacks the C2 carbocation from the *Re* face to generate **1**, whereas the alkyl shift proceeds from the *Si*



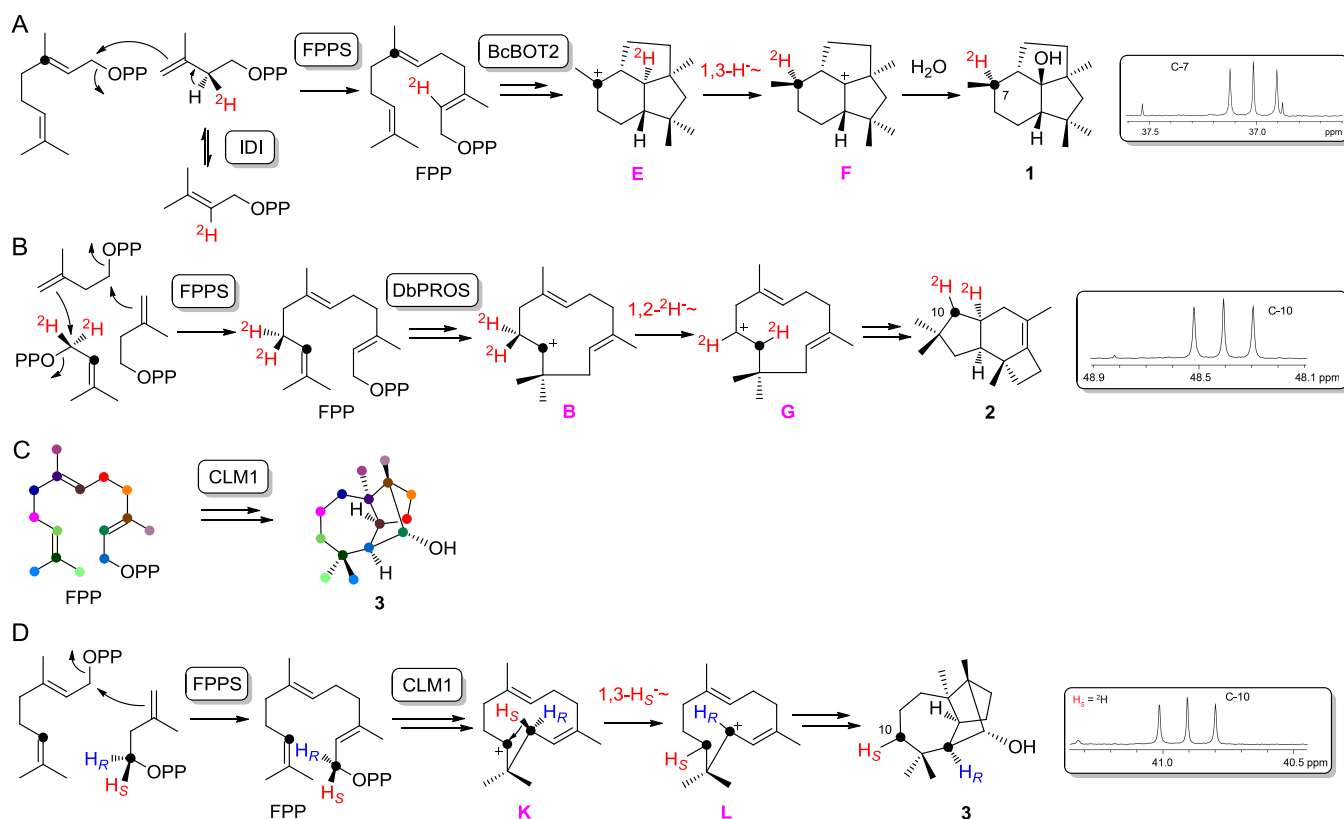
**Figure 5.** Biosynthesis of **3** by CLM1. (A) Cyclization of FPP to **3** (full arrows) and to shunt products obtained from enzyme variants (dashed arrows). QM/MM MD simulations of CLM1 with carbocation intermediates are shown in dashed boxes. Carbocation intermediates are shown in yellow sticks with the carbocation positions highlighted in magenta, and the interactions toward the carbocation positions are indicated with blue dashed lines. Interaction distances (Å) are shown in blue, and angles are shown in green. (B) GC–MS analyses of the products obtained from CLM1 and its enzyme variants.

face to the C2 carbocation, followed by methyl group migration and deprotonation to generate **10**. The relative energy profile for these steps shows low barriers and a notable free energy release (see Figure S8).

The biosynthesis of **2** from FPP by DbPROS proceeds through carbocations **A**, **B**, and **G–I** (Figure 4A).<sup>32b</sup> QM/MM analyses of DbPROS with these intermediates revealed their interactions with key residues. Specifically, L70, M90, L229, W309, and W316 are involved in the interactions with **A**, **B**, **G**, and **I** (Figure 4A), and exchange of these residues against Ala resulted in a significantly decreased production of **2** in all cases (Figure 4B). I193 interacts with most carbocation intermediates, and especially forms a dipole–cation interaction with **G** (Figure 4A). This interaction may be important for the

stabilization of **G**, and, consequently, the I193A exchange leads to a decreased production of **2** and concomitant formation of **9** from **A** (Figure 4B). This finding is well explained by the QM/MM MD analysis for DbPROS<sup>I193A</sup> with **A** showing that the distance from C1 to C10 becomes much shorter (2.6 Å) in the mutant than in the wild type (3.6 Å) (Figure S9), which opens the reaction channel for a 1,10-cyclization to **9**. F93 also interacts with most carbocation intermediates, and is especially engaged in the  $\pi$ -cation interactions with **B** and **G**. This residue is also ideally positioned to guide the migrating hydrogen through C–H $\cdots\pi$  stabilization, similar to the recently reported observations based on crystal structures of pentalenene synthase and its F76W variant.<sup>33a</sup> The F93A mutation leads to a strongly depleted production of **2** (Figure





**Figure 6.** Key isotopic labeling experiments on the biosynthesis of the wild-type enzyme products 1–3. (A) A selective deuteration and  $^{13}\text{C}$ -labeling revealed the 1,3-hydride shift in the biosynthesis of 1 by a triplet signal for C-7 in the  $^{13}\text{C}$ -NMR. (B) An analogous strategy was used to study the 1,2-hydride shift in the biosynthesis of 2. (C) Single  $^{13}\text{C}$ -labeling experiments show the origin of each carbon from FPP in 3 (also see Figure S22). Colored dots represent  $^{13}\text{C}$ -labeled carbons. (D) A stereoselective labeling in conjunction with  $^{13}\text{C}$ -labeling revealed the stereochemical course of the 1,3-hydride shift in the biosynthesis of 3.

4B), and also this exchange may facilitate the access of a water molecule to attack G (Figure 4B), thereby generating the hydroxylated product 5 (Figure 4A). Interestingly, the F93Y exchange retains activity for the production of 2 and shows no formation of 5 (Figure S13), demonstrating that Tyr can functionally substitute Phe at this position. The residues N313 and F67 interact with I, but not directly with its cationic center, and the substitution of either of these two residues by Ala leads to a change in the active site cavity with the consequence of a decrease of 2 and the additional formation of small amounts of 11 and 12 from I (Figure 4). In contrast, the exchange against a bulkier residue as in the F67L variant results in a retained high production of 2 with no formation of 12 (Figure S13), demonstrating that a large hydrophobic residue, but not necessarily the aromatic Phe, is required at this position for the efficient biosynthesis of 2. The cation in I interacts directly through the  $\pi$ -cation interaction with W309 and dipole–cation interaction with M90, which may lead to a significant decrease of the relative energy of I for the efficient conversion into 2 (Figure S8).

The biosynthesis of 3 by CLM1 from FPP proceeds through a series of carbocation intermediates J–N (Figure 5A). The GC–MS analysis of extracts from an engineered *E. coli* strain expressing wild-type CLM1 shows the production of 3 together with a minor amount of 13 (Figure 5B). Based on the QM/MM analyses, very important active site residues are Y212 and H315 that strongly interact with all intermediates J–N (Figure 5A). Accordingly, the production of 3 is completely abolished for the enzyme variants Y212A and H315A (Figure

5B). The aromatic residue Y212 is specifically devoted to the  $\pi$ -cation interactions with K and M, and both exchanges Y212A and Y212L fail to stabilize these cations with the consequence of an abolished production of 3 (Figure S13). In contrast, the Y212F variant can functionally substitute and shows a retained production of 3 (Figure S13). Also, T111 is involved in the interactions with most intermediates (K–N, Figure 5A), but in this case, the T111A exchange, besides a decreased production of 3, leads to the production of several other sesquiterpenes (14–17, Figure 5B). The T111 side-chain hydroxy group is engaged in a dipole–cation interaction with K (Figure 5A), and the missing functional group in the T111A variety may prevent the formation of this intermediate, leading to the shunt products 15–17 from J (Figure 5A). This view is further supported by the finding that the T111S variant does not form the shunt products 16 and 17 (Figure S13), suggesting that the hydroxy group of Ser can partially complement the dipole–cation interaction with K. Furthermore, the residues V83, Y87, I239, and Y311 interact with M (Figure 5A), and their substitution with Ala may widen or reshape the active site cavity to generate the shunt products 14 and 18–20 from M (Figure 5B). Notably, Y87 is located below Y212 and may assist this residue to form the  $\pi$ -cation interaction with M. When Y87 is exchanged to Phe or Leu, the production of 3 is retained and the shunt products 18 and 19 are not observed (Figure S13), confirming that a sterically demanding residue is sufficient at this position to push Y212 toward M for the  $\pi$ -cation interaction. The enzyme variants A110G and M314A show a slightly decreased production of 3

in comparison to wild-type CLM1 (Figure 5B), suggesting that these two residues are not deeply involved in CLM1 catalysis. The binding modes of the intermediates from the QM/MM MD are consistent with the stereoselective ring closure: C6 attacks the C1 carbocation from the *Si* face in L to form the C1–C6 bond, and C3 attacks the C7 carbocation from the *Re* face in M to form the C3–C7 bond. This is further proven by the gradually decreasing relative energy from state J to state N and the low reaction barriers for these transformations (Figure S8).

The biosynthetic pathways toward the shunt products 13–20 were also proposed. In the intermediate J, C6 attacks the C1 carbocation to generate the branching point O. Its deprotonation at C8 forms 15 (path c, Figure 5A), while two successive hydride shifts lead to P (path d), which forms 16 upon ring closure and deprotonation. The deprotonation at C6 and reprotonation at C10 in O result in the intermediate Q (path e). Alternatively, as proposed by Hong and Tantillo, a direct intramolecular proton transfer may apply.<sup>33b</sup> Downstream ring closure and deprotonation form 17. Another important bifurcation point is the intermediate M that undergoes hydroxylation to 18 or deprotonation to 19 (path f), respectively. M can also undergo ring closure between C2 and C7 to generate the intermediate R (path g), which is the direct precursor to 14 and 20, respectively. Finally, the intermediate N can undergo a Wagner–Meerwein rearrangement followed by deprotonation to produce 13 (Figure 5A).

**2.3. Isotopic Labeling Studies toward Key Steps in the Production of 1–3 and Sesquiterpene Variants.** In order to understand the cyclization mechanisms by the three wild-type STSs and their enzyme variants more deeply, isotopic labeling experiments were performed. These experiments addressed all hydride shifts and stereochemical problems, which are associated with deprotonation steps and the migrations of one of the geminal methyl groups of FPP.

In the biosynthesis of 1 by BcBOT2, several cyclization reactions occur. A key step is the proposed 1,3-hydride shift from E to F, which was investigated using (3-<sup>13</sup>C)GPP and (2-<sup>2</sup>H)DMAPP<sup>34</sup> that were converted with *E. coli* isopentenyl diphosphate isomerase (IDI)<sup>35</sup> and *Streptomyces coelicolor* FPP synthase (FPPS)<sup>36</sup> into (7-<sup>13</sup>C,2-<sup>2</sup>H)FPP, followed by the cyclization with BcBOT2. The proposed 1,3-hydride shift was evident from a slightly upfield-shifted triplet resulting from <sup>13</sup>C-<sup>2</sup>H spin coupling (Figure 6A and Figure S14). Instead, in the biosynthesis of 8 by BcBOT2<sup>F99A</sup> from E (Figure 3A), a 1,2-hydride migration takes place that was investigated analogously using (3-<sup>13</sup>C,2-<sup>2</sup>H)GPP<sup>37</sup> in conjunction with IPP and FPPS (Figure S15). For compound 4 the stereoselectivity of the deprotonation from B was studied using (*R*)- and (*S*)-(1-<sup>13</sup>C,1-<sup>2</sup>H)IPP<sup>38</sup> that were converted with IDI, FPPS, and BcBOT2<sup>W94A</sup>, showing selective abstraction of the *pro-S* proton (Figure S16). For compound 5 produced by BcBOT2<sup>W94A</sup>, the 1,2-hydride shift in B was investigated in two experiments: first, the incubation of (2-<sup>13</sup>C,1,1-<sup>2</sup>H<sub>2</sub>)DMAPP<sup>36</sup> and IPP with FPPS and BcBOT2<sup>W94A</sup> revealed the migration of deuterium from C-9 to C-10 (Figure S17); in a second experiment, the conversion of (*R*)- and (*S*)-(1-<sup>13</sup>C,1-<sup>2</sup>H)IPP with IDI, FPPS, and BcBOT2<sup>W94A</sup> demonstrated that selectively the *pro-S* hydride shifts from C-9 to C-10 (Figure S18). In the biosynthesis of 10 by BcBOT2<sup>N325A</sup>, the second last step involves a migration of one of the two geminal methyl groups (Figure 3A). Through enzymatic conversion of

(12-<sup>13</sup>C)FPP and (13-<sup>13</sup>C)FPP,<sup>39</sup> the migrating methyl group was identified as C-12 (Figure S19).

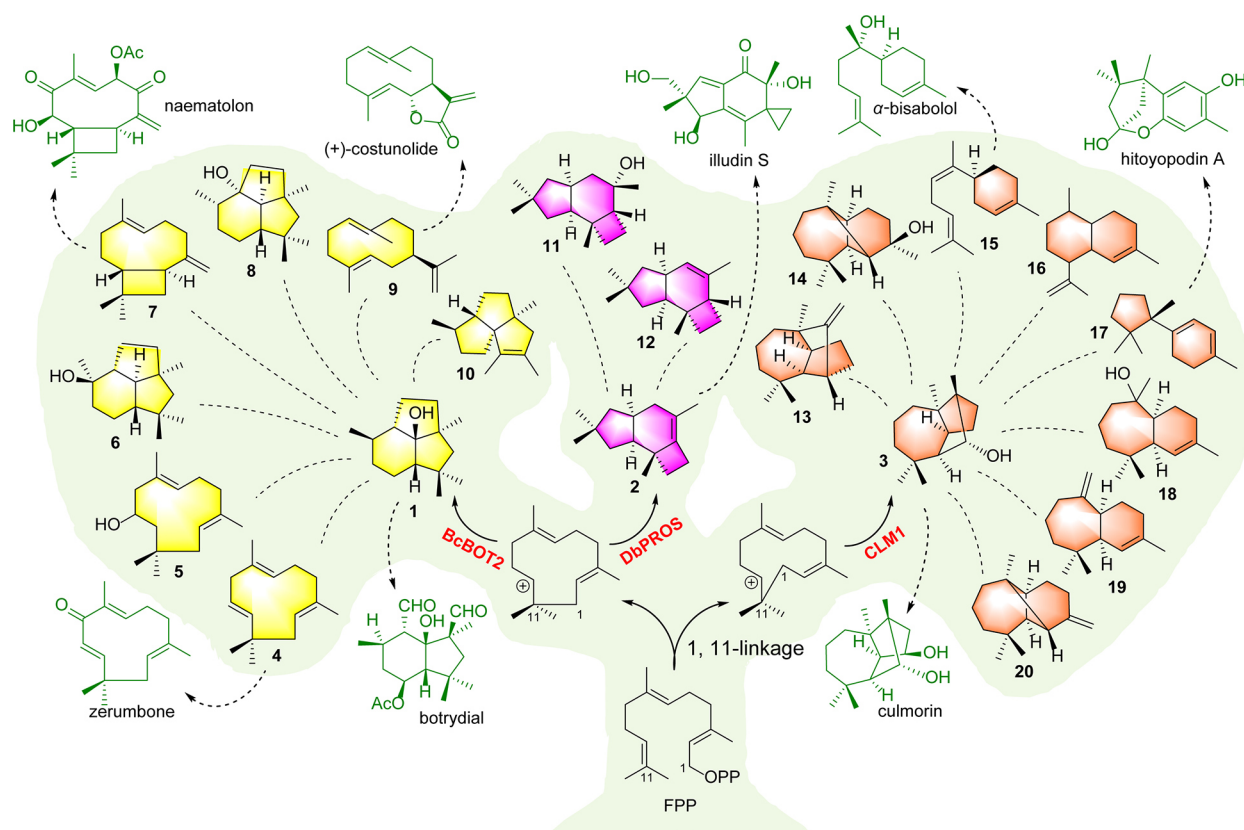
Along the cyclization cascade from FPP to 2 by DbPROS, a 1,2-hydride shift from B to G (Figure 4A) is proposed that is similar to the step from B toward 5 by BcBOT2<sup>W94A</sup> (Figure 3A). Also in this case, two experiments were performed to follow this step: the incubation of (2-<sup>13</sup>C,1,1-<sup>2</sup>H<sub>2</sub>)DMAPP and IPP with FPPS and DbPROS confirmed the migration of a hydride from C-9 to C-10 (Figure 6B and Figure S20), and subsequent conversion of (*R*)- and (*S*)-(1-<sup>13</sup>C,1-<sup>2</sup>H)IPP with IDI, FPPS, and DbPROS revealed the selective shift of the *pro-S* hydride from C-9 to C-10 (Figure S21). This is the same stereochemical course as observed for the formation of 5 by BcBOT2<sup>W94A</sup>.

Regarding the cyclization of FPP to 3 by CLM1, each of the 15 carbons was followed by incubation of the 15 isotopomers of (<sup>13</sup>C)FPP<sup>39</sup> (Figure 6C and Figure S22). The results were fully consistent with the proposed cyclization mechanism in Figure 5A. Subsequently, the unique hydride shift from K to L was investigated by incubation of (6-<sup>13</sup>C)GPP and (*R*)- and (*S*)-(1-<sup>13</sup>C, 1-<sup>2</sup>H)IPP with FPPS and CLM1, uncovering the specific migration of the 1-*pro-S* hydride to C-10 (Figure 6D and Figure S23). The formation of 16 by CLM1<sup>T111A</sup> involves a sequence of two 1,2-hydride shifts, as was evident from the enzymatic conversion of (3-<sup>13</sup>C, 2-<sup>2</sup>H)GPP and IPP with FPPS and CLM1<sup>T111A</sup> showing the first of these hydride shifts (Figure S24). However, because of the low production of 16, the triplet signal observed in the <sup>13</sup>C-NMR of the product obtained in this experiment was very weak and therefore this assignment remains tentative.

The absolute configurations of 1–3 were assigned using a stereoselective deuteration approach (Figures S25–S27), pointing to the absolute configurations as shown in Figures 3–5. While the same absolute configurations for 1 and 3 were previously secured through chemical correlations,<sup>32a,40</sup> the absolute configuration of 2 seems to be only assigned based on the assumed biosynthetic relationship to illudin S and related compounds of known absolute configuration<sup>41</sup> but has never been firmly established.

### 3. DISCUSSION

The three STSs BcBOT2, DbPROS, and CLM1 all catalyze an initial 1,11-cyclization of FPP. Herein, BcBOT2 and DbPROS proceed through the (*E,E*)-humulyl cation (B), while CLM1 generates the (*Z,E*)-humulyl cation (K). These differences are also supported by the QM/MM MD analyses of intermediates A (in BcBOT2 and DbPROS) and J (in CLM1), in which the C-1 carbocations are close to C-11 in all three cases and well stabilized by the diphosphate (Figures 3A, 4A, and 5A). Interestingly, although BcBOT2 and DbPROS share the same early intermediates A and B, the two enzymes next catalyze reactions to distinct directions: BcBOT2 converts B into C whose C-3 carbocation is stabilized by N325 and F99 (Figure 3A), whereas DbPROS converts B into G whose C-9 carbocation is stabilized by F93 and I193 (Figure 4A). The many distinct cascade cyclization steps toward 1–3 in the three pathways were evidenced by isotopic labeling experiments: the key hydride and methyl migrations in the formation of 1–3 and shunt pathway products were thoroughly addressed, and the cyclization of FPP to 3 by CLM1 was confirmed in detail by the cyclization of the 15 isotopomers of (<sup>13</sup>C)FPP. Thus, the characterization of the three STSs' catalytic mechanisms by our crystal structures, QM/MM



**Figure 7.** Biosynthesis of 1–3 by BcBOT2, DbPROS, and CLM1 through the 1,11-linkage of FPP, and the shunt products 4–20 generated from the three STS enzyme variants. Bioactive sesquiterpenoids derived from sesquiterpenes are shown in green. Sesquiterpenes 1, 2, and 3 and their congeners obtained from enzyme variants, linked with dashed lines, are shown in yellow, magenta, and orange, respectively.

analyses, and isotopic labeling experiments may allow for future rational STS engineering and de novo design of catalysts for polycyclic sesquiterpene generation.

The mutagenesis of key active site residues not only confirmed their roles in the cyclization process but also generated structurally diverse sesquiterpene variants 4–20 (Figure 7). Most of these compounds including 4–10, 13–15, and 17–20, although generated from the mutations of three fungal STSs in this study, are important essential oil components from various plants. Several of these sesquiterpenes are of great value for their biological activity and fragrance properties. For example,  $\alpha$ -humulene (4) possesses antibacterial and antibiofilm activities against *Bacteroides fragilis*,<sup>42</sup> while  $\beta$ -caryophyllene (7) selectively binds to the G-protein-coupled receptor (GPCR) CB2 and shows broad biological activities.<sup>43</sup> Moreover, oral administration of 7 reduces the glucose level in blood and increases plasma insulin, providing an important medicinal benefit to oral preparations.<sup>44</sup> Compound 11 has not been isolated from natural sources so far but is known as a synthetic compound with only proton NMR data reported.<sup>45</sup> Here, we first provide full NMR data and a complete structure elucidation. Compound 12 was recently reported as a product of a STS from social amoebae,<sup>36</sup> and our fungal-originated DbPROS<sup>F67A</sup> provides an alternative choice for the preparation of 12.

Compounds 1–3 are important precursors of bioactive molecules (i.e., botrydial, illudin S, and culmorin, Figures 1B and 7), and some sesquiterpenes obtained from enzyme variants in this study constitute the carbon skeletons of additional important bioactive sesquiterpenoids (Figure 7).

Compound 4 can be oxidized to zerumbone that is found in the rhizomes of shampoo ginger and possesses potent immunomodulatory and antitumor activities.<sup>46</sup> Compound 7 can be oxidized and acetylated to the antibacterial sesquiterpenoid naematolon (Figure 7).<sup>47</sup> Compound 9 is the precursor of (+)-costunolide, which has been investigated for a wide range of biological activities that are owing to its modulation of various intracellular signaling pathways by targeting intracellular kinases and redox-regulated transcription factors.<sup>48</sup>  $\alpha$ -Bisabolol derived from compound 15 has a strong time- and dose-dependent cytotoxic effect on human and rat glioma cells.<sup>49</sup> Hitoyopodin A derived from compound 17 shows growth inhibition against human leukemia cells HL-60 and the malarial parasite *Plasmodium falciparum* 3D7.<sup>50</sup> Taken together, our generation of sesquiterpenes 1–20 from the wild type and mutants of BcBOT2, DbPROS, and CLM1 may facilitate the preparation of 1–20 and their derivatives in pharmaceutical agent development.

The BcBOT2<sup>W79A</sup> variant forms 6 as a single product (Figure 3B), which arises by premature water quenching and points to a less tight control of the intermediate E in comparison to the wild-type BcBOT2, for which E undergoes another hydride shift prior to the water quenching (Figure 3A). Furthermore, the expanded cavity of the BcBOT2<sup>W79A</sup> variant can accommodate three water molecules that form a stable hydrogen-bonding network with nearby residues after long-time classical MD simulations (Figure S10). The hydrogen bonding network of the water chain was also well preserved in QM/MM MD simulations. As a result, the nearest

water molecule adjacent to the cation in **E** kept a specific orientation for the stereospecific quenching to generate **6**.

Several enzyme variants with single residue exchanges including BcBOT2<sup>W94A</sup>, BcBOT2<sup>F99A</sup>, BcBOT2<sup>N325A</sup>, DbPROS<sup>F67A</sup>, DbPROS<sup>N313A</sup>, CLM1<sup>T111A</sup>, CLM1<sup>V83A</sup>, and CLM1<sup>Y87A</sup> turned the highly selective wild-type biocatalysts into promiscuous enzymes, demonstrating that the difference between selectivity and promiscuity in terpene synthase-catalysis is a ridge walk. This may also be important in evolution: sometimes a high-fidelity enzyme is beneficial, if the target compound of a pathway is highly active and provides an ecological advantage to the producing organism, but the evolution of new enzyme functions may proceed through multiproduct stages that allow for the “testing” of several new compounds at once. The catalytic promiscuity and fidelity of the STSs TEAS and ATAS may give an example:<sup>51</sup> the high-fidelity catalysis of ATAS (that produces aristolochene as a single product) is contrasted by the multiproduct aristolochene synthase TEAS. For ATAS, the aromatic residues in the active site provide the required steric hindrance and  $\pi$ -cation interactions for tight substrate control to guarantee high enzyme fidelity. In contrast, the active site of TEAS only loosely shapes the substrate and has less control of the cationic intermediates, leading to multiple side products. We deduced that our mutations mentioned above may reduce the steric hindrance or abolish the  $\pi$ -cation interactions in the active site, which looses the active site cavity and increases the free energy barrier during the intermediate transformations, as observed for TEAS, and then enhancing the chances to form shunt pathway products. In every case, the thin line between enzyme selectivity and promiscuity will be one of the major challenges for the rational design of selective terpene synthases with new functions.

Another interesting phenomenon is that the wild-type enzyme product (**1**, **2**, or **3**) is almost always detected in the investigated enzyme variants reported here. This phenomenon was also observed in previous studies in which it was suggested that the terpene cyclization cascade may be driven by the intrinsic reactivity of the carbocation intermediates, during which the enzyme does not need to intervene at all stages.<sup>8,52</sup> Also, supramolecular capsules can mimic terpene synthases to generate cyclized terpenes.<sup>53</sup> For catalysis by the diterpene synthase CotB2, the free energy profiles of the reaction cascade in the presence or absence of CotB2 are similar, which is probably attributed to the balanced interaction between the enzyme and carbocationic intermediates.<sup>54</sup> In our own previous work, we proposed that three factors determine catalytic promiscuity versus specificity: the substrate folding mode, the presence of key active site residues, and the flexibility of intermediates that is mainly a consequence of the available space in the active site.<sup>55,56</sup> Thus, the production of **1**, **2**, or **3** in the variants of BcBOT2, DbPROS, and CLM1, respectively, and their promiscuity observed in several cases provide new insights that may provoke future investigations of these phenomena.

#### 4. CONCLUSIONS

In summary, we determined the crystal structures of three STSs BcBOT2, DbPROS, and CLM1 in the biosynthesis of tricyclic sesquiterpenes presilphiperfolan-8 $\beta$ -ol (**1**),  $\Delta^6$ -protoilludene (**2**), and longiborneol (**3**), respectively. QM/MM studies, in combination with site-directed mutagenesis, revealed distinct cascade cyclization mechanisms by the three

enzymes toward **1–3** biosynthesis. Exchange of key residues in the active sites of the three enzymes generated 17 shunt products (**4–20**), expanding the chemical space that can be reached enzymatically. Isotopic labeling studies revealed the key hydride and methyl migration events during the formation of **1–3** and shunt pathway products. Our study not only provides a better understanding of tricyclic sesquiterpene biosynthesis by STSs but also facilitates the synthetic biology research of **1–20** in the future, given that many of them are of great values to be pharmaceutically active and perfumery molecules.

#### ■ ASSOCIATED CONTENT

##### SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/jacs.3c00278>.

All experiment details including general experimental procedures, gene expression and protein purification, crystallization and structural elucidation, QM/MM MD simulation methods, construction of engineered *E. coli* strains, site-directed mutagenesis and GC-MS analysis, large-scale fermentation and sesquiterpene isolation, structural elucidation of sesquiterpenes, synthesis of (3-<sup>13</sup>C)GPP, primers/plasmids/strains used in this study, data collection and refinement statistics of crystal structures, detailed labeling experiments, overall structures of STSs and comparison, NMR data of compounds **1–20**, and the <sup>1</sup>H and <sup>13</sup>C NMR spectra of compounds **1–20** (PDF)

Carbocation intermediates (ZIP)


#### ■ AUTHOR INFORMATION

##### Corresponding Authors

**Ruibao Wu** – School of Pharmaceutical Sciences, Sun Yat-sen University, Guangzhou 510006, China; Email: [wurb3@mail.sysu.edu.cn](mailto:wurb3@mail.sysu.edu.cn)

**Jeroen S. Dickschat** – Kekulé-Institute for Organic Chemistry and Biochemistry, University of Bonn, 53121 Bonn, Germany;  [orcid.org/0000-0002-0102-0631](https://orcid.org/0000-0002-0102-0631); Email: [dickschat@uni-bonn.de](mailto:dickschat@uni-bonn.de)

**Donghui Yang** – State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University, Beijing 100191, China; Email: [ydhui@bjmu.edu.cn](mailto:ydhui@bjmu.edu.cn)

**Ming Ma** – State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University, Beijing 100191, China;  [orcid.org/0000-0001-8311-3892](https://orcid.org/0000-0001-8311-3892); Email: [mma@bjmu.edu.cn](mailto:mma@bjmu.edu.cn)

##### Authors

**Tingting Lou** – State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University, Beijing 100191, China

**Annan Li** – State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University, Beijing 100191, China

**Houchao Xu** – Kekulé-Institute for Organic Chemistry and Biochemistry, University of Bonn, 53121 Bonn, Germany

**Jingfeng Pan** – School of Pharmaceutical Sciences, Sun Yat-sen University, Guangzhou 510006, China

Baiying Xing – State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University, Beijing 100191, China

Complete contact information is available at:  
<https://pubs.acs.org/10.1021/jacs.3c00278>

### Author Contributions

<sup>#</sup>T.L., A.L., H.X., and J.P. contributed equally.

### Notes

The authors declare no competing financial interest.

### ACKNOWLEDGMENTS

We thank Hongli Jia and Fuling Yin from the State Key Laboratory of Natural and Biomimetic Drugs (Peking University) for X-ray diffraction tests. We thank the staffs from BL02U1/BL17B/BL10U2/BL19U1 beamlines of the National Facility for Protein Science in Shanghai (NFPS) at Shanghai Synchrotron Radiation Facility, for assistance during X-ray diffraction data collection. This research was supported in part by the National Natural Science Foundation of China (grant numbers 22077007, 81991525, 82273829, 81573326, 22107007), the Key Project at Central Government Level: the Ability Establishment of Sustainable Use for Valuable Chinese Medicine Resources (2060302-2201-17), and the German Research Foundation DFG (DI1536/7-2).

### REFERENCES

- (1) Chen, F.; Tholl, D.; Bohlmann, J.; Pichersky, E. The Family of Terpene Synthases in Plants: a Mid-Size Family of Genes for Specialized Metabolism That Is Highly Diversified throughout the Kingdom. *Plant J.* **2011**, *66*, 212–229.
- (2) Quin, M. B.; Flynn, C. M.; Schmidt-Dannert, C. Traversing the Fungal Terpenome. *Nat. Prod. Rep.* **2014**, *31*, 1449–1473.
- (3) Rudolf, J. D.; Alsup, T. A.; Xu, B.; Li, Z. Bacterial Terpenome. *Nat. Prod. Rep.* **2021**, *38*, 905–980.
- (4) Klapschinski, T. A.; Rabe, P.; Dickschat, J. S. Pristinol, a Sesquiterpene Alcohol with an Unusual Skeleton from *Streptomyces Pristinaespiralis*. *Angew. Chem., Int. Ed. Engl.* **2016**, *55*, 10141–10144.
- (5) Cane, D. E. Enzymic Formation of Sesquiterpenes. *Chem. Rev.* **1990**, *90*, 1089–1103.
- (6) Minami, A.; Ozaki, T.; Liu, C.; Oikawa, H. Cyclopentane-Forming Di/Sesterterpene Synthases: Widely Distributed Enzymes in Bacteria, Fungi, and Plants. *Nat. Prod. Rep.* **2018**, *35*, 1330–1346.
- (7) Tao, H.; Lauterbach, L.; Bian, G.; Chen, R.; Hou, A.; Mori, T.; Cheng, S.; Hu, B.; Lu, L.; Mu, X.; Li, M.; Adachi, N.; Kawasaki, M.; Moriya, T.; Senda, T.; Wang, X.; Deng, Z.; Abe, I.; Dickschat, J. S.; Liu, T. Discovery of Non-Squalene Triterpenes. *Nature* **2022**, *606*, 414–419.
- (8) Christianson, D. W. Structural and Chemical Biology of Terpenoid Cyclases. *Chem. Rev.* **2017**, *117*, 11570–11648.
- (9) Xu, H.; Dickschat, J. S. Mechanistic Investigations on Microbial Type I Terpene Synthases through Site-Directed Mutagenesis. *Synthesis* **2022**, *54*, 1551–1565.
- (10) Dickschat, J. S. Bacterial Diterpene Biosynthesis. *Angew. Chem., Int. Ed. Engl.* **2019**, *58*, 15964–15976.
- (11) Hare, S. R.; Tantillo, D. J. Dynamic Behavior of Rearranging Carbocations—Implications for Terpene Biosynthesis. *Beilstein J. Org. Chem.* **2016**, *12*, 377–390.
- (12) Lesburg, C. A.; Zhai, G.; Cane, D. E.; Christianson, D. W. Crystal Structure of Pentalene Synthase: Mechanistic Insights on Terpenoid Cyclization Reactions in Biology. *Science* **1997**, *277*, 1820–1824.
- (13) Aaron, J. A.; Lin, X.; Cane, D. E.; Christianson, D. W. Structure of Epi-Isozaene Synthase from *Streptomyces Coelicolor* A3(2), a Platform for New Terpenoid Cyclization Templates. *Biochemistry* **2010**, *49*, 1787–1797.
- (14) Blank, P. N.; Pemberton, T. A.; Chow, J. Y.; Poulter, C. D.; Christianson, D. W. Crystal Structure of Cucumene Synthase, a Terpenoid Cyclase That Generates a Linear Triquinane Sesquiterpene. *Biochemistry* **2018**, *57*, 6326–6335.
- (15) Chen, C. C.; Malwal, S. R.; Han, X.; Liu, W.; Ma, L.; Zhai, C.; Dai, L.; Huang, J. W.; Shillo, A.; Desai, J.; Ma, X.; Zhang, Y.; Guo, R. T.; Oldfield, E. Terpene Cyclases and Prenyltransferases: Structures and Mechanisms of Action. *ACS Catal.* **2021**, *11*, 290–303.
- (16) Whitehead, J. N.; Leferink, N. G. H.; Komati Reddy, G.; Levy, C. W.; Hay, S.; Takano, E.; Scrutton, N. S. How a 10-*epi*-Cubebol Synthase Avoids Premature Reaction Quenching to Form a Tricyclic Product at High Purity. *ACS Catal.* **2022**, *12*, 12123–12131.
- (17) Pinedo, C.; Wang, C. M.; Pradier, J. M.; Dalmais, B.; Choquer, M.; Le Pâcheur, P.; Morgant, G.; Collado, I. G.; Cane, D. E.; Viaud, M. Sesquiterpene Synthase from the Botrydial Biosynthetic Gene Cluster of the Phytopathogen *Botrytis Cinerea*. *ACS Chem. Biol.* **2008**, *3*, 791–801.
- (18) Collado, I. G.; Sánchez, A. J.; Hanson, J. R. Fungal Terpene Metabolites: Biosynthetic Relationships and the Control of the Phytopathogenic Fungus *Botrytis Cinerea*. *Nat. Prod. Rep.* **2007**, *24*, 674–686.
- (19) Williamson, B.; Tudzynski, B.; Tudzynski, P.; van Kan, J. A. *Botrytis Cinerea*: the Cause of Grey Mould Disease. *Mol. Plant Pathol.* **2007**, *8*, 561–580.
- (20) Choquer, M.; Fournier, E.; Kunz, C.; Levis, C.; Pradier, J. M.; Simon, A.; Viaud, M. *Botrytis Cinerea* Virulence Factors: New Insights into a Necrotrophic and Polyphageous Pathogen. *FEMS Microbiol. Lett.* **2007**, *277*, 1–10.
- (21) Siewers, V.; Viaud, M.; Jimenez-Teja, D.; Collado, I. G.; Gronover, C. S.; Pradier, J. M.; Tudzynski, B.; Tudzynski, P. Functional Analysis of the Cytochrome P450 Monooxygenase Gene *bcbot1* of *Botrytis Cinerea* Indicates That Botrydial Is a Strain-Specific Virulence Factor. *Mol. Plant-Microbe Interact.* **2005**, *18*, 602–612.
- (22) Moraga, J.; Dalmais, B.; Izquierdo-Bueno, I.; Aleu, J.; Hanson, J. R.; Hernández-Galán, R.; Viaud, M.; Collado, I. G. Genetic and Molecular Basis of Botrydial Biosynthesis: Connecting Cytochrome P450-Encoding Genes to Biosynthetic Intermediates. *ACS Chem. Biol.* **2016**, *11*, 2838–2846.
- (23) Porquier, A.; Morgant, G.; Moraga, J.; Dalmais, B.; Luyten, I.; Simon, A.; Pradier, J. M.; Amselem, J.; Collado, I. G.; Viaud, M. The Botrydial Biosynthetic Gene Cluster of *Botrytis Cinerea* Displays a Bipartite Genomic Structure and Is Positively Regulated by the Putative Zn(II)<sub>2</sub>Cys<sub>6</sub> Transcription Factor BcBot6. *Fungal Genet. Biol.* **2016**, *96*, 33–46.
- (24) Zhang, C.; Chen, X.; Orban, A.; Shukul, S.; Birk, F.; Too, H. P.; Rühl, M. *Agrocybe Aegerita* Serves as a Gateway for Identifying Sesquiterpene Biosynthetic Enzymes in Higher Fungi. *ACS Chem. Biol.* **2020**, *15*, 1268–1277.
- (25) (a) Anchel, M.; Hervey, A.; Robbins, W. J. Antibiotic Substances from Basidiomycetes. VII. Clitocybe Illudens. *Proc. Natl. Acad. Sci. U. S. A.* **1950**, *36*, 300–305. (b) Schobert, R.; Knauer, S.; Seibt, S.; Biersack, B. Anticancer Active Illudins: Recent Developments of a Potent Alkylating Compound Class. *Curr. Med. Chem.* **2011**, *18*, 790–807.
- (26) McMorris, T. C.; Kelner, M. J.; Wang, W.; Yu, J.; Estes, L. A.; Taetle, R. (Hydroxymethyl)acylfulvene: an Illudin Derivative with Superior Antitumor Properties. *J. Nat. Prod.* **1996**, *59*, 896–899.
- (27) (a) McCormick, S. P.; Alexander, N. J.; Harris, L. J. CLM1 of *Fusarium Graminearum* Encodes a Longiborneol Synthase Required for Culmorin Production. *Appl. Environ. Microbiol.* **2010**, *76*, 136–141. (b) Strongman, D. B.; Miller, J. D.; Calhoun, L.; Findlay, J. A.; Whitney, N. J. The Biochemical Basis for Interference Competition among Some Lignicolous Marine Fungi. *Bot. Mar.* **1987**, *30*, 21–26.
- (28) Lusi, R. F.; Sennari, G.; Sarpong, R. Total Synthesis of Nine Longiborneol Sesquiterpenoids Using a Functionalized Camphor Strategy. *Nat. Chem.* **2022**, *14*, 450–456.
- (29) Mirdita, M.; Schütze, K.; Moriwaki, Y.; Heo, L.; Ovchinnikov, S.; Steinegger, M. ColabFold: Making Protein Folding Accessible to All. *Nat. Methods* **2022**, *19*, 679–682.

- (30) Rabe, P.; Schmitz, T.; Dickschat, J. S. Mechanistic Investigations on Six Bacterial Terpene Cyclases. *Beilstein J. Org. Chem.* **2016**, *12*, 1839–1850.
- (31) Baer, P.; Rabe, P.; Fischer, K.; Citron, C. A.; Klapschinski, T. A.; Groll, M.; Dickschat, J. S. Induced-Fit Mechanism in Class I Terpene Cyclases. *Angew. Chem., Int. Ed. Engl.* **2014**, *53*, 7652–7656.
- (32) (a) Wang, C. M.; Hopson, R.; Lin, X.; Cane, D. E. Biosynthesis of the Sesquiterpene Botrydial in *Botrytis cinerea*. Mechanism and Stereochemistry of the Enzymatic Formation of Presilphiperfolan-8 $\beta$ -ol. *J. Am. Chem. Soc.* **2009**, *131*, 8360–8361. (b) Zu, L.; Xu, M.; Lodewyk, M. W.; Cane, D. E.; Peters, R. J.; Tantillo, D. J. Effect of Isotopically Sensitive Branching on Product Distribution for Pentalenene Synthase: Support for a Mechanism Predicted by Quantum Chemistry. *J. Am. Chem. Soc.* **2012**, *134*, 11369–11371.
- (33) (a) Matos, J. O.; Kumar, R. P.; Ma, A. C.; Patterson, M.; Krauss, I. J.; Oprian, D. D. Mechanism Underlying Anti-Markovnikov Addition in the Reaction of Pentalenene Synthase. *Biochemistry* **2020**, *59*, 3271–3283. (b) Hong, Y. J.; Tantillo, D. J. Which Is More Likely in Trichodiene Biosynthesis: Hydride or Proton Transfer? *Org. Lett.* **2006**, *8*, 4601–4604.
- (34) Rinkel, J.; Rabe, P.; Chen, X.; Köllner, T. G.; Chen, F.; Dickschat, J. S. Mechanisms of the Diterpene Cyclases  $\beta$ -Pinacene Synthase from *Dictyostelium Discoideum* and Hydroxyrene Synthase from *Streptomyces Clavuligerus*. *Chem. – Eur. J.* **2017**, *23*, 10501–10505.
- (35) Hahn, F. M.; Hurlburt, A. P.; Poulter, C. D. *Escherichia Coli* Open Reading Frame 696 Is *idi*, a Nonessential Gene Encoding Isopentenyl Diphosphate Isomerase. *J. Bacteriol.* **1999**, *181*, 4499–4504.
- (36) Rabe, P.; Rinkel, J.; Nubbemeyer, B.; Köllner, T. G.; Chen, F.; Dickschat, J. S. Terpene Cyclases from Social Amoebae. *Angew. Chem., Int. Ed. Engl.* **2016**, *55*, 15420–15423.
- (37) Rinkel, J.; Lauterbach, L.; Rabe, P.; Dickschat, J. S. Two Diterpene Synthases for Spiroalbatene and Cembrene A from *Allokutzneria Albata*. *Angew. Chem., Int. Ed.* **2018**, *57*, 3238–3241.
- (38) Rinkel, J.; Dickschat, J. S. Addressing the Chemistry of Germacrene A by Isotope Labeling Experiments. *Org. Lett.* **2019**, *21*, 2426–2429.
- (39) Rabe, P.; Barra, L.; Rinkel, J.; Riclea, R.; Citron, C. A.; Klapschinski, T. A.; Janusko, A.; Dickschat, J. S. Conformational Analysis, Thermal Rearrangement, and EI-MS Fragmentation Mechanism of (1(10)*E*,4*E*,6*S*,7*R*)-Germacradien-6-ol by <sup>13</sup>C-Labeling Experiments. *Angew. Chem., Int. Ed. Engl.* **2015**, *54*, 13448–13451.
- (40) Kuo, D. L.; Money, T. An Enantiospecific Synthesis of Longiborneol and Longifolene. *J. Chem. Soc., Chem. Commun.* **1986**, 1691–1692.
- (41) Furusaki, A.; Shirahama, H.; Matsumoto, T. Determination of the Chirality of Illudin S by the Bijvoet X-Ray Method. *Chem. Lett.* **1973**, *2*, 1293–1294.
- (42) Jang, H. I.; Rhee, K. J.; Eom, Y. B. Antibacterial and Antibiofilm Effects of  $\alpha$ -Humulene against *Bacteroides Fragilis*. *Can. J. Microbiol.* **2020**, *66*, 389–399.
- (43) Hartsel, J. A.; Eades, J.; Hickory, B.; Makriyannis, A. *Cannabis Sativa* and Hemp. *Nutraceuticals* **2016**, 735–754.
- (44) Chaturvedi, D.; Dwivedi, P. K. Recent Developments on the Antidiabetic Sesquiterpene Lactones and Their Semisynthetic Analogues. *Discovery and Development of Antidiabetic Agents from Natural Products.* **2017**, 185–207.
- (45) Ohfuné, Y.; Shirahama, H.; Matsumoto, T. Syntheses of 7 $\alpha$ - and 7 $\beta$ -Protoilludanol, and 7(13)-Protoilludene, Possible Biogenetic Intermediates for Illudoid Sesquiterpenes. *Tetrahedron Lett.* **1975**, *49*, 4377–4380.
- (46) Mohamed, S. I. A.; Jantan, I.; Haque, M. A. Naturally Occurring Immunomodulators with Antitumor Activity: an Insight on Their Mechanisms of Action. *Int. Immunopharmacol.* **2017**, *50*, 291–304.
- (47) Erkel, G.; Anke, T. (Eds.: Rehm, H. J.; Reed, G.) Products from Basidiomycetes. *Biotechnology: Products of Secondary Metabolism.* **1997**, 489–533.
- (48) Kim, D. Y.; Choi, B. Y. Costunolide-A Bioactive Sesquiterpene Lactone with Diverse Therapeutic Potential. *Int. J. Mol. Sci.* **2019**, *20*, 2926.
- (49) Cavalieri, E.; Mariotto, S.; Fabrizi, C.; de Prati, A. C.; Gottardo, R.; Leone, S.; Berra, L. V.; Lauro, G. M.; Ciampa, A. R.; Suzuki, H. Alpha-Bisabolol, a Nontoxic Natural Compound, Strongly Induces Apoptosis in Glioma Cells. *Biochem. Biophys. Res. Commun.* **2004**, *315*, 589–594.
- (50) Otaka, J.; Shimizu, T.; Futamura, Y.; Hashizume, D.; Osada, H. Structures and Synthesis of Hitoyopodins: Bioactive Aromatic Sesquiterpenoids Produced by the Mushroom *Coprinopsis Cinerea*. *Org. Lett.* **2018**, *20*, 6294–6297.
- (51) Zhang, F.; An, T.; Tang, X.; Zi, J.; Luo, H. -B.; Wu, R. Enzyme Promiscuity versus Fidelity in Two Sesquiterpene Cyclases (TEAS versus ATAS). *ACS Catal.* **2020**, *10*, 1470–1484.
- (52) Tantillo, D. J. Importance of Inherent Substrate Reactivity in Enzyme-Promoted Carbocation Cyclization/Rearrangements. *Angew. Chem., Int. Ed.* **2017**, *56*, 10040–10045.
- (53) Zhang, Q.; Catti, L.; Syntrivanis, L. -D.; Tiefenbacher, K. En Route to Terpene Natural Products Utilizing Supramolecular Cyclase Mimetics. *Nat. Prod. Rep.* **2019**, *36*, 1619–1627.
- (54) Raz, K.; Driller, R.; Dimos, N.; Ringel, M.; Bruck, T.; Loll, B.; Major, D. T. The Impression of a Nonexisting Catalytic Effect: The Role of CotB2 in Guiding the Complex Biosynthesis of Cyclooctat-9-en-7-ol. *J. Am. Chem. Soc.* **2020**, *142*, 21562–21574.
- (55) Zha, W.; Zhang, F.; Shao, J.; Ma, X.; Zhu, J.; Sun, P.; Wu, R.; Zi, J. Rationally Engineering Santalene Synthase to Readjust the Component Ratio of Sandalwood Oil. *Nat. Commun.* **2022**, *13*, 2508.
- (56) Zhang, F.; Wang, Y. -H.; Tang, X.; Wu, R. Catalytic Promiscuity of the Non-Native FPP Substrate in the TEAS Enzyme: Non-Negligible Flexibility of the Carbocation Intermediate. *Phys. Chem. Chem. Phys.* **2018**, *20*, 15061–15073.

#### NOTE ADDED AFTER ASAP PUBLICATION

This paper was published on April 5, 2023. Figures 1, 5, and 7 have been updated. The revised version was re-posted on April 10, 2023.

## Appendix R

### **Crystal Structure Based Mutagenesis of Cattleylene Synthase Leads to the Generation of Rearranged Polycyclic Diterpenes**

*Angew. Chem. Int. Ed.* **2022**, *61*, e202209785

DOI: 10.1002/ange.202209785

*Angew. Chem.* **2022**, *134*, e202209785

DOI: 10.1002/anie.202209785

VIP Enzymes Very Important Paper

 How to cite: *Angew. Chem. Int. Ed.* **2022**, *61*, e202209785

International Edition: doi.org/10.1002/anie.202209785

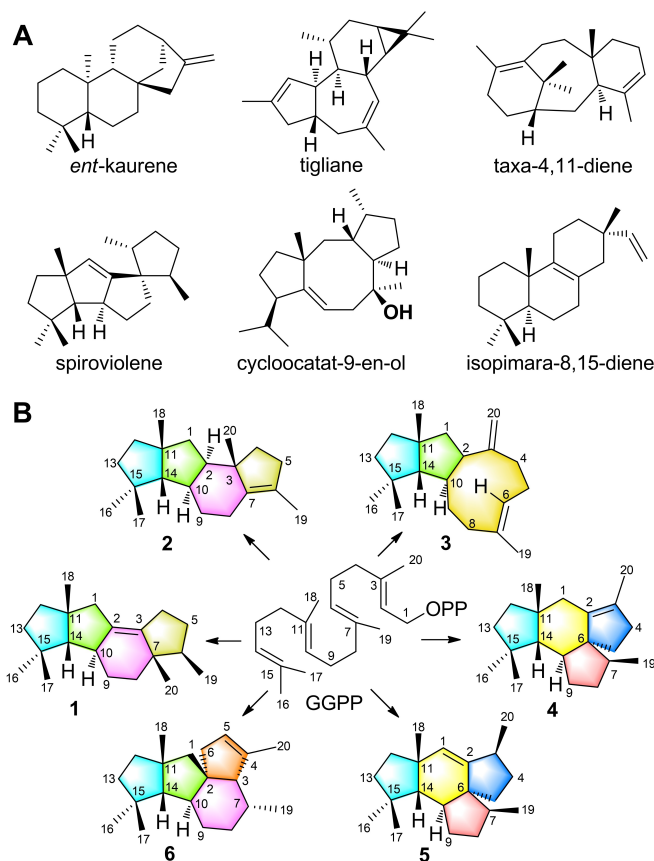
German Edition: doi.org/10.1002/ange.202209785

# Crystal Structure Based Mutagenesis of Cattleylene Synthase Leads to the Generation of Rearranged Polycyclic Diterpenes

Baiying Xing<sup>+</sup>, Houchao Xu<sup>+</sup>, Annan Li, Tingting Lou, Meng Xu, Kaibiao Wang, Zhengren Xu, Jeroen S. Dickschat,<sup>\*</sup> Donghui Yang,<sup>\*</sup> and Ming Ma<sup>\*</sup>

**Abstract:** The crystal structures of cattleylene synthase (apo-CyS), and CyS complexed with geranylgeranyl pyrophosphate (GGPP) were solved. The CyS<sup>C59A</sup> variant exhibited an increased production of cattleylene and other diterpenes with diverse skeletons. Its structure showed a widened active site cavity explaining the relaxed selectivity. Isotopic labeling experiments revealed a remarkable cyclization mechanism involving several skeletal rearrangements for one of the novel diterpenes.

Diterpenoids exhibit diverse chemical skeletons and important biological activities.<sup>[1]</sup> Because of the larger number of possible reactions for geranylgeranyl pyrophosphate (GGPP) as compared to geranyl (GPP) and farnesyl pyrophosphate (FPP), diterpenoids usually exhibit more complex skeletons than mono- and sesquiterpenoids (Figure 1A). Some polycyclic diterpenoids have attracted increasing attention, e.g. gibberellins are phytohormones derived from *ent*-kaurene,<sup>[2]</sup> while phorbol esters exhibiting the tigliane skeleton are currently in phase II clinical trials for the treatment of acute myeloid leukemia (Figure 1A).<sup>[3,4]</sup> Their biosynthesis is attributed to diterpene synthases (DTSs), the type I of which catalyzes the conversion of GGPP through diphosphate abstraction and cationic cascade reactions. Substrate ionization is mediated by a Lewis acidic trinuclear Mg<sup>2+</sup> cluster, bound itself to a highly conserved Asp-rich motif (DDXX(X)D) and an NSE/DTE triad



**Figure 1.** Representative diterpenes. A) Compounds from previous studies, B) products of CyS and its variants.

[\*] B. Xing,<sup>+</sup> A. Li, T. Lou, M. Xu, K. Wang, Dr. Z. Xu, Dr. D. Yang, Dr. M. Ma

State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University  
 38 Xueyuan Road, Haidian District, Beijing 100191 (China)  
 E-mail: ydhui@bjmu.edu.cn  
 mma@bjmu.edu.cn

H. Xu,<sup>+</sup> Prof. Dr. J. S. Dickschat  
 Kekulé-Institute for Organic Chemistry and Biochemistry, University of Bonn  
 Gerhard-Domagk-Strasse 1, 53121 Bonn (Germany)  
 E-mail: dickschat@uni-bonn.de

[†] These authors contributed equally to this work.

© 2022 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

((N,D)DLX(S,T)XXXE), to which the substrate's pyrophosphate docks,<sup>[5]</sup> with assistance of a highly conserved Arg residue (pyrophosphate sensor).<sup>[6]</sup> A main chain carbonyl oxygen in the effector triad is involved in the stabilization of the initially formed allyl cation<sup>[6]</sup> and serves as a catalytic base and acid in the formation and reprotonation of neutral intermediates.<sup>[7]</sup> The mechanisms of these multistep processes can be probed by isotopic labeling experiments,<sup>[8]</sup> revealing an astonishing complexity associated with a single enzymatic transformation. This enzymatic power is often superior to the long and laborious routes<sup>[9]</sup> with low overall yields<sup>[10]</sup> to diterpenes by chemical synthesis. Several DTSs with structurally complex products have been reported,<sup>[11–20]</sup> but only a few structures of class I DTSs have been solved, including taxadiene synthase from *Taxus brevifolia*,<sup>[21]</sup>



spiroviolene synthase (SvS) from *Streptomyces violens*<sup>[22]</sup> and cyclooctat-9-en-7-ol synthase (CotB2) from *S. melanosporofaciens*,<sup>[23–25]</sup> ent-kaurene synthase from *Bradyrhizobium japonicum*,<sup>[26]</sup> and isopimarane synthases Sat1646 from *Salinispora* sp. and Stt4548 from *Streptomyces* sp. (Figure 1A).<sup>[27]</sup> Structural knowledge of DTSs is of interest to deepen our mechanistic understanding of these enzymes and allows for structure based site-directed mutagenesis.<sup>[25,28,29]</sup>

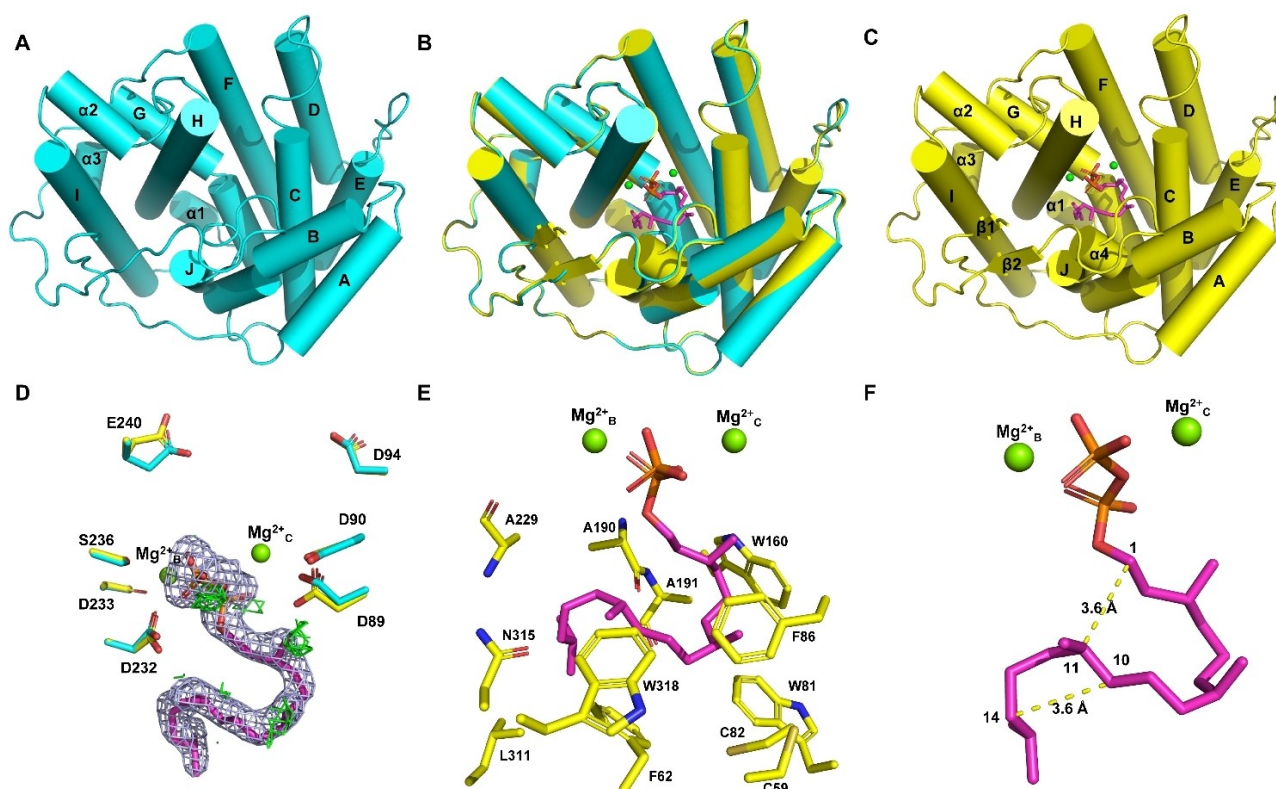
We recently discovered the cattleyene (**1**) synthase (CyS) from *Streptomyces cattleya* (Figure 1B) and studied its cyclization mechanism through isotopic labelings.<sup>[18]</sup> Here we report on the crystal structures of apo-CyS, CyS complexed with GGPP and Mg<sup>2+</sup> (CyS-GGPP-Mg<sup>2+</sup>), and the CyS<sup>C59A</sup> enzyme variant. Modellings in conjunction with site-directed mutagenesis and additional labeling experiments are discussed that provide a deeper understanding of cattleyene production by CyS.

High-quality crystals of purified CyS (Figure S1) were obtained and the structure of apo-CyS was solved at 2.00 Å, using the structure of SvS<sup>[22]</sup> as template (PDB ID: 6TBD). Crystals of CyS-GGPP-Mg<sup>2+</sup> were obtained by soaking and the structure was solved at 1.87 Å using the apo-CyS structure as template. CyS-GGPP-Mg<sup>2+</sup> is the first structure of a terpene synthase (TS) in complex with the native

substrate GGPP, providing an ideal opportunity for analyzing GGPP binding and interactions in the active site.

Apo-CyS adopts the classical  $\alpha$ -helical fold of class I TSs, with ten core (A–J) and three short  $\alpha$  helices ( $\alpha$ 1– $\alpha$ 3, Figure 2A). The Asp-rich motif (D<sup>89</sup>DVHCD<sup>94</sup>) is located on helix C and the NSE triad (D<sup>232</sup>DLFS<sup>236</sup>YGKE<sup>240</sup>) on helix H. The CyS-GGPP-Mg<sup>2+</sup> structure shows a similar fold to apo-CyS (Figure 2B), with a root-mean-square deviation (RMSD) of 0.24 Å for C $\alpha$  atoms. Upon GGPP and Mg<sup>2+</sup> binding one more  $\alpha$  helix ( $\alpha$ 4) and two  $\beta$  strands ( $\beta$ 1 and  $\beta$ 2) close to the active site become ordered (Figure 2C). Further differences are observed for residues involved in Mg<sup>2+</sup>-binding (Figure 2D), but only two Mg<sup>2+</sup> ions are found (Mg<sup>2+</sup><sub>C</sub> coordinated by R324, D89 and D90, and Mg<sup>2+</sup><sub>B</sub> coordinated by R186, N232 and S236; Figures 2D and S1B), while Mg<sup>2+</sup><sub>A</sub> as observed in selinadiene synthase<sup>[6]</sup> is missing. The pyrophosphate moiety of GGPP binds to both Mg<sup>2+</sup> and the conserved C-terminal RY (R324 and Y325).

GGPP is surrounded by five aromatic (F62, W81, F86, W160, and W318), four aliphatic (A190, A191, A229, and L311) and three polar residues (C59, C82, and N315, Figure 2E). These interactions render GGPP folded into a specific conformation, in which C-11 and C-14 are close to C-1 and C-10 with both distances of 3.6 Å (Figure 2F), allowing the formation of the 5/11 bicyclic intermediate in the first cyclization steps. The GGPP conformation is P-



**Figure 2.** Crystal structures of A) apo-CyS, B) superimposition of apo-CyS and CyS-GGPP-Mg<sup>2+</sup>, and C) CyS-GGPP-Mg<sup>2+</sup>. D) Superimposition of aspartate-rich motifs in apo-CyS (cyan) and CyS-GGPP-Mg<sup>2+</sup> (yellow). The  $2F_o - F_c$  and  $F_o - F_c$  electron density maps of GGPP, contoured at  $2\sigma$ , are shown as light blue and green meshes (only positive densities are found). E) Residues surrounding GGPP. F) Conformation of GGPP in the active site. Mg<sup>2+</sup> ions are shown in green, GGPP is shown in magenta.

helical from C-1 to C-11 and M-helical from C-10 to C-14, which explains the observed stereoselectivity of the C-1/C-11 and C-10/C-14 bond formations to generate the 10*S*, 11*R* and 14*S* configurations.

Key intermediates were modelled into the CyS active site using CyS-GGPP-Mg<sup>2+</sup> as the macromolecule in AutoDock Vina 1.1.2, by removing water and the geranyl-geranyl moiety of GGPP. Intermediates **A–H** were prepared by using default parameters (Supporting Information). The intermediates **A–G** are stabilized by cation- $\pi$  interactions and van der Waals forces with aromatic residues (F62, W318, W81, F86, and W160, Figure 3). The last intermediate **H** is stabilized by a cation-dipole interaction with the A190 main chain carbonyl group (effector), in an equivalent position to the effector G182 in selinadiene synthase.<sup>[6]</sup> Notably, no polar residue or water is found near the cation at C-3, but the pyrophosphate is only 4.1 Å away from C-2 and may abstract the C-2 proton to form **1** (Figure 3).<sup>[18]</sup>

The above modellings show how CyS catalyzes the biosynthesis of **1**. The roles of the identified key residues were then investigated by site-directed mutagenesis (Figure 4). For this purpose, an engineered *E. coli* strain containing a reconstructed isopentenol utilization pathway (IUP) to produce isopentenyl (IPP) and dimethylallyl pyrophosphate (DMAPP) was used (Figure S2). The F62A and W318A enzyme variants showed a substantially decreased or completely abolished production of **1**, confirming the important roles of F62 and W318 in the biosynthesis of **1**. The W81A variant only produced minor amounts of **1**. Considering the interaction of W81 with GGPP, this residue may be required to keep the substrate and the intermediates in suitable conformations.

The variations of A190G and A191G caused the formation of an additional product **2**, while the exchanges of F86A and W160A resulted in decreased levels of **1** with simultaneous production of several new compounds (**3–6**). Diterpene **3** was only observed in trace amounts, and yields of **4**, **5** and **6** were about 30 %, 20 % and 20 % of that of **1**, with an overall higher production by the W160A in comparison to the F86A variant. Gratifyingly, the C59A variant gave a similar product profile, but with a much better production (about 6-fold in comparison to the F86A variant except for **3**). Exchanges of other active site residues (C82A, A229G and N315A) led to no significant product change compared to wild-type CyS (Figure S3).

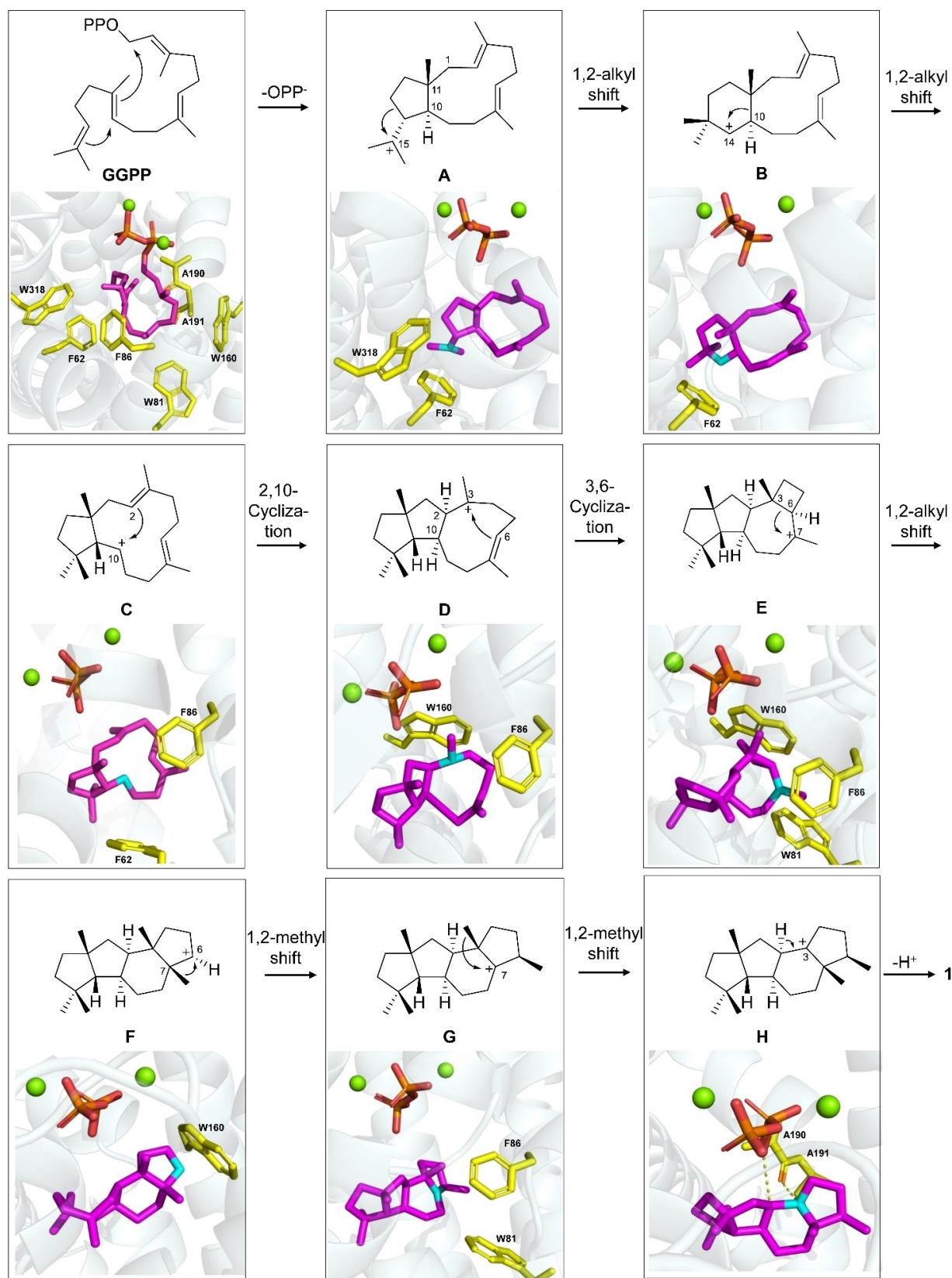
Compounds **2–6** were isolated from large scale fermentations of engineered *E. coli* expressing the CyS A190G or C59A variety, and their structures were elucidated by NMR spectroscopy. All five diterpenes exhibit different carbon skeletons (Figure 1B) and are formed from several of the proposed pathway intermediates. Compound **2** arises by deprotonation at C-6 of **G**, **3** is generated through proton abstraction at C-20 of **D**, and allokutznerene<sup>[16]</sup> (**4**) and **5** originate from alternative deprotonations of **C** (Figure 3 and S6). The production of **3–6** from the enzyme variants of F86A and W160A supports the role of these residues in stabilizing intermediates **D** and **E**. The F86A and W160A variants are incompetent to stabilize these intermediates

which consequently leads to shunt products derived from **C** and **D** (Figure 3).

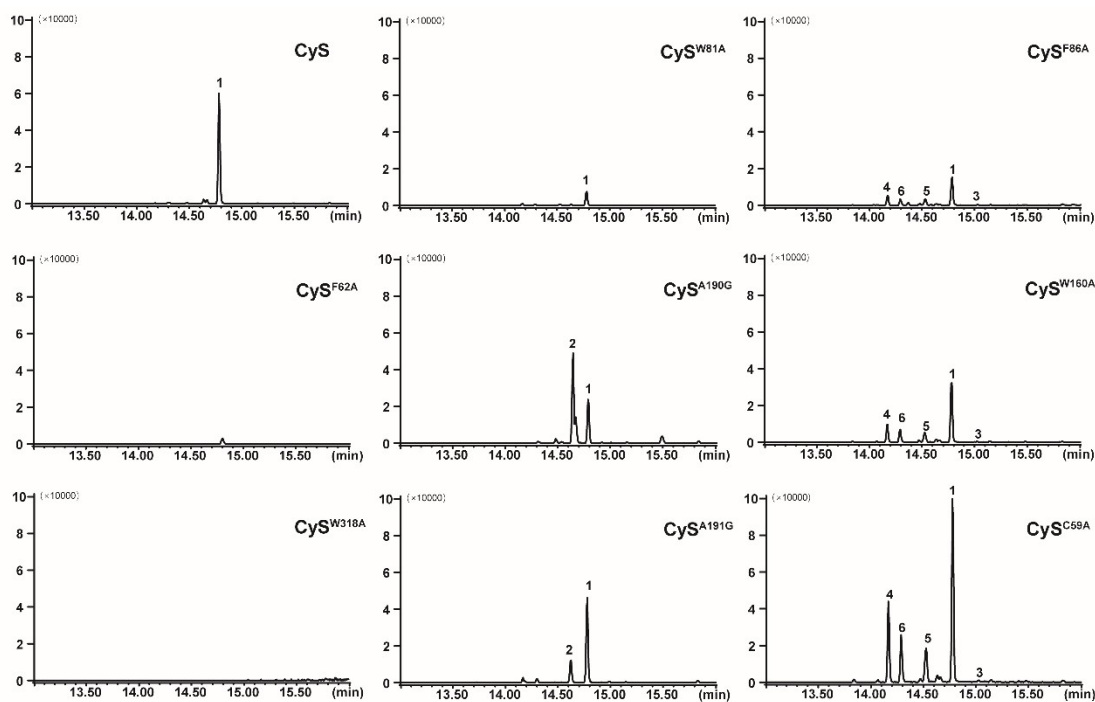
For a deeper understanding of the increased but less selective production by the C59A variant (Cys<sup>C59A</sup>), its crystal structure was solved at 2.30 Å. The structure is highly similar to the apo-CyS and CyS-GGPP-Mg<sup>2+</sup> structures (Figure 5A, RMSD of 0.16 Å and 0.15 Å for Ca atoms). The active site residues of the three structures superimpose well, with the exception of F86 located between C59 and the substrate binding pocket (Figure 5B). In the apo-CyS and CyS-GGPP-Mg<sup>2+</sup> structures, the thiol of C59 is close to the phenyl ring of F86 (3.4 Å) and renders it towards the active site, while in Cys<sup>C59A</sup> this interaction is disrupted. As a result, F86 moves a bit away from the active site and the phenyl ring rotates 24° clockwise (Figure 5B). This slightly widens the active site cavity which may lead to an improved uptake of GGPP, albeit on the expense of selectivity because of a less tight substrate control through cation- $\pi$  interactions. Similar observations have been made before for SvS,<sup>[22]</sup> which together with our results provides a basis for future TS engineering strategies.

Compound **6** possesses a novel skeleton and the mechanism of its formation was further investigated. For this purpose, all 20 isotopomers of (<sup>13</sup>C)GGPP, prepared enzymatically from <sup>13</sup>C-labeled FPP, GPP or IPP precursors (Table S8), were enzymatically converted with Cys<sup>C59A</sup>, followed by extraction with C<sub>6</sub>D<sub>6</sub> and analysis of the product mixture through <sup>13</sup>C NMR (Figure S34 and S35). All 20 experiments resulted in the detection of the labeled carbons of the six products, with one signal matching the NMR data of **6** in each experiment. The results revealed a remarkable mechanism for its formation with the first steps towards **C** being the same as for the other products, but then branching out through **I** to **O** with involvement of multiple ring closures, 1,2-hydride shifts, and skeletal rearrangements (Figure 6). The 1,2-hydride shift from intermediate **I** to **J** was investigated with (3-<sup>13</sup>C,2-<sup>2</sup>H)FPP<sup>[30]</sup> and IPP with GGPP synthase (GGPPS)<sup>[13]</sup> and Cys<sup>C59A</sup>, resulting in a slightly upfield shifted triplet for C-7 of **6** (Figure S36A and B) due to a direct <sup>13</sup>C-<sup>2</sup>H bond in the product. The 1,2-hydride shifts from **K** to **L** and from **N** to **O** represent a forward and backward movement of the same hydrogen. Consequently, when using (3-<sup>13</sup>C,2-<sup>2</sup>H)GGPP<sup>[13]</sup> with Cys<sup>C59A</sup> the deuterium atom will end up in its starting position, in agreement with the observed minor upfield shift for the signal of C-3 of **6**, indicating a deuterium atom in a neighbouring position (Figure S36C and D). The stereoselectivity of the final deprotonation was investigated by conversion of DMAPP and (*E*)- or (*Z*)-(4-<sup>2</sup>H, 4-<sup>13</sup>C)IPP<sup>[16]</sup> with GGPPS and Cys<sup>C59A</sup>, showing loss of deuterium from (*E*)-(4-<sup>2</sup>H,4-<sup>13</sup>C)IPP and retainment from (*Z*)-(4-<sup>2</sup>H,4-<sup>13</sup>C)IPP, i.e. loss of the  $\alpha$ -oriented proton in **O** (Figure S37).

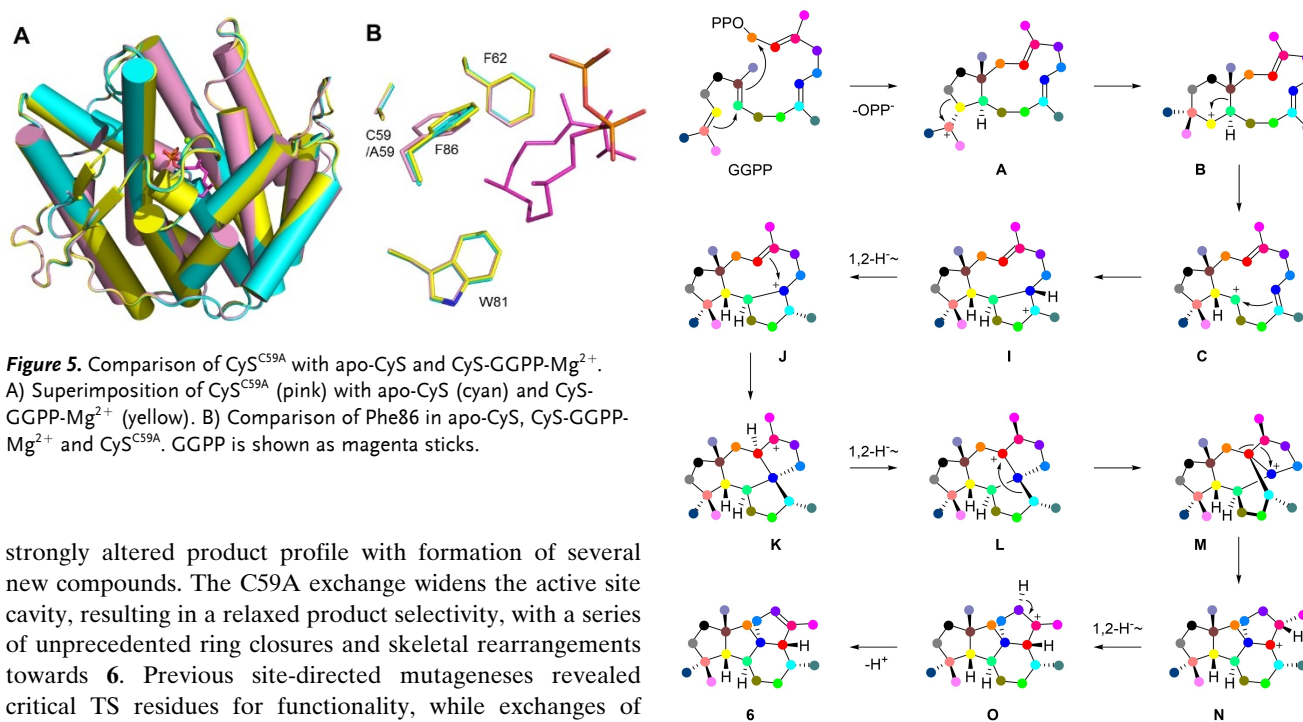
In summary, the structures of CyS and CyS-GGPP-Mg<sup>2+</sup>, representing the first example of a TS in complex with its native substrate, were solved. Intermediate modelling and site-directed mutagenesis gave detailed insights into the biosynthesis of **1**. Based on the structure, several CyS variants were designed, and especially Cys<sup>C59A</sup> showed a



**Figure 3.** Biosynthesis of **1** and modelling of intermediates into the active site of Cys. Colour code: Active site residues (yellow), geranylgeranyl chain and intermediates **A–H** (magenta, cationic centers in cyan), diphosphate (red), and  $\text{Mg}^{2+}$  ions (green). A previously suggested mechanism avoids secondary cation **B** by a concerted mechanism from **A** to **C**.<sup>[13,18]</sup>



**Figure 4.** GC-MS analysis (extracted ion chromatograms at  $m/z$  272) of products by CyS and its variants.



**Figure 5.** Comparison of CyS<sup>C59A</sup> with apo-CyS and CyS-GGPP-Mg<sup>2+</sup>. A) Superimposition of CyS<sup>C59A</sup> (pink) with apo-CyS (cyan) and CyS-GGPP-Mg<sup>2+</sup> (yellow). B) Comparison of Phe86 in apo-CyS, CyS-GGPP-Mg<sup>2+</sup> and CyS<sup>C59A</sup>. GGPP is shown as magenta sticks.

strongly altered product profile with formation of several new compounds. The C59A exchange widens the active site cavity, resulting in a relaxed product selectivity, with a series of unprecedented ring closures and skeletal rearrangements towards **6**. Previous site-directed mutageneses revealed critical TS residues for functionality, while exchanges of other residues lead to changed product profiles.<sup>[31]</sup> Amino acid sequence alignments to selinadiene synthase, for which the structure and active site residues are known, allow for an identification of residues that presumably contour the active site cavities of other enzymes. This enabled the generation of enzyme variants of polytrichastrene synthase to obtain novel products.<sup>[20]</sup> However, such alignment based targetings cannot fully substitute for the structure based identification of residues, and in fact the position analogous to C59 in CyS

**Figure 6.** Cyclization mechanism from GGPP to **6**.

has not been targeted in any other TS before. Even more structural information will be required to deepen our understanding of TS catalysis and to open the possibility of sequence-function predictions and rational enzyme engineering.

## Acknowledgements

We thank Fuling Yin and Hongli Jia (Peking University) for X-ray diffraction tests, Haiyan Tao and Wen Ma (Peking University), for GC-MS analyses, and the staff from beamlines BL17U1/BL19U1 (NFPS, Shanghai) for assistance during data collection. This research was funded by the National Natural Science Foundation of China (grants 21877002, 22077007, 81991525, 81573326, 22107007), key project at central government level (2060302-2201-17), and the German Research Foundation DFG (DI1536/7-2). 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 Supporting Information of this article. Crystal structures have been deposited in the Protein Data Bank with PDB IDs 7Y50 (apo-CyS), 7Y88 (CyS-GGPP-Mg<sup>2+</sup>), and 7Y87 (CyS<sup>C59A</sup>).

**Keywords:** Biosynthesis · Diterpenes · Enzymes · Natural Products · Reaction Mechanisms

- [1] J. D. Rudolf, T. A. Alsup, B. Xu, Z. Li, *Nat. Prod. Rep.* **2021**, *38*, 905–980.
- [2] B. Tudzynski, *Appl. Microbiol. Biotechnol.* **2005**, *66*, 597–611.
- [3] S. Kawamura, H. Chu, J. Felding, P. S. Baran, *Nature* **2016**, *532*, 90–93.
- [4] H. J. Mackay, C. J. Twelves, *Nat. Rev. Cancer* **2007**, *7*, 554–562.
- [5] C. M. Starks, K. Back, J. Chappell, J. P. Noel, *Science* **1997**, *277*, 1815–1820.
- [6] P. Baer, P. Rabe, K. Fischer, C. A. Citron, T. A. Klapschinski, M. Groll, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2014**, *53*, 7652–7656; *Angew. Chem.* **2014**, *126*, 7783–7787.
- [7] Y. H. Wang, H. Xu, J. Zou, X. B. Chen, Y. Q. Zhuang, W. L. Liu, E. Celik, G. D. Chen, D. Hu, H. Gao, R. Wu, P. H. Sun, J. S. Dickschat, *Nat. Catal.* **2022**, *5*, 128–135.
- [8] J. S. Dickschat, *Angew. Chem. Int. Ed.* **2019**, *58*, 15964–15976; *Angew. Chem.* **2019**, *131*, 16110–16123.
- [9] T. Asaba, Y. Katoh, D. Urabe, M. Inoue, *Angew. Chem. Int. Ed.* **2015**, *54*, 14457–14461; *Angew. Chem.* **2015**, *127*, 14665–14669.
- [10] O. Corminboeuf, L. E. Overman, L. D. Pennington, *J. Am. Chem. Soc.* **2003**, *125*, 6650–6652.
- [11] T. Toyomasu, A. Kaneko, T. Tokiwano, Y. Kanno, Y. Kanno, R. Niida, S. Miura, T. Nishioka, C. Ikeda, W. Mitsuhashi, T. Dairi, T. Kawano, H. Oikawa, N. Kato, T. Sassa, *J. Org. Chem.* **2009**, *74*, 1541–1548.
- [12] Y. Yamada, T. Kuzuyama, M. Komatsu, K. Shin-ya, S. Omura, D. E. Cane, H. Ikeda, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 857–862.
- [13] P. Rabe, J. Rinkel, E. Dolja, T. Schmitz, B. Nubbemeyer, T. H. Luu, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2017**, *56*, 2776–2779; *Angew. Chem.* **2017**, *129*, 2820–2823.
- [14] G. Bian, Y. Han, A. Hou, Y. Yuan, X. Liu, Z. Deng, T. Liu, *Metab. Eng.* **2017**, *42*, 1–8.
- [15] J. Rinkel, L. Lauterbach, P. Rabe, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2018**, *57*, 3238–3241; *Angew. Chem.* **2018**, *130*, 3292–3296.
- [16] L. Lauterbach, J. Rinkel, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2018**, *57*, 8280–8283; *Angew. Chem.* **2018**, *130*, 8412–8415.
- [17] T. Mitsuhashi, T. Kikuchi, S. Hoshino, M. Ozeki, T. Awakawa, S. P. Shi, M. Fujita, I. Abe, *Org. Lett.* **2018**, *20*, 5606–5609.
- [18] J. Rinkel, S. T. Steiner, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2019**, *58*, 9230–9233; *Angew. Chem.* **2019**, *131*, 9328–9332.
- [19] Z. Li, Y. Jiang, X. Zhang, Y. Chang, S. Li, X. Zhang, S. Zheng, C. Geng, P. Men, L. Ma, Y. Yang, Z. Gao, Y. J. Tang, S. Li, *ACS Catal.* **2020**, *10*, 5846–5851.
- [20] A. Hou, B. Goldfuss, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2021**, *60*, 20781–20785; *Angew. Chem.* **2021**, *133*, 20949–20953.
- [21] M. Köksal, Y. Jin, R. M. Coates, R. Croteau, D. W. Christianson, *Nature* **2011**, *469*, 116–120.
- [22] K. Schriever, P. Saenz-Mendez, R. S. Rudraraju, N. M. Hendrikse, E. P. Hudson, A. Biundo, R. Schnell, P. O. Syrén, *J. Am. Chem. Soc.* **2021**, *143*, 3794–3807.
- [23] R. Janke, C. Görner, M. Hirte, T. Brück, B. Loll, *Acta Crystallogr. Sect. D* **2014**, *70*, 1528–1537.
- [24] T. Tomita, S. Y. Kim, K. Teramoto, A. Meguro, T. Ozaki, A. Yoshida, Y. Motoyoshi, N. Mori, K. Ishigami, H. Watanabe, M. Nishiyama, T. Kuzuyama, *ACS Chem. Biol.* **2017**, *12*, 1621–1628.
- [25] R. Driller, S. Janke, M. Fuchs, E. Warner, A. R. Mhashal, D. T. Major, M. Christmann, T. Brück, B. Loll, *Nat. Commun.* **2018**, *9*, 3971.
- [26] W. Liu, X. Feng, Y. Zheng, C. H. Huang, C. Nakano, T. Hoshino, S. Bogue, T. P. Ko, C. C. Chen, Y. Cui, J. Li, I. Wang, S. T. Hsu, E. Oldfield, R. T. Guo, *Sci. Rep.* **2014**, *4*, 6214.
- [27] B. Xing, J. Yu, C. Chi, X. Ma, Q. Xu, A. Li, Y. Ge, Z. Wang, T. Liu, H. Jia, F. Yin, J. Guo, L. Huang, D. Yang, M. Ma, *Commun. Chem.* **2021**, *4*, 140.
- [28] K. Raz, R. Driller, N. Dimos, M. Ringel, T. Brück, B. Loll, D. T. Major, *J. Am. Chem. Soc.* **2020**, *142*, 21562–21574.
- [29] S. Edgar, F. S. Li, K. Qiao, J. K. Weng, G. Stephanopoulos, *ACS Synth. Biol.* **2017**, *6*, 201–205.
- [30] T. A. Klapschinski, P. Rabe, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2016**, *55*, 10141–10144; *Angew. Chem.* **2016**, *128*, 10296–10299.
- [31] H. Xu, J. S. Dickschat, *Synthesis* **2022**, *54*, 1551–1565.

Manuscript received: July 5, 2022

Accepted manuscript online: July 12, 2022

Version of record online: August 1, 2022

## Appendix S


### **Mechanistic Investigations on Microbial Type I Terpene Synthases through Site-Directed Mutagenesis**

*Synthesis* **2022**, *54*, 1551

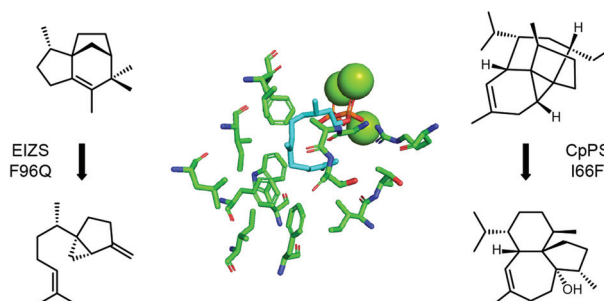
DOI: [10.1055/a-1675-8208](https://doi.org/10.1055/a-1675-8208)

# Mechanistic Investigations on Microbial Type I Terpene Synthases through Site-Directed Mutagenesis

Houchao Xu

Jeroen S. Dickschat\* 

Kekulé-Institute for Organic Chemistry and Biochemistry,  
University of Bonn, Gerhard-Domagk-Straße 1, 53121 Bonn,  
Germany  
dickschat@uni-bonn.de



Received: 14.09.2021

Accepted after revision: 21.10.2021

Published online: 21.10.2021

DOI: 10.1055/a-1675-8208; Art ID: ss-2021-r0544-sr

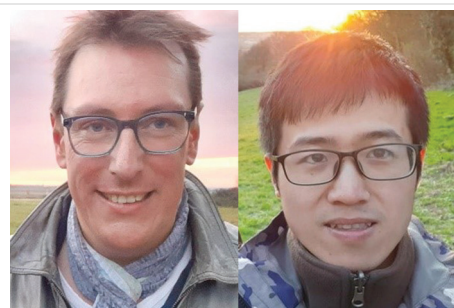
**Abstract** During the past three decades many terpene synthases have been characterised from all kingdoms of life. Enzymes of type I, from bacteria, fungi and protists, commonly exhibit several highly conserved motifs and single residues, and the available crystal structures show a shared  $\alpha$ -helical fold, while the overall sequence identity is generally low. Several enzymes have been studied by site-directed mutagenesis, giving valuable insights into terpene synthase catalysis and the intriguing mechanisms of terpene synthases. Some mutants are also preparatively useful and give higher yields than the wild type or a different product that is otherwise difficult to access. The accumulated knowledge obtained from these studies is presented and discussed in this review.

- 1 Introduction
- 2 Residues for Substrate Binding and Catalysis
- 3 Residues with Structural Function
- 4 Residues Contouring the Active Site Cavity
- 5 Other Residues
- 6 Conclusions

**Key words** terpenes, biosynthesis, mutagenesis, enzyme mechanisms

## 1 Introduction

Despite the fact that they are biosynthesised from only two building blocks, the C1 electrophile dimethylallyl diphosphate (DMAPP) and the C4 nucleophile isopentenyl diphosphate (IPP), terpenes constitute the largest class of natural products, with more than 80,000 compounds known to date. Terpenes exhibit many ecological functions, have important medicinal applications, and can be transformed into useful fine chemicals such as fragrances and pharmaceuticals.<sup>1,2</sup> Their biosynthesis proceeds through coupling of the monomers DMAPP and IPP, catalysed by an oligoprenyl diphosphate synthase (Scheme 1), leading in the first step to the C<sub>10</sub> monoterpene precursor geranyl diphosphate

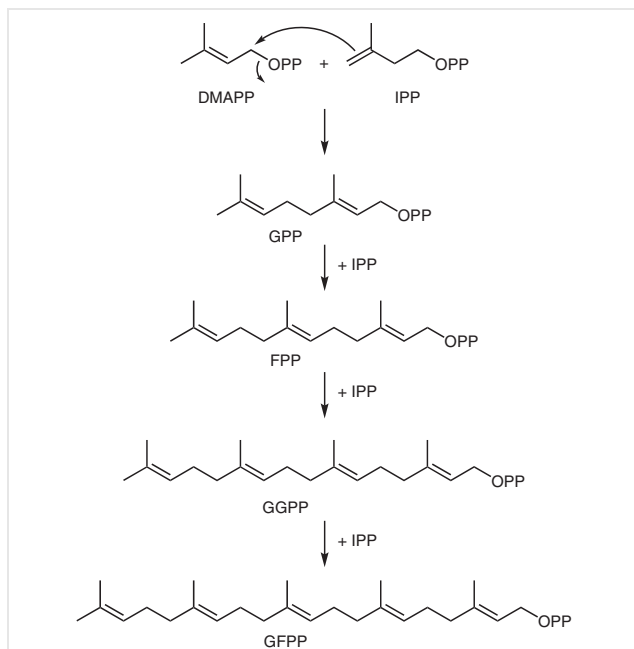


**Jeroen S. Dickschat** studied chemistry at TU Braunschweig (Germany), where he got involved in natural product chemistry working with Prof. Stefan Schulz to obtain his PhD in 2004. He then moved for postdoctoral research stays with Prof. Rolf Müller to Saarland University (Germany) and with Prof. Peter Leadlay to the University of Cambridge (UK). In 2008 he became an independent group leader at TU Braunschweig and received his *venia legendi* in 2013. Since 2014 he has been a Professor of Organic Chemistry and Biochemistry at the Kekulé Institute at the University in Bonn (Germany). Since 2016 he has been an honorary fellow of the Netherlands Institute of Ecology in Wageningen and holds an Honorary Professorship at Zhejiang University of Technology (P. R. of China). He serves as an Editor for the Beilstein Journal of Organic Chemistry and the Chemical and Pharmaceutical Bulletin. His research interests span all aspects of natural product biosynthesis, with a special emphasis on terpenes, volatile natural products and enzyme mechanisms.

**Houchao Xu** obtained his B.Sc. degree in medicinal chemistry from the Kunming Institute of Botany, Chinese Academy of Sciences (P. R. of China). He joined Prof. Dickschat's group as a PhD student at the University of Bonn (Germany) in September 2019. His research focuses on the biosynthesis of terpenes and polyketides by using isotope labelling techniques.

(GPP). Additional couplings of IPP subsequently lead to the sesquiterpene precursor farnesyl diphosphate (C<sub>15</sub>, FPP), geranylgeranyl diphosphate (C<sub>20</sub>, GGPP) for diterpene biosynthesis, and geranylgeranyl diphosphate (C<sub>25</sub>, GFPP) towards sesterterpenes. These acyclic precursors can then be

cyclised by terpene synthases (TPSs) to yield structurally complex terpenes that often contain multiple rings and stereogenic centres.<sup>3,4</sup> Remarkably, these cyclisation reactions can be realised by a single enzyme, the role of which is to ionise the substrate and force it inside the active site into a reactive conformation that determines the structure of the final product. After ionisation, a cationic cascade reaction is promoted with elementary steps that are typical for carbocation chemistry, such as intramolecular attacks of olefinic double bonds to the cationic centre with ring closure, or, less often, ring openings through the reverse process, Wagner–Meerwein rearrangements, hydride or proton migrations, ultimately leading to terpene hydrocarbons in a final deprotonation step. Eventually, water can be incorporated by nucleophilic attack at a cationic intermediate, which explains the formation of terpene alcohols and, more rarely, ethers.



**Scheme 1** Biosynthesis of terpene precursors by oligoprenyl diphosphate synthases

How is substrate ionisation achieved? This is possible through two alternative mechanisms: Type I TPSs ionise their substrates by the abstraction of diphosphate, producing a reactive allyl cation, while type II TPSs act by protonation of the substrate at a double bond or an epoxide installed by a preceding oxidation step.<sup>5</sup> In this review article we will focus on type I enzymes from microorganisms (bacteria, fungi and protists) to give a summary of mutational studies that not only shaped our mechanistic understanding of these biocatalysts, but also resulted in enzyme variants that can be used for the synthesis of new compounds in vitro. While most work on TPSs has been conducted us-

ing site-directed mutagenesis (SDM), the technique has been further developed towards saturation mutagenesis and directed evolution with main contributions inter alia by Reetz, Hilvert and Arnold, culminating in the Nobel prize in Chemistry 2018.<sup>6,7</sup> Rational enzyme design by Focused Rational Iterative Site-specific Mutagenesis (FRISM) has also been successful for obtaining highly stereoselective biocatalysts.<sup>8,9</sup> The accumulated knowledge obtained by SDM of TPSs presented in this article will be organised according to the targeted amino acid residues, starting with residues that are located within highly conserved motifs and directly involved in substrate binding and enzyme catalysis, followed by conserved residues of structural importance. Of course, there can be a smooth transition between these categories. The last sections will include residues contouring the active site cavity, and other residues that cannot easily be classified as members of one of the other groups.

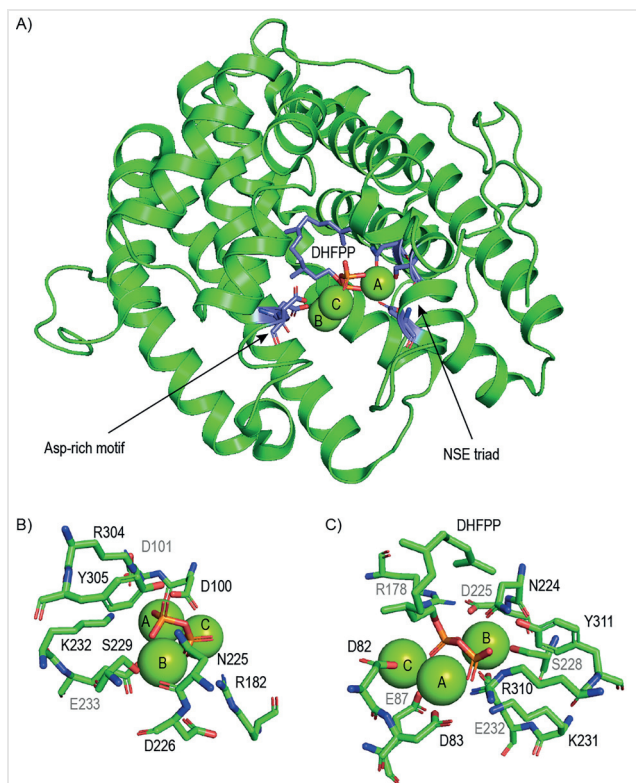
## 2 Residues for Substrate Binding and Catalysis

Microbial type I TPSs show an  $\alpha$ -helical ‘terpenoid synthase fold’,<sup>10</sup> similar to the structural observations first made for avian FPP synthase (Figure 1A).<sup>11</sup> Their active site is composed of several highly conserved motifs and residues, including the aspartate-rich motif DDXX(X)D<sup>5</sup> observed for trichodiene synthase (TS) from *Fusarium sporotrichioides* at the C-terminal end of helix D (starting at position 100), and the NSE triad ND(L,I,V)XSXX(R,K)E<sup>12</sup> on helix H (starting at position 225) that are involved in binding of the  $Mg^{2+}$  cofactor.<sup>13</sup> Specifically, D100 coordinates to two  $Mg^{2+}$  cations ( $Mg^{2+}_A$  and  $Mg^{2+}_C$ ) and N225, S229 and E233 bind to the third  $Mg^{2+}_B$  that, in turn, are coordinated by diphosphate (Figure 1B). In the structure of selina-4(15),7(11)-diene synthase (SdS) from *Streptomyces pristinaespiralis* in complex with 2,3-dihydro-FPP (DHFPP) the active site is completely closed, showing binding of the first Asp and the third Glu (D82 and E87) of the Asp-rich motif (<sup>82</sup>DDGHCE) to  $Mg^{2+}_A$  and  $Mg^{2+}_C$  and of N224, S228 and E232 of the NSE triad to  $Mg^{2+}_B$  (Figure 1C).<sup>14</sup> Strikingly, as first observed in the crystal structure of tobacco 5-*epi*-aristolochene synthase (TEAS),<sup>15</sup> plant TPSs exhibit a very similar  $\alpha$ -helical fold and active site architecture, despite the fact that the relative positioning of the highly conserved motifs within the amino acid sequence is different.

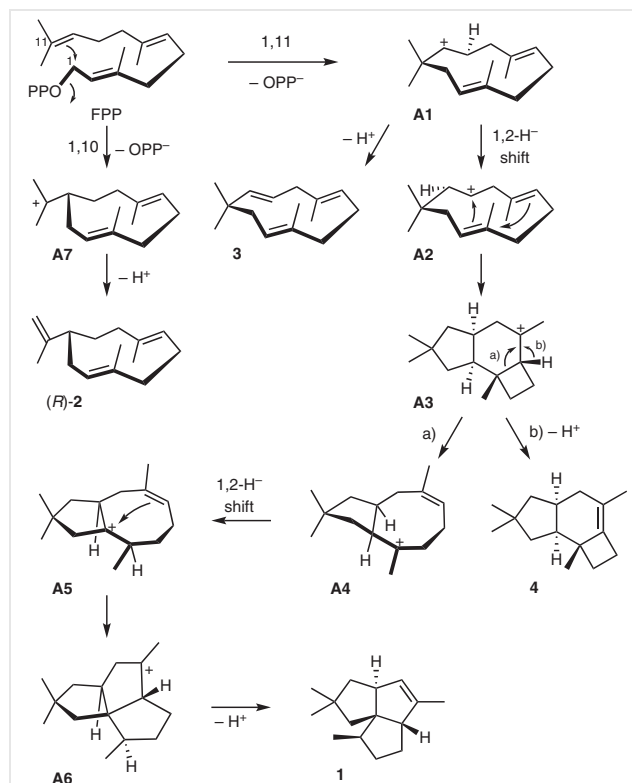
### 2.1 The Asp-Rich Motif

Pentalenene synthase (PS) from *Streptomyces exfoliatus* catalyses the conversion of FPP into pentalenene (**1**, Scheme 2).<sup>16–18</sup> This enzyme is not only the first characterised bacterial TPS, but also the first bacterial case that was investigated by site-directed mutagenesis (SDM).<sup>19</sup> Mutations





**Figure 1** (A) Structure of selina-4(15),7(11)-diene synthase (SdS, PDB: 4OKZ) with highlighted Asp-rich motif and NSE triad. Active site architectures of (B) trichodiene synthase (TS, PDB: 1JFG) and (C) SdS. Fore-ground residues are labelled in black, background residues are shown in grey. Green spheres =  $\text{Mg}^{2+}$ . DHFPP = 2,3-dihydro-FPP.

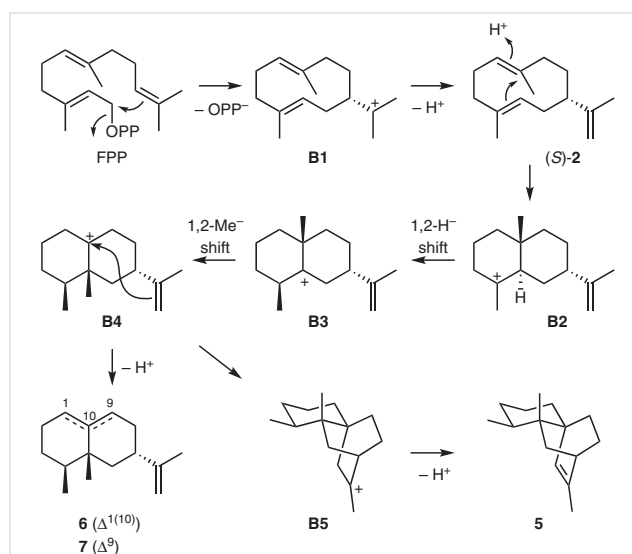


**Scheme 2** Cyclisation mechanism for pentalenene (**1**) by pentalenene synthase (PS)

within the Asp-rich motif of PS resulted for D80E and D81E in a strong loss of activity, but only minor effects were observed for the third Asp (D84E), in agreement with the  $\text{Mg}^{2+}$  complexation by the first two Asp residues (Figure 1C). For the D80E variant also a minor change in the product profile was found, leading to ca. 5% of the 1,10-cyclisation product (*R*)-(+)-germacrene A (**2**).<sup>20</sup> Similarly, for *epi*-isozizaene synthase (EIZS) from *Streptomyces coelicolor*,<sup>21</sup> for hedycaryol synthase (HcS) from *Kitasatospora setae*,<sup>22</sup> SdS<sup>14</sup> and both domains of the bifunctional geosmin synthase from *Streptomyces coelicolor* (SCO6073),<sup>23</sup> exchanges within the first two Asp residues disturbed catalytic activity. Among fungal enzymes, aristolochene synthase from *Penicillium roqueforti* (PrAS) showed strongly reduced or even abolished activity for exchanges within the first two Asp residues, but not the third position, which is naturally occupied by Glu and was exchanged against Asp or Gln.<sup>24</sup> Analogous observations were made for TS, the first fungal TPS investigated by SDM, for which all three Asp residues were tested by exchange with Glu.<sup>25</sup>

In rare cases, the Asp-rich motif is naturally altered, as found for the isoishwarane synthase from *Streptomyces lincolnensis* (IWS)<sup>26</sup> that catalyses the conversion of FPP into

isoishwarane (**5**) and minor amounts of the widespread biosynthetic intermediate (*S*)-(-)-**2**<sup>27</sup> and valencene (**6**, Scheme 3). IWS exhibits a <sup>80</sup>DDLHT sequence with the third



**Scheme 3** Cyclisation mechanism for isoishwarane (**5**) by isoishwarane synthase (IWS) and for aristolochene (**7**) by aristolochene synthase (PrAS and AtAS)

Asp missing, raising the question whether enzyme activity could be improved for the T84D variant. The experiment showed that for this variant the overall activity remained nearly unchanged, but the product profile was shifted with a decrease of **5** and increase of (*S*)-**2**.<sup>26</sup> For the sesterterpene synthase from *Streptomyces mobaraensis* (SmTS1), which produces seven known and three unknown sesterterpenes, the natural sequence deviation is even more dramatic and also affects the first Asp, which is usually lethal (<sup>86</sup>NDLTV). Also for this enzyme, the N86D variant had an activity similar to that of the wild type (104±9%) and is associated with moderate changes in the relative proportions of sesterterpenes.<sup>28</sup> Structural work on IWS and SmTS1 will be required to understand the contrasting effects for these enzymes.

## 2.2 The NSE Triad

Residues of the NSE triad have also been investigated by SDM. First investigations were made for pentalene synthase (PS), showing inactivity for the N219A and N219L variants and strongly reduced activity for N219D.<sup>20</sup> The analogous residue was also found to be critical for PrAS, for which the N244L variant was inactive and N244D activity was strongly reduced.<sup>24</sup> Also mutations within the conserved Ser or Glu of the NSE triad resulted in a reduced catalytic efficiency of PrAS, while the S248A/E252D double mutant was inactive. The importance of the NSE triad for catalytic activity was also studied for *Aspergillus terreus* aristolochene (**7**) synthase (AtAS, Scheme 3), showing reduced or abolished activity in several cases.<sup>24</sup>

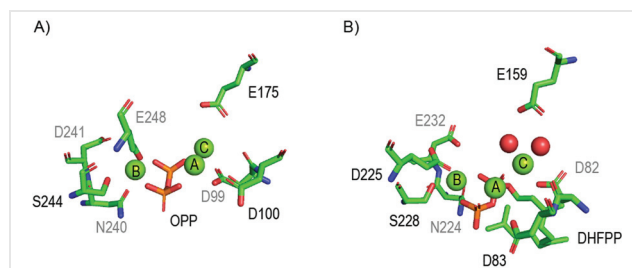
Additional work has been performed to mutate the NSE triad of EIZS, showing reduced activity (<5%) for the N240D, S244A and E248D variants.<sup>21</sup> The bifunctional geosmin synthase is an interesting case, because its N-terminal domain contains two NSE triads, but substitutions were only critical within the first one and within the NSE triad of the C-terminal domain.<sup>23</sup>

For the NSE triad a few inverse cases have also been studied in which natural substitutions of the otherwise highly conserved residues are found. For IWS (<sup>222</sup>NDLHSHLD) the second last position is occupied by Leu, but usually found to be Arg or Lys. The L229R variant and the double mutant T84D/L229R combining 'corrections' in the Asp-rich motif and the NSE triad showed retained activity, but a product shift towards **2**.<sup>26</sup> The sesterterpene synthase SmTS1 (<sup>226</sup>NQRYSYFKE) reveals a Gln instead of the usual Asp in the second position. In this case the Q227D variant showed a 1.5-fold increased catalytic activity, but no changes in the product profile.<sup>28</sup> For the diterpene synthase spiroalbatene synthase (SaS) from *Allokutzneria albata* (<sup>225</sup>NDVYGLEKD) the highly conserved Ser is exchanged by Gly, but the G229S variant could not be obtained as a soluble protein, suggesting that substitution of the conforma-

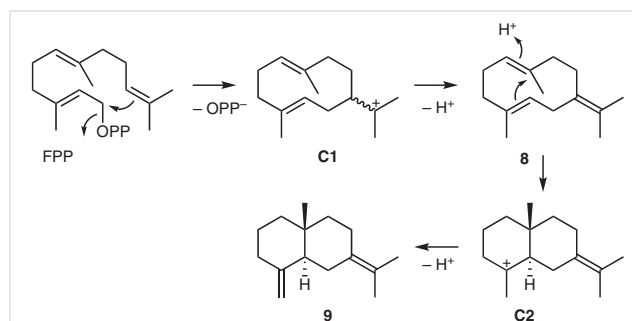
tionally flexible Gly in this position may prevent proper enzyme folding.<sup>29</sup> A crystal structure of the wild-type enzyme could confirm this hypothesis.

## 2.3 Another Conserved Glu for Mg<sup>2+</sup> Binding

Another highly conserved residue is a Glu that is usually located 19 positions upstream of the pyrophosphate sensor (the function of the pyrophosphate sensor is explained below). The structures of TS and SdS show its involvement in stabilisation of the trinuclear Mg<sup>2+</sup> cluster (Figure 2), with hydrogen bonds to two Mg<sup>2+</sup> coordinated water molecules in SdS that are not observed in TS.<sup>13,14</sup> The exchanges E159Q and E159D resulted in a loss of production of selina-4(15),7(11)-diene (**9**), but the biosynthetic intermediate germacrene B (**8**) was retained (Scheme 4).<sup>13</sup> Also for spata-13,17-diene synthase (SpS) from *Streptomyces xinghaiensis* the E184Q mutant was nearly inactive (0.5%).<sup>30</sup> For cembrene A synthase (CAS) from *Allokutzneria albata* the E179Q showed a strongly reduced activity (ca. 25%), while the E179K exchange disrupted activity.<sup>29</sup> Notably, this position is naturally filled with Gln in SaS and with Lys in bonnadiene synthase (BdS) from *Allokutzneria albata*.<sup>29,31</sup> SaS installation of the usually observed Glu in the Q160E enzyme variant resulted in a complete loss of activity,<sup>29</sup> while for



**Figure 2** A highly conserved Glu for stabilisation of the Mg<sup>2+</sup> cluster. Active sites of (A) TS (PDB: 1JFG) and (B) SdS (PDB: 4OKZ). Foreground residues are labelled in black, background residues are shown in grey. Green spheres = Mg<sup>2+</sup>, red spheres = H<sub>2</sub>O. DHFPP = 2,3-dihydro-FPP, OPP = diphosphate.



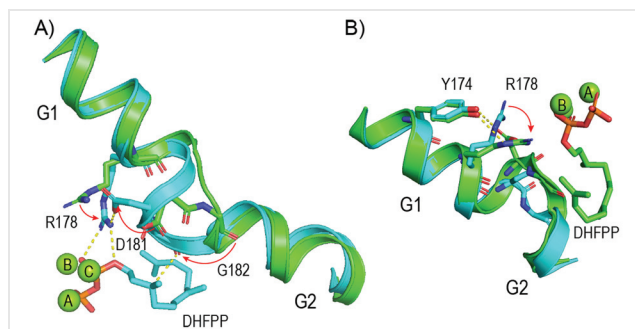
**Scheme 4** Cyclisation mechanism for selina-4(15),7(11)-diene (**9**) by selinadiene synthase (SdS)

BdS the result for the K169E mutant depended on the metal cofactor: activity was mostly retained with  $Mg^{2+}$  (75%), whereas with  $Mn^{2+}$  the activity was strongly decreased.<sup>31</sup> These results demonstrate that metal coordination in the active site of TPSs is finely balanced and can be disturbed significantly by minor structural changes in the position of E159 (SdS numbering).

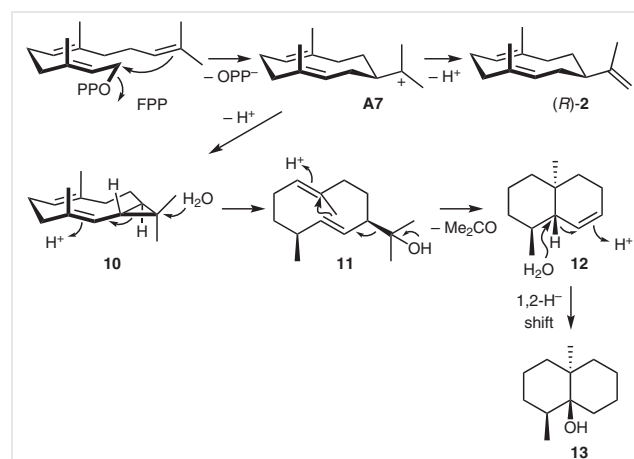
## 2.4 The Pyrophosphate Sensor and the Effector Triad

Further important active site residues that are highly conserved in microbial type I TPSs include the pyrophosphate sensor,<sup>14</sup> a single Arg usually located 46 positions upstream of the NSE triad, and the RY pair at the C-terminus, two conserved residues that reside in the 'basic motif' <sup>302</sup>DRRYR of trichodiene synthase.<sup>13</sup> The pyrophosphate sensor (R182 in TS and R178 in SdS, Figure 1) is located near a break of helix G<sup>15</sup> that undergoes a major conformational change upon substrate binding and active-site closure to form hydrogen bridges with the substrate's diphosphate (Figure 3A). Simultaneously, a polar residue (often Asp, Asn, Ser or Thr, the 'linker') installs hydrogen bridges to the pyrophosphate sensor, and a main chain carbonyl oxygen of a small residue (in many cases Gly, the 'effector') moves inwards and points directly to C3. The hydrogen bridges from the positively charged pyrophosphate sensor to the negatively charged substrate's diphosphate and the electron density of the effector donated into the  $\pi^*$  orbital of the allyl diphosphate trigger substrate ionisation.<sup>14</sup> These structural characteristics can be observed in the TEAS structure.<sup>15</sup> The dynamics of this process can be understood on the basis of a comparison of the open and the closed conformation of SdS, and by comparison to other enzyme structures, allowed for a generalised view on TPS catalysis.<sup>14</sup> The role of the effector triad R178-D181-G182 for catalysis by SdS was investigated by SDM. The R178Q variant showed a

complete loss of activity, and also R178K lost the ability to produce **9**, while some residual formation of **8** was still observed. Exchange of the linker (D181S, D181N) showed no effect, in agreement with the observation of Ser and Asn in the corresponding position of other TPSs. For germacradien-11-ol synthase (Gd11oIS), the N-terminal domain of geosmin synthase from *Streptomyces coelicolor*, the linker exchange V187N yielded the stable intermediate isolepidozene (**10**), while V187A gave moderately reduced activity with formation of germacradien-11-ol (**11**), and V187D was inactive (Scheme 5).<sup>32</sup> Exchange of the effector in SdS against a small residue (G182A) gave slightly reduced activity, with accumulation of the neutral intermediate **8**, while larger residues (G182V, G182P) resulted in inactivity.<sup>14</sup> For Gd11oIS, similar observations were made with G188A yielding the neutral intermediate **11** with strongly reduced activity, while G188V was inactive.<sup>32</sup> Thus, in both cases, exchange of the effector against Ala seems to prevent reprotonation of the neutral intermediate, while exchange against Val completely disrupts enzyme activity. Also for sestertermobaraene synthase, exchange of the effector (G184L) resulted in abolished production. This finding is particularly interesting because this enzyme naturally does not contain a pyrophosphate sensor, but a Gly instead, and its artificial introduction (G180R) also leads to a moderately reduced catalytic efficiency ( $51 \pm 16\%$ ).<sup>28</sup> The structural basis for a deeper understanding of these findings is currently lacking.



**Figure 3** Conformational change of SdS helix G upon active site closure. (A) Perspective showing movement of R178, D181 and G182. (B) Y174 hydrogen binds to R178, but does not move during active-site closure. Green: apo structure of SdS (4OKM), cyan: SdS structure in complex with three  $Mg^{2+}$  (green spheres) and DHFPF (4OKZ). Red arrows indicate movements within the effector triad R178, D181 and G182.



**Scheme 5** Cyclisation mechanism for geosmin (**13**) by the bifunctional geosmin synthase. The N-terminal domain is also known as germacradien-11-ol synthase (Gd11oIS)

A highly conserved Tyr is located four positions upstream of the pyrophosphate sensor. This residue forms hydrogen bridges to the pyrophosphate sensor, but does not show a conformational change during active-site closure (Figure 3B). This residue may be involved in stabilisation of the observed hydrogen-bond network, but is, in cases such

as SaS, naturally replaced with a Phe. However, the F175Y variant of SaS exhibits reduced catalytic efficiency (ca. 30% of wild-type level).<sup>29</sup>

## 2.5 The RY Pair in the Basic Motif

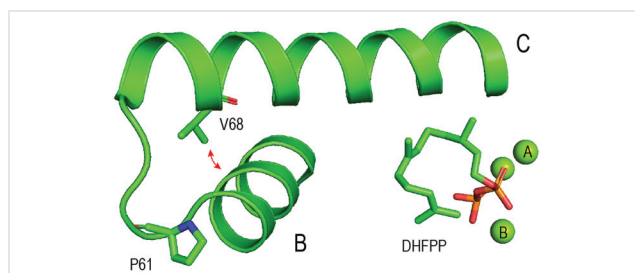
Mutational work has also been done for the ‘basic motif’<sup>30,2</sup> DRRYR of TS. Exchanges in the first two (D302N, D302E, R303I, R303E) or in the last position (R306K) were without significant effect; only for R306E was a reduced activity observed (14.3% of wild-type level). In contrast, mutations within the conserved RY pair (R304K, R304E, Y305T) resulted in drastically reduced activity, with the exception of Y305F showing 85.5% activity.<sup>33</sup> Similarly, for SdS the Y311L variant was inactive, while Y311F showed nearly wild-type activity and Y311W reduced activity.<sup>14</sup> For hedycaryol synthase (HcS) no changes in activity were observed for the R315K and Y316F varieties, showing that Lys can functionally substitute for Arg.<sup>22</sup> Overall, these findings demonstrate the importance of the observed hydrogen bonding between the substrate’s diphosphate and especially the Arg of the RY pair (Figure 1).

## 3 Residues with Structural Function

Microbial type I TPSs contain several highly conserved residues in specific positions, some of which are far from the active site and seem to have structural functions. They may serve as structural anchors that are important for installing the generally observed terpenoid synthase fold, while the amino acid sequences between these anchors may be much more flexible, which explains the poor overall sequence identity and simultaneous high structural conservation of the enzyme family.

### 3.1 The Conserved Pro and Leu in Helices B and C

Two highly conserved residues are a Leu and a Pro, 14 and 21 positions, respectively, upstream of the Asp-rich motif (Figure 4). In SdS, Pro61 is located at the N-terminal

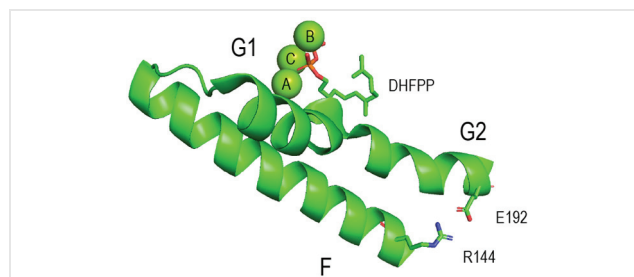


**Figure 4** A highly conserved Pro at the end of helix C and a highly conserved Leu (substituted in SdS by Val) at the beginning of helix D that serves as a steric spacer. Figure is based on the structure of SdS.<sup>14</sup> DHFPP = 2,3-dihydro-FPP, green spheres = Mg<sup>2+</sup>.

end of helix C and may cause a helix turn, while V68, with its steric bulk (substituting for the more common Leu), acts as a spacer between helices B and C. The corresponding residues have been mutated for SaS (P65A, L72A),<sup>29</sup> SpS (P83A, L90A),<sup>30</sup> BdS (P66A, V73A),<sup>31</sup> and phomopsene synthase (PmS) from *Allokutzneria albata* (P58A, L65A),<sup>31</sup> showing inactivity or at least strongly reduced activity in all cases, sometimes associated with poor yields during enzyme expression. In rare cases, the Pro is naturally replaced by another residue; for example, an Ala is found in CAS. However, the mutant A64P gave unclear results, as incubation experiments with Mg<sup>2+</sup>, but not with Mn<sup>2+</sup>, resulted in a moderately reduced catalytic efficiency.<sup>29</sup> Also for the (4*S*,7*R*)-germacra-(1(10)*E*,5*E*)-dien-11-ol synthase from the protist *Dictyostelium purpureum* (DpTPS9) a sequence deviation is observed, but the C60P, M67L and C60P/M67L variants did not show any effect on catalytic activity.<sup>34</sup> Taken together, these results demonstrate that for enzymes in which the usual Pro/Leu combination is established, these residues are important for proper enzyme folding and catalysis, but in cases where these residues are naturally substituted, the introduction of mutations has no or little effect.

### 3.2 The Arg-Glu Salt Bridge between Helices F and G

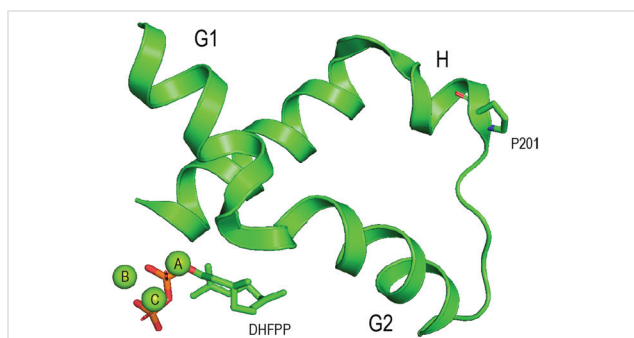
Another conserved structural motif in microbial type I TPSs is a salt bridge between an Arg at the beginning of helix F and a Glu at the end of helix G (Figure 5). Mutations within this salt bridge with chain shortening by an exchange of Arg with Lys or of Glu with Asp gave for BdS (R154K, E202D) and PmS (R137K) inactive enzymes,<sup>31</sup> and strongly reduced activity for SaS (R145K, E193D).<sup>29</sup> Also substitutions of Arg against Met were critical for BdS (R154M) and PmS (R137M),<sup>31</sup> whereas SaS (R145M) retained almost full activity.<sup>29</sup> In some cases the Glu of the salt bridge is naturally substituted by Asp, and the exchange of Asp against Glu can lead to improved enzyme efficiency, as observed for the D217E variant of SpS (173±54%).<sup>30</sup> However, for CAS (D212E) a strongly reduced activity was obtained,<sup>29</sup> showing that this is a promising, but not a generalised, approach to enhance catalytic function.



**Figure 5** A highly conserved Arg at the beginning of helix F and a highly conserved Glu at the end of helix G that form a salt bridge. Figure is based on the structure of SdS.<sup>14</sup> DHFPP = 2,3-dihydro-FPP, green spheres = Mg<sup>2+</sup>.

### 3.3 The Conserved Pro at the Entrance of Helix H

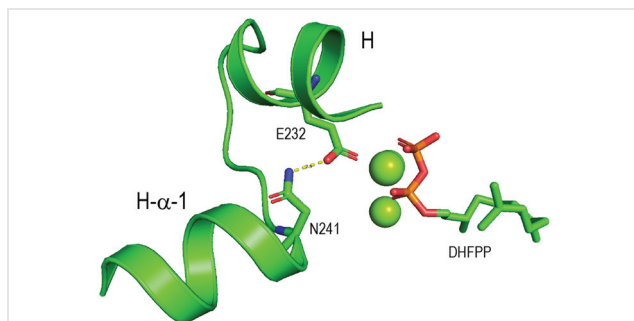
The SdS structure shows a highly conserved Pro at the beginning of helix H where a disordered region translates into a helical fold (Figure 6). This Pro residue is located ca. 21 positions downstream of the pyrophosphate sensor and is likely of structural relevance, although it is not observed in all microbial type I TPSs and is sometimes substituted by Thr, Ser or Ala. Mutation of the corresponding residue in spiroalbatene synthase (SaS) to Thr (P201T) yielded an enzyme with strongly decreased catalytic activity (ca. 15%).<sup>29</sup>



**Figure 6** The highly conserved P201 at the distorted N-terminal end of helix H. Figure is based on the structure of SdS.<sup>14</sup> DHFFP = 2,3-dihydro-FPP, green spheres = Mg<sup>2+</sup>.

### 3.4 The Conserved Asn in the H- $\alpha$ -1 Loop

In the closed conformation of SdS the metal coordinating E232 of the NSE triad forms a hydrogen bridge to N241 in the second coordination sphere (Figure 7). This residue is highly conserved in microbial type I TPSs and is usually located eight or nine positions downstream of the NSE triad and resides within the H- $\alpha$ -1 loop that undergoes a major conformational change during active-site closure.<sup>13</sup> This can also be observed for SdS, for which this region in the open conformation is disordered, and after active-site closure the interaction between E232 and N241 is established.<sup>14</sup>

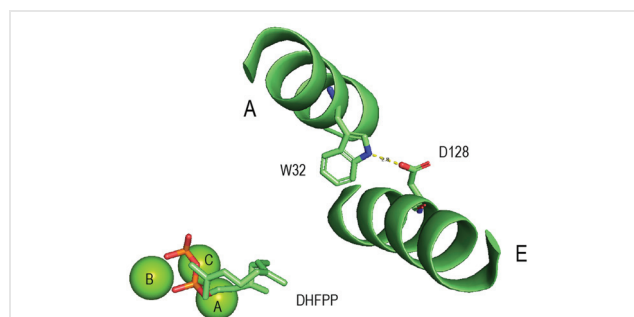


**Figure 7** The highly conserved N241 in the H- $\alpha$ -1 loop forms a hydrogen bridge to E232 of the NSE triad. Figure is based on the structure of SdS.<sup>14</sup> DHFFP = 2,3-dihydro-FPP, green spheres = Mg<sup>2+</sup>.

SDM of the conserved Asn has been performed for four enzymes from social amoebae, including the (-)- $\beta$ -barbatene synthase DdTPS9 from *Dictyostelium discoideum* (N236A, N236D),<sup>35</sup> and the synthases for (4*S*,7*R*)-germacra-(1(10)*E*,5*E*)-dien-11-ol (DpTPS9; N239A, N239D),<sup>33</sup> (-)- $\beta$ -araneosene (DpTPS10; N249A, N249D)<sup>36</sup> and (*S*)-(+)-naphthenol (DpTPS11; N243A, N243D)<sup>36</sup> from *Dictyostelium purpureum*. In almost all cases, no soluble protein was obtained, demonstrating a critical role for proper enzyme folding.<sup>34,35</sup> Only for DdTPS9 did the N236A variant turn out to be soluble, but showed a strongly reduced activity.<sup>35</sup> In the sestermobaraene synthase (SmTS1), an Arg residue is found instead of the conserved Asn, but for this enzyme the R242N variant did not show any changes in the product profile or activity.<sup>28</sup>

### 3.5 The Conserved Trp and Asp in Helices A and E

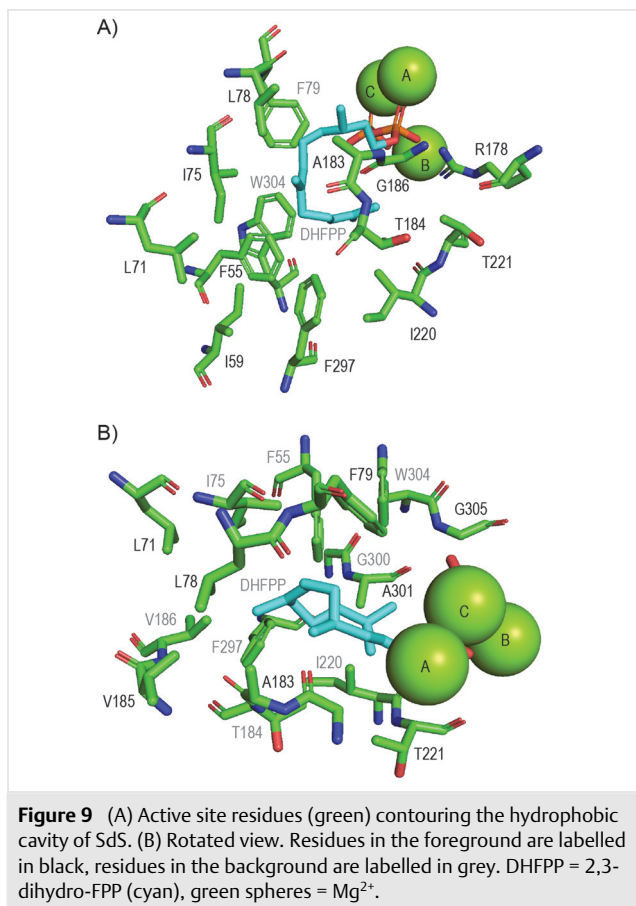
Two highly conserved motifs in microbial type I TPSs, W31 and D128 according to SdS numbering, residing on helices A and E, form a hydrogen bridge to each other (Figure 8). This interaction is likely of relevance for the overall enzyme structure and for correct enzyme folding. In some cases, the conserved Asp residue is naturally substituted by Asn, as observed for isoishwarane synthase (IWS). Notably, the N124D gave improved enzyme yields (142 $\pm$ 20% of wild-type level).<sup>26</sup>



**Figure 8** The highly conserved W32 on helix A and D128 on helix E hydrogen bridge to each other. Figure is based on the structure of SdS.<sup>14</sup> DHFFP = 2,3-dihydro-FPP, green spheres = Mg<sup>2+</sup>.

## 4 Residues Contouring the Active-Site Cavity

The active site of microbial type I TPSs is contoured by several non-polar and aromatic residues that serve as a template for the substrate, forcing it into a reactive conformation and thereby defining the reaction path that is entered after substrate ionisation. Aromatic residues may also be involved in the stabilisation of cationic intermediates through cation- $\pi$  interactions. Moreover, the hydrophobic cavity is important for the exclusion of water, which is a prerequisite to allow a cationic cyclisation cascade in an aqueous environment. Figure 9 gives an overview of the ac-

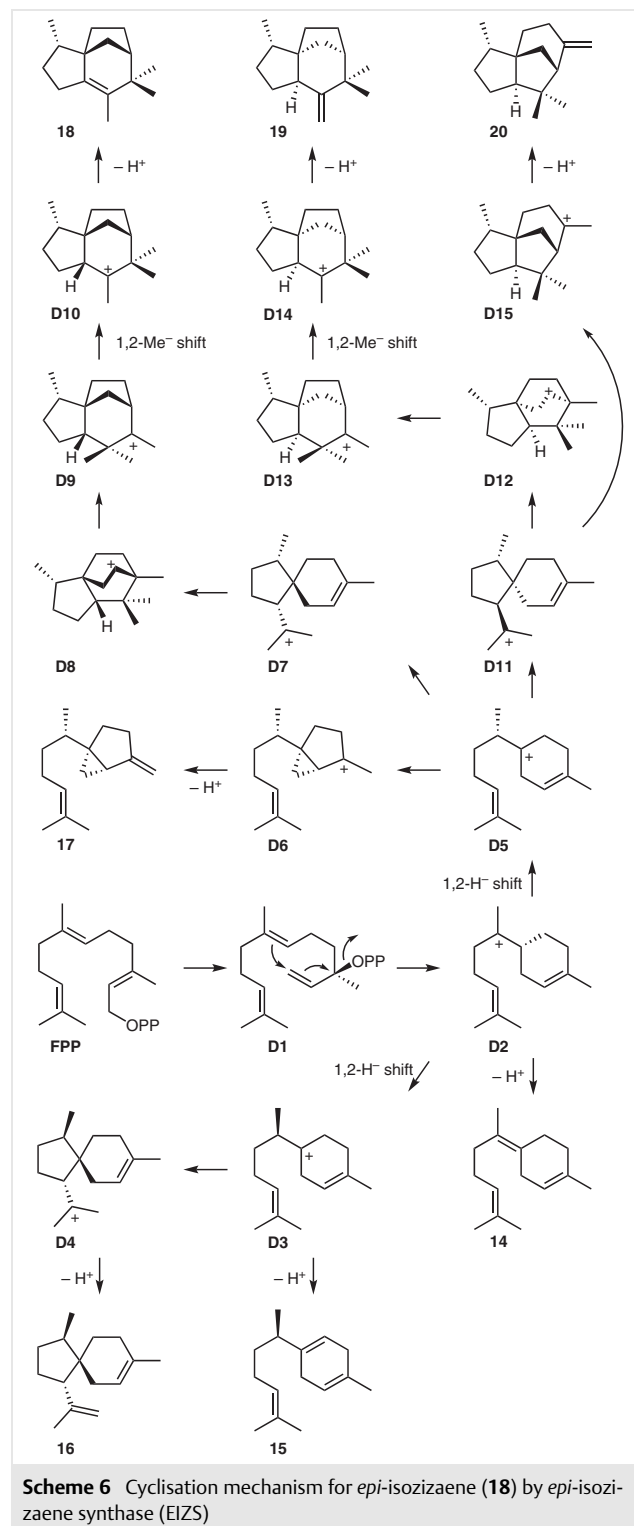


tive site residues lining the hydrophobic cavity of SdS. The corresponding residues in other characterised microbial type I TPSs have recently been tabulated and it has been demonstrated that their added calculated van der Waals volumes correlate negatively with the size of the preferred substrate.<sup>36</sup>

#### 4.1 Three and Four Positions Upfront of the Asp-Rich Motif

In one of the first mutational studies, Cane and co-workers noted two active-site Phe residues located on helix D of pentalenene synthase that are found three and four positions before the Asp-rich motif. The mutation of these residues (F76A, F77A; corresponding to L78 and F79 in Figure 9) only gave insoluble protein, while for the F77Y variant a 20-fold reduced catalytic activity was observed.<sup>20</sup> Recent structural work has demonstrated that F76 is involved in a C–H... $\pi$  interaction by which the intermediate secondary cations **A1** and **A2** are stabilised by hyperconjugation and the hydride shift between them is mediated (Scheme 2). SDM experiments demonstrated that F76Y can functionally substitute for the formation of **1**. With the F76W mutant a mixture of **1** and  $\alpha$ -humulene (**3**) was obtained, while

F76H gave a partial functional switch, resulting in a mixture of the anti-Markovnikov products **1** and **3**, besides the Markovnikov adduct (*S*)-**2**.<sup>37</sup>

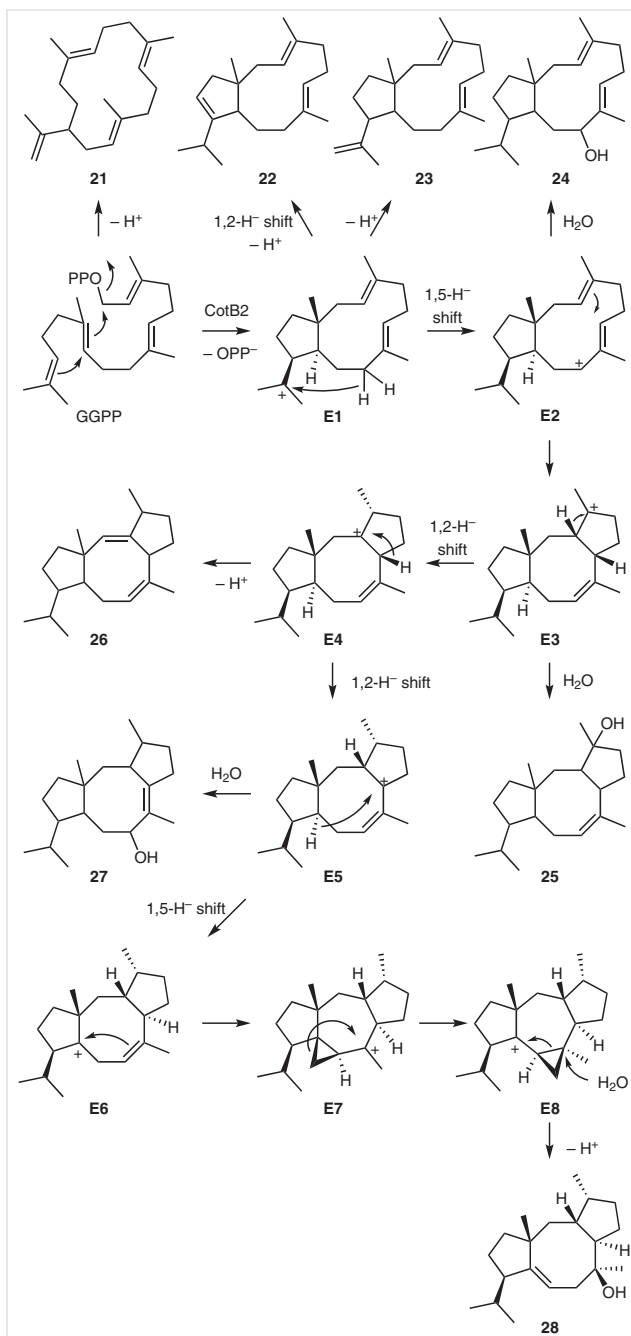


The analogous positions were also extensively tested for other microbial type I TPSs, including EIZS. Herein, the F96A variant gave a reduced activity and acyclic (*E*)- $\beta$ -farnesene as main product, suggesting that, in this mutant, cationic intermediates are not well stabilised.<sup>10</sup> This work was extended by investigations of a large number of variants in positions F95 (F95A, F95V, F95M, F95H, F95N, F95Y, F95Q and F95C) and F96 (F96V, F96L, F96Y, F96W, F96S, F96T, F96H, F96N, F96M and F96Q) that all showed a reduced catalytic activity and, in most cases, complex product mixtures. The wild type produces – depending on the incubation conditions – up to 99% *epi*-isozizaene (**18**, unless otherwise stated, the numbers given here and in the following sections do not refer to isolated yields, but are relative average percentages), which was still the main product only for the F95V, F95Y and F96Y variants (Scheme 6). Selective product formation was observed with some of the investigated mutants, yielding (*Z*)- $\gamma$ -bisabolene (**14**, 44% with F96V),  $\beta$ -curcumene (**15**, 50% with F95H),  $\beta$ -acoradiene (**16**, 68% with F95M), sesquisabinene A (**17**, 78% in F96S, 91% in F96M, and 97% in F96Q) and zizaene (**19**, 65% in F96W).<sup>38,39</sup>

Structural data for SdS suggested the involvement of the corresponding F79 in a cation stabilisation at C3 of FPP. The F79L variant resulted in the formation of farnesenes, besides **8** and **9** (Scheme 4), while F79Y and particularly F79W showed a product shift from **9** towards **8**.<sup>14</sup> For the F77I variant of isoishwarane synthase a reduced activity and shifted production from **5** towards intermediate (*S*)-**2** was observed (Scheme 3),<sup>26</sup> and production of farnesenes and the intermediate **2** has also been reported for PrAS (F112A).<sup>40</sup> Taken together, these data demonstrate that in all three cases the Phe residue is of importance for the second cyclisation of the neutral intermediate to yield the final product. The (4*S*,7*R*)-germacra-(1(10)*E*,5*E*)-dien-11-ol synthase DdTPS9 catalyses only a 1,10-cyclisation, but no reprotonation-induced secondary cyclisation. Accordingly, for this enzyme, F78V and F78I only showed reduced activity, but no change in the product profile.<sup>34</sup>

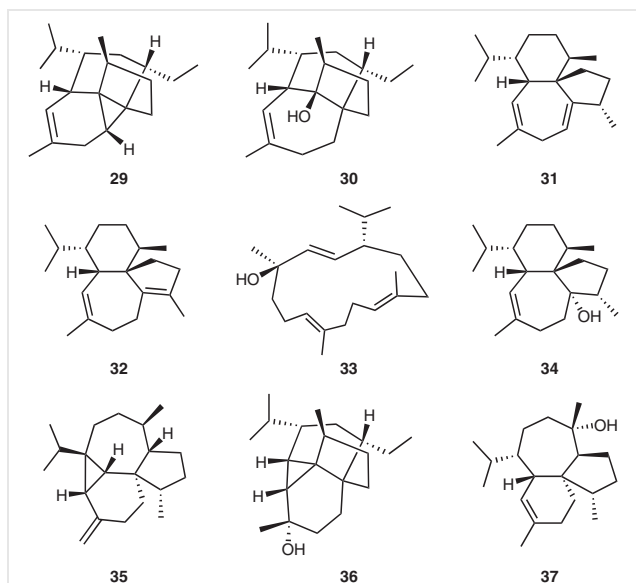
Similar findings regarding the importance of a Phe residue directly upstream of the Asp-rich motif have been reported for cyclooctat-9-en-7-ol synthase (CotB2) from *Streptomyces melanosporofaciens*.<sup>41</sup> This enzyme catalyses the multistep cascade from GGPP to the diterpene cyclooctat-9-en-7-ol (**28**) with involvement of an interesting cyclopropylcarbanyl cation rearrangement from **E7** to **E8** (Scheme 7).<sup>42</sup> For CotB2, the F107A and F107G variants resulted in the formation of cembrene A (**21**), while the F107Y exchange yielded cyclooctat-1,7-diene (**26**) and another unidentified diterpene,<sup>43</sup> and F107L gave, besides **28**, several new compounds including 3,7-dolabelladiene-9-ol (**24**) and cyclooctat-6-en-8-ol (**27**).<sup>44</sup>

A recently characterised multiproduct diterpene synthase from *Cryptosporangium polytrichastri* (CpPS) catalyses conversion of GGPP into the main product polytrichas-



**Scheme 7** Cyclisation mechanism for cyclooctat-9-en-7-ol (**28**) by cyclooctat-9-en-7-ol synthase (CotB2)

trene A (**29**), besides polytrichastrol A (**30**), wanju-2,5-diene (**31**), wanju-2,6-diene (**32**) and thunbergol (**33**, Figure 10). In this enzyme the positions corresponding to L78 and F79 in SdS are occupied by small residues (G90 and A91). For the G90F exchange a reduced activity (23±2%) with abolished production of **29** and **31** as main product was observed, suggesting that the introduction of steric bulk in the active site of this enzyme hampers GGPP acceptance and ef-



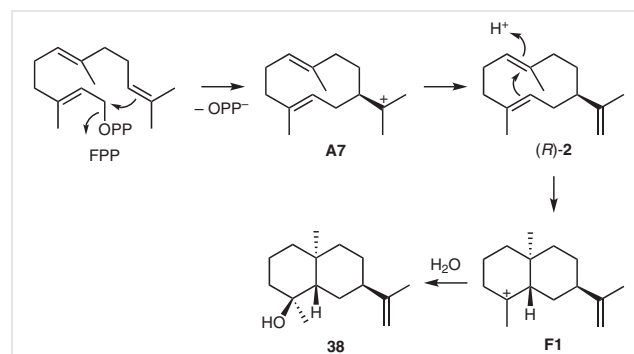
**Figure 10** Structures of diterpene products obtained from wild-type and mutant *Cryptosporangium polytrichastri* polytrichastrene A synthase (CpPS).

ficient catalysis.<sup>36</sup> In the 2-methylisoborneol synthase (MIBS) from *Streptomyces coelicolor* two polar residues are observed in the analogous positions, E193 and N194, with especially the residue E193 being conserved in many MIBSs. However, none of the mutants E193A, E193L, and E193D showed significant effects on the production of 2-methylisoborneol.<sup>45</sup>

#### 4.2 A Conserved Trp Upstream of the RY pair

Another extensively tested active-site residue is a highly conserved tryptophan (W304 in SdS, showing  $\pi$  stacking with F55, Figure 9) located six or seven positions upstream of the C-terminal RY pair. For PS, the W307F mutant showed almost completely retained activity with respect to the production of **1** and minor amounts of (*R*)-**2** (Scheme 2).<sup>20</sup> In contrast, for PrAS, a strongly reduced catalytic efficiency was reported for the W334F and also the W334Y, W334H, W334V and W334L variants, but similar observations were made regarding the product spectrum that shifted from **7** to **2** in all cases, especially for the exchanges against aliphatic residues yielding >95% **2**.<sup>46,47</sup> Also non-canonical substitutions (*p*-substituted Phe derivatives and the naphthyl analogue) have been introduced in this position, with similar effects on enzyme efficiency and product distribution.<sup>47</sup> With the SdS variant W304F and W305Y only small effects were observed, while the W304L substitution resulted in an almost completely abolished production of **9**, but the formation of **8** was retained.<sup>14</sup> Also for the W335F mutant of (+)-intermedeol synthase (STC4) from the basidiomycete *Termitomyces* sp. the cyclisation of FPP stopped at

the intermediate (*R*)-**2** and did not proceed to the wild-type product intermedeol (**38**, Scheme 8). In contrast, the W314F variant of the (–)- $\gamma$ -cadinene synthase (STC9) from the same fungus showed no effect, while the W314F substitution of the (+)-germacrene D-4-ol synthase (STC15) did not yield soluble enzyme. Also the exchanges of the conserved Trp against Ala was critical for correct enzyme folding in all three cases.<sup>48</sup> For hedycaryol synthase (HcS) none of the exchanges W309F, W309Y and W309L proved to be critical, which is in line with the observation that also for other enzymes with mutations in this position, the ability to catalyse 1,10-cyclisations is retained.<sup>22</sup> With the W288G mutant of CotB2 the ability to catalyse the initial 1,11-10,14-cyclisation was retained, leading to the formation of dolabella-3,7,18-triene (**23**, Scheme 7), but no downstream products requiring reprotonation-induced secondary cyclisations were observed.<sup>49</sup> For sesterterpene synthase (SmTS1), the analogous position is naturally occupied by Phe. For this enzyme the F307W exchange resulted in a decreased yield (53±6% of wild-type level) with unchanged product distribution, suggesting that the introduction of steric bulk moderately blocks substrate acceptance by this sesterterpene synthase.<sup>28</sup>



**Scheme 8** Cyclisation mechanism for (+)-intermedeol (**38**) by (+)-intermedeol synthase (STC4)

Taken together, these mutational studies on the highly conserved C-terminal Trp indicate its general importance for secondary cyclisations of neutral intermediates, while this residue does not seem to be relevant for the initial cyclisation reactions performed with the oligoprenyl diphosphate. In particular, its substitution against another aromatic residue can result in a retained catalytic activity also for secondary cyclisations, but exchange against aliphatic residues is usually critical for this step.

#### 4.3 The Neighbourhood of the Conserved Trp

Several studies have also addressed active-site residues in the direct neighbourhood of this highly conserved Trp, but these residues are less conserved in microbial type I TPSs. In pentalene synthase, H309, adjacent to W308,



protrudes into the active site<sup>50</sup> and was suspected to act as the catalytic base in the deprotonation of **A1** to **3** (Scheme 2).<sup>19</sup> Following this hypothesis, in a next step, the same proton would be re-introduced to form **A2** for the downstream steps towards **1** (as discussed above, recent investigations<sup>37</sup> favour a direct proton transfer from **A1** to **A2** mediated by F76). However, SDM yielding the enzyme variants H309A, H309S, H309C and H309F revealed a substantially retained enzyme activity in all cases, with strongest decreases for the H309F variant. Notably, besides the production of **1**, minor amounts of (*R*)-**2** and protoillud-6-ene (**4**) were detected for all four mutants, which is explainable by deprotonation of intermediate **A3**.<sup>19</sup> The formation of **4** also indicated that **A3** could be an intermediate towards **1**, which was confirmed by DFT calculations in conjunction with an isotopic labelling study in which the proton that is removed from **A3** towards **4** was substituted by deuterium. The H309A mutant revealed a pronounced kinetic isotope effect for the distribution of the products **1** and **4**, which supported **A3** as an on-path intermediate towards both products.<sup>51,52</sup>

In contrast to the findings with PS, mutation of the His in the analogous position of hedycaryol synthase (H310S) resulted in an inactive enzyme. Based on structural data for HcS it was suggested that H310 in the protonated state assists in the abstraction of the negatively charged diphosphate.<sup>22</sup> For Gd11oIS, the corresponding H320 has recently been described as part of an 'RQH site', with residues R228, Q313, and H320 being located adjacent to C11 of FPP. Especially, the mutation of H320 gave pronounced effects with significantly reduced activity and a product shift towards isolepidozene (**10**, Scheme 5).<sup>32</sup> In the corresponding position of EIZS a Phe residue is found. Its exchange (F332A) showed a strongly reduced catalytic activity with formation of 37% **19** (Scheme 6).<sup>38</sup> In contrast, for the (–)- $\gamma$ -cadinene synthase STC9, the mutations Y315G and Y315S revealed no effect on the product spectrum and activity.<sup>48</sup>

Active-site residues two and three positions upstream of the conserved C-terminal Trp were also investigated for EIZS and STC9, respectively. The EIZS variants V329G, V329I and V329M, but not V329A, exhibited a reduced activity with **18** as main product in all cases, and 33% production of **19** for V329M.<sup>38</sup> For STC9 from *Termitomyces* sp. the C311N and C311S substitutions resulted in a moderate and C311H in a strong reduction of enzyme activity, but unchanged product profile.<sup>48</sup>

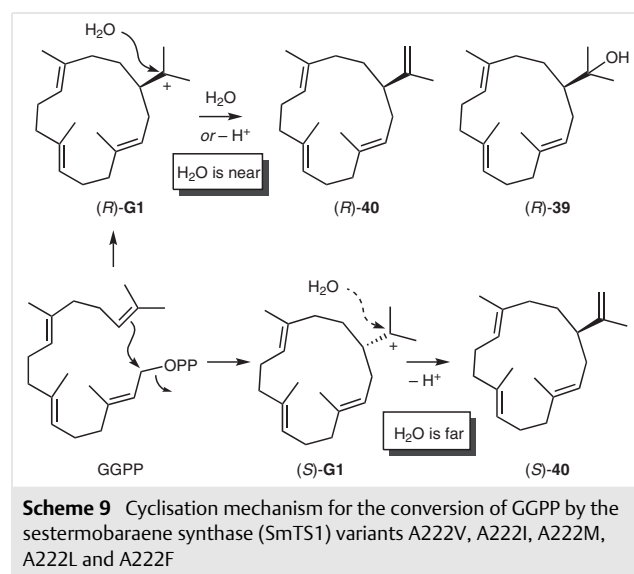
#### 4.4 Thirteen Positions Upstream of the RY Pair

Microbial type I TPSs usually contain an aromatic or aliphatic residue thirteen positions in front of the RY pair (equal to F297 in SdS, Figure 9). The corresponding residue in EIZS was modified by SDM to obtain the W325F mutant, which exhibited lowered catalytic efficiency with forma-

tion of 50% **18** and 45% **19**.<sup>38</sup> For the germacradien-11-ol synthase (Gd11oIS) the corresponding position was addressed by W312A, leading to significant formation of the acyclic compound nerolidol, while W312F gave **2** as the main product. Both variants showed moderately reduced activity.<sup>32</sup>

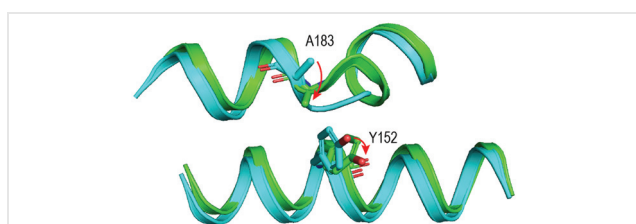
#### 4.5 Three and Four Positions Upfront of the NSE Triad

Two more hydrophobic active-site contouring residues are located three and four positions upstream of the NSE triad, represented by I220 and T221 in SdS (Figure 9). The residue A236 of EIZS, corresponding to I220 of SdS, was investigated in detail, showing a strongly reduced activity for A236F and inactivity for A236M. These data suggested that this position should be occupied by a small residue for efficient catalysis, and, accordingly, the A236G variant retained almost full catalytic potential as compared to the wild type.<sup>38</sup> Particularly interesting results were obtained with the sesterterpene synthase SmTS1. Exchange of A222 showed a reduced (A222M, A222L) or completely abolished activity (A222V, A222I, A222F, A222Y and A222W) with GGPP as substrate, demonstrating that the introduction of steric bulk in this position is also critical for sesterterpene synthase activity. However, the exchanges A222V, A222I, A222M, A222L and A222F opened a new reactivity towards GGPP that was not observed with the wild type, leading to the production of cembrene A (**21**) of varying enantiomeric compositions, but always enantiomerically pure (*R*)-naphthenol (**39**) in the five variants. These findings are explainable by two different reaction paths through the (*R*)- and (*S*)-cembranyl cation (**G1**) with an active-site water that is close to the cationic centre only for (*R*)-**G1** (Scheme 9).<sup>28</sup>



#### 4.6 The Region Downstream of the Pyrophosphate Sensor

Another series of hydrophobic active-site residues is located 5, 6, 8 and 9 positions after the pyrophosphate sensor (A183, T184, V186 and V187 in SdS, Figure 9). SDM experiments on these residues were performed for several TPSs. Based on structural insights for this enzyme, the EIZS variants W203F,<sup>10</sup> W203Y and W203H were addressed.<sup>39</sup> All three substitutions gave reduced enzyme yields, with **14** as main product, ranging from 38% in W203H to 47% in W203F (Scheme 6).<sup>10,39</sup> Furthermore, F198A was investigated, showing reduced activity and a complex product pattern.<sup>10</sup> Also F198V, F198L and F198Y exhibited a diminished catalytic performance, with formation of a complex product mixture by F198V, while F198L yielded 61% of  $\beta$ -cedrene (**20**) and F198Y produced **18** with high selectivity (89%). For HcS, the variants M181H and M181K were made available, with M181K showing inactivity, whereas M181H was active at pH 8.5, but not at pH 7.0. The HcS crystal structure reveals that M181 is close to C3 of the ligand nerolidol, suggesting that a positively charged residue in position 181 can destabilise the allyl cation formed upon diphosphate abstraction from FPP. While Lys in the M181K mutant is protonated at pH 7.0 and 8.5, the less basic His in the M181H mutant is protonated at pH 7.0, but not at 8.5.<sup>22</sup> For SdS, A183 was investigated. This residue undergoes a movement towards Y152 during active-site closure, which is pushed away (Figure 11). The introduction of steric hindrance in this position (A183V, A183F) destroys enzyme function, while the alterations Y152L, Y152F and Y152W only influence the product ratio (**8/9**), but catalytic activity is retained.



**Figure 11** Plasticity of A183 and Y152 in SdS. During the change from the open (cyan) to the closed conformation (green) A183 moves towards Y152, which is pushed away.

For CotB2, residues F185 and W186, corresponding to V185 and V186 of SdS (Figure 9), were examined by SDM, leading to the formation of interesting new products. The F185A and W186F variants yielded original **28**, in addition to **24** and **27**. The W186H mutant gave **28**, cyclooctat-7-en-3-ol (**25**) and **23**, and the W186L mutant produced **28**, **21** and **23**.<sup>44</sup>

The bacterial polytrichastrene synthase (CpPS)<sup>36</sup> from *Chryseobacterium polytrichastri* and the wanjudiene synthase from *Chryseobacterium wanjuense* (CwWS)<sup>53</sup> differ in

only two of their active site residues: A192 and A87 of CpPS are altered to T86 and V191 in CwWS (=T184 and I75 in SdS, Figure 9). Besides the main product **29**, the wild type of CpPS also produces **31**, the main product of CwWS. Interestingly, for the A87T, A192V and double A87T/A192V variants of CpPS a shift towards **31** as main product was observed, which impressively demonstrates the role of active-site residues for determining the product outcome of TPSs.<sup>36</sup>

#### 4.7 Another Aromatic Residue in Front of the Pyrophosphate Sensor for Cation- $\pi$ -Stabilisation

The crystal structure of CotB2 revealed F149 as an active-site residue.<sup>49</sup> Mutation of this residue resulted for the F149L, F149V and F149H variants in the formation of **25** as the only product.<sup>43,44</sup> Comparison of the wild-type and F149L X-ray structures showed a high structural similarity of both enzymes, but different water networks,<sup>49</sup> which may explain the production of different terpene alcohols **25** versus **28** by these enzymes. The structure of PrAS shows that the active-site residue F178 is ideally oriented for stabilisation of the eudesmane cation **B2** (Scheme 3).<sup>54</sup> SDM yielded the enzyme variants F178Y, F178W, F178V, F178I and F178C, which all exhibited reduced activity, especially for the cases in which F178 was exchanged against a non-aromatic residue. For the F178Y and F178W mutants, **7** was still the main product, while the F178V, F178I and F178C exchanges gave only minor amounts of **7**, but mainly (*S*)-**2** instead, and also substantial amounts of **6** (20%) for F178C.<sup>40,55</sup> These findings are in line with the suggested cation- $\pi$  stabilisation of the downstream intermediate **B2** by the F178 residue, but its involvement in a mediation of the hydride shift from **B2** to **B3** has also been proposed.<sup>55</sup>

#### 4.8 Further Hydrophobic Residues in Front of the Asp-Rich Motif

Several more hydrophobic active-site residues are found upstream of the Asp-rich motif, including for SdS the positions of F55, I59, L71 and I75 (Figure 9). Particularly interesting is the position of F55. This residue has been tested for SdS itself, leading to a product shift from **9** to the neutral intermediate **8** for all investigated cases of F55L, F55Y and F55W (Scheme 4). Noteworthy is also the production of farnesenes by the F55L variant, suggesting a role of aromatic F55 in cation stabilisation.<sup>14</sup> In contrast, no effect was observed for the Y169F mutant in the analogous position of MIBS.<sup>45</sup> In other enzymes, aliphatic residues are located in this position; for example, PrAS exhibits a V88 here. Its alteration (V88A) gave a reduced enzyme efficiency and a moderate product shift towards intermediate (*S*)-**2** (12%). Swapping to an aromatic residue (V88F) also reduced the enzyme activity, but had more pronounced effects on the product distribution, yielding only 18% **7**, but 58% **6** and

24% (*S*)-**2** (Scheme 3).<sup>56</sup> EIZS contains an L72 in this position. Its exchange showed a significantly disturbed enzyme performance for L72A and L72M, while the activity for the structurally more conservative L72V exchange was only moderately reduced. In all three variants, **18** was retained as main product, but the L72V mutant also produced 38% of an unidentified sesquiterpene.<sup>38</sup> CpPS also contains an aliphatic I66 residue. Its exchange in the I66F variant gave particularly interesting results, because this mutant not only showed an increased catalytic efficiency (195±43% of wild-type activity) with production of the diterpenes **29**, **31** and **33** as also obtained with the wild type, but also several more compounds of high structural complexity including **32** and **34–37** (Figure 10).<sup>36</sup> In HcS the analogous position is taken by a polar residue (S55) that was challenged by the S55W exchange, resulting in an inactive enzyme, likely because the active site was blocked by the large Trp residue.<sup>22</sup>

The next hydrophobic active-site residue is located four positions downstream, observed as I59 in SdS (Figure 9). In PrAS, this position is filled by Y92 that has been investigated extensively by SDM. Based on the finding that in the Y92F mutant the relative production of germacrene A (**2**) is higher than in the wild type, it has been suggested that this residue acts as the catalytic acid in the reprotonation of **2** for its further conversion into **7**.<sup>57</sup> but this interpretation can be questioned by the fact that **7** is still the main product in this mutant.<sup>24</sup> As for the Y92F variant, also the Y92V, Y92A and Y92C mutants exhibited strongly reduced catalytic activity. Notably, the smaller the residue in this position, the more acyclic products (farnesenes) were obtained, with completely abolished production of **7** in the Y92A variant.<sup>58,59</sup>

The position analogous to I75 of SdS (Figure 9) has been investigated in PrAS (L108) and CotB2 (N103). For PrAS, the mutants L108A, L108S, L108V and L108F have been generated, which all gave lowered enzyme activity. In this case also, the exchange against small residues (L108A, L108S) resulted in the substantial formation of farnesenes, suggesting that the enlarged space in the active site of these variants does not enforce a proper substrate folding, with the consequence that the substrate becomes ionised by diphosphate abstraction, but cannot react by cyclisation; instead, simple deprotonation products are obtained. In contrast, L108V and L108F only showed moderately changed product profiles.<sup>56</sup> Structural data imply that the unusual polar residue N103 in CotB2 is involved in a dipole–charge interaction of the side-chain carbonyl oxygen with C8 of the substrate analogue geranylgeranyl thiolodiphosphate (GGSP), suggesting stabilisation of the intermediate **E2** by this residue. Accordingly, the N103A mutant resulted in the formation of dolabella-3,7,12-triene (**22**); that is, **E2** is not formed in this mutant.<sup>44</sup>

Taken together, the mutagenesis experiments on hydrophobic active-site residues show very interesting effects. While mutations on residues directly involved in catalysis

and substrate binding (Chapter 2) often only give inactive enzymes, the residues of the hydrophobic cavity are the playground of TPS catalysis. These residues determine the size of the active site and the conformational fold of the substrate, and many experiments have demonstrated that swapping of these residues can result in changed substrate folding and consequently the formation of new products. Even the substrate specificity may be altered, if the cavity is enlarged (by exchange of large against small residues) or narrowed (by the opposite changes).

## 5 Other Residues

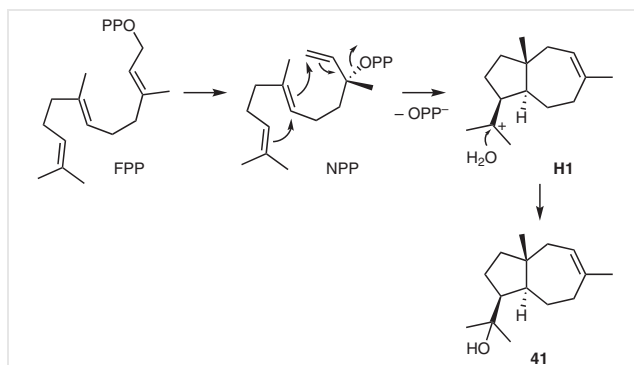
This section will summarise mutational studies on residues that do not fall under the main categories described above. For trichodiene synthase (TS), several enzyme variants were investigated before its structure had been clarified.<sup>12</sup> Treatment of TS with methyl methanethiosulfonate resulted in an inactivation of the enzyme, which was interpreted as showing that active-site Cys residues were involved in catalysis. Their chemical modification would then lead to a loss of function. Although it is today clear that these residues lie outside the active site, the C146A and C190A variants gave reduced activity, with retained production of trichodiene by C190A.<sup>33</sup>

### 5.1 Residues Involved in a Water Network in AtAS

For AtAS, several residues near the active site were found to form hydrogen bridges with water molecules, including Q151, N299 and S303. Since this enzyme only produces terpene hydrocarbons, but no alcohols, there is no obvious catalytic role for these water molecules. Mutagenesis of these residues gave only moderate effects on catalytic performance, but resulted in the formation of farnesol, nerolidol and germacrene A (**2**) in several cases. Thus, the network between these residues and active-site waters are important for template formation and maybe a tight control of the water molecules that does not allow them to quench the cationic intermediates during terpene cyclisation.<sup>60</sup>

### 5.2 Surface Residues of DcS

Dauc-8-en-11-ol synthase (DcS) from *Streptomyces venezuelae* catalyses an unusual 1,7-6,10-cyclisation of FPP to dauc-8-en-11-ol (**41**, Scheme 10). For this enzyme, three surface residues (H58, H197 and Y241) were identified by enzyme homology modelling and investigated by SDM. Interestingly, the H197F/Y241F double mutant exhibited a 4-fold higher enzyme yield that turned out to be preparatively useful (46% isolated yield of **41** from 80 mg FPP, in comparison to 11% obtained with the wild type).<sup>61</sup> It is possible that these exchanges cause some structural rearrangement in the mutant, because the non-polar Phe residues may turn



**Scheme 10** Cyclisation mechanism for dauc-8-en-11-ol (**41**) by dauc-8-en-11-ol synthase (DcS).

from the surface towards the inside of the enzyme. Structural work on DcS and the double mutant will be required for a detailed understanding of the observed effects.

## 6 Conclusions

As summarised in this review article, site-directed mutagenesis experiments have been performed on many microbial type I terpene synthases. Depending on the residues addressed in these experiments, different effects can be observed. If residues are exchanged that are directly involved in substrate binding and ionisation, for example residues of the Asp-rich motif or the NSE triad, this usually leads to inactivation. These experiments deepened our understanding of the mechanisms of TPS catalysis, but are not preparatively useful. The same is true for exchanges of residues that are important for the enzyme structure and proper folding. As reviewed here, several residues have been identified based on structural work that are highly conserved and that stabilise TPS structures, for example by specific interactions between two helices. Contrary to these findings, TPSs contain several hydrophobic residues that contour the active site. They determine substrate folding, and if exchanges are made within these residues, often new products are formed. Therefore, these residues can be regarded as a playground for chemists and biochemists to generate new biocatalysts that open the way to interesting and sometimes even novel terpenes. Several examples were also reported for which an improved catalytic efficiency with higher enzyme yield has been obtained by site-directed mutagenesis. Future research will hopefully deepen our understanding of the complex reactions catalysed by TPSs so that tailored catalysts can be designed for the specific enzymatic synthesis of novel terpenes.

## Conflict of Interest

The authors declare no conflict of interest.

## References

- (1) Gershenzon, J.; Dudareva, N. *Nat. Chem. Biol.* **2007**, *3*, 408.
- (2) Schwab, W.; Fuchs, C.; Huang, F.-C. *Eur. J. Lipid Sci. Technol.* **2013**, *115*, 3.
- (3) Dickschat, J. S. *Nat. Prod. Rep.* **2016**, *33*, 87.
- (4) Christianson, D. W. *Chem. Rev.* **2017**, *117*, 11570.
- (5) Wendt, K. U.; Schulz, G. E. *Structure* **1998**, *6*, 127.
- (6) Qu, G.; Li, A.; Acevedo-Rocha, C. G.; Sun, Z.; Reetz, M. T. *Angew. Chem. Int. Ed.* **2020**, *59*, 13204.
- (7) Arnold, F. H. *Angew. Chem. Int. Ed.* **2019**, *58*, 14420.
- (8) Xu, J.; Cen, Y.; Singh, W.; Fan, J.; Wu, L.; Lin, X.; Zhou, J.; Huang, M.; Reetz, M. T.; Wu, Q. *J. Am. Chem. Soc.* **2019**, *141*, 7934.
- (9) Li, D.; Wu, Q.; Reetz, M. T. *Methods Enzymol.* **2020**, *643*, 225.
- (10) Aaron, J. A.; Lin, X.; Cane, D. E.; Christianson, D. W. *Biochemistry* **2010**, *49*, 1787.
- (11) Tarshis, L. C.; Yan, M.; Poulter, D. C.; Sacchettini, J. C. *Biochemistry* **1994**, *33*, 10871.
- (12) Cane, D. E.; Kang, I. *Arch. Biochem. Biophys.* **2000**, *376*, 354.
- (13) Rynkiewicz, M. J.; Cane, D. E.; Christianson, D. W. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 13543.
- (14) Baer, P.; Rabe, P.; Fischer, K.; Citron, C. A.; Klapschinski, T. A.; Groll, M.; Dickschat, J. S. *Angew. Chem. Int. Ed.* **2014**, *53*, 7652.
- (15) Starks, C. M.; Back, K.; Chappell, J.; Noel, J. P. *Science* **1997**, *277*, 1815.
- (16) Cane, D. E.; Sohng, J.-K.; Lamberson, C. R.; Rudnicki, S. M.; Wu, Z.; Lloyd, M. D.; Oliver, J. S.; Hubbard, B. R. *Biochemistry* **1994**, *33*, 5846.
- (17) Gutta, P.; Tantillo, D. J. *J. Am. Chem. Soc.* **2006**, *128*, 6172.
- (18) Zu, L.; Xu, M.; Lodewyk, M. W.; Cane, D. E.; Peters, R. J.; Tantillo, D. J. *J. Am. Chem. Soc.* **2012**, *134*, 11369.
- (19) Seemann, M.; Zhai, G.; Umezawa, K.; Cane, D. E. *J. Am. Chem. Soc.* **1999**, *121*, 591.
- (20) Seemann, M.; Zhai, G.; de Kraker, J.-W.; Paschall, C. M.; Christianson, D. W.; Cane, D. E. *J. Am. Chem. Soc.* **2002**, *124*, 7681.
- (21) Lin, X.; Cane, D. E. *J. Am. Chem. Soc.* **2009**, *131*, 6332.
- (22) Baer, P.; Rabe, P.; Citron, C. A.; de Oliveira Mann, C. C.; Kaufmann, N.; Groll, M.; Dickschat, J. S. *ChemBioChem* **2014**, *15*, 213.
- (23) Jiang, J.; He, X.; Cane, D. E. *Nat. Chem. Biol.* **2007**, *3*, 711.
- (24) Felicetti, B.; Cane, D. E. *J. Am. Chem. Soc.* **2004**, *126*, 7212.
- (25) Cane, D. E.; Xue, Q.; Fitzsimons, B. C. *Biochemistry* **1996**, *35*, 12369.
- (26) Xu, H.; Rinkel, J.; Dickschat, J. S. *Org. Chem. Front.* **2021**, *8*, 1177.
- (27) Xu, H.; Dickschat, J. S. *Chem. Eur. J.* **2020**, *26*, 17318.
- (28) Hou, A.; Dickschat, J. S. *Beilstein J. Org. Chem.* **2021**, *17*, 2441.
- (29) Rinkel, J.; Lauterbach, L.; Rabe, P.; Dickschat, J. S. *Angew. Chem. Int. Ed.* **2018**, *57*, 3238.
- (30) Rinkel, J.; Lauterbach, L.; Dickschat, J. S. *Angew. Chem. Int. Ed.* **2017**, *56*, 16385.
- (31) Lauterbach, L.; Rinkel, J.; Dickschat, J. S. *Angew. Chem. Int. Ed.* **2018**, *57*, 8280.
- (32) Srivastava, P. L.; Escorcía, A. M.; Huynh, F.; Miller, D. J.; Allemann, R. K.; van der Kamp, M. W. *ACS Catal.* **2021**, *11*, 1033.
- (33) Cane, D. E.; Shim, J. H.; Xue, Q.; Fitzsimons, B. C.; Hohn, T. M. *Biochemistry* **1995**, *34*, 2480.

- (34) Xu, H.; Rinkel, J.; Chen, X.; Köllner, T. G.; Chen, F.; Dickschat, J. S. *Org. Biomol. Chem.* **2021**, *19*, 370.
- (35) Rinkel, J.; Köllner, T. G.; Chen, F.; Dickschat, J. S. *Chem. Commun.* **2019**, *55*, 13255.
- (36) Hou, A.; Goldfuss, B.; Dickschat, J. S. *Angew. Chem. Int. Ed.* **2021**, *60*, 20781.
- (37) Matos, J. O.; Kumar, R. P.; Ma, A. C.; Patterson, M.; Krauss, I. J.; Oprian, D. D. *Biochemistry* **2020**, *59*, 3271.
- (38) Li, R.; Chou, W. K. W.; Himmelberger, J. A.; Litwin, K. M.; Harris, G. G.; Cane, D. E.; Christianson, D. W. *Biochemistry* **2014**, *53*, 1155.
- (39) Blank, P. N.; Barrow, G. H.; Chou, W. K. W.; Duan, L.; Cane, D. E.; Christianson, D. W. *Biochemistry* **2017**, *56*, 5798.
- (40) Forcat, S.; Allemann, R. K. *Org. Biomol. Chem.* **2006**, *4*, 2563.
- (41) Kim, S.-Y.; Zhao, P.; Igarashi, M.; Sawa, R.; Tomita, T.; Nishiyama, M.; Kuzuyama, T. *Chem. Biol.* **2009**, *16*, 736.
- (42) Meguro, A.; Motoyoshi, Y.; Teramoto, K.; Ueda, S.; Totsuka, Y.; Ando, Y.; Tomita, T.; Kim, S.-Y.; Kimura, T.; Igarashi, M.; Sawa, R.; Shinada, T.; Nishiyama, M.; Kuzuyama, T. *Angew. Chem. Int. Ed.* **2015**, *54*, 4353.
- (43) Görner, C.; Häuslein, I.; Schrepfer, P.; Eisenreich, W.; Brück, T. *ChemCatChem* **2013**, *5*, 3289.
- (44) Tomita, T.; Kim, S.-Y.; Teramoto, K.; Meguro, A.; Ozaki, T.; Yoshida, A.; Motoyoshi, Y.; Mori, N.; Ishigami, K.; Watanabe, H.; Nishiyama, M.; Kuzuyama, T. *ACS Chem. Biol.* **2017**, *12*, 1621.
- (45) Köksal, M.; Chou, W. K. W.; Cane, D. E.; Christianson, D. W. *Biochemistry* **2012**, *51*, 3011.
- (46) Deligeorgopoulou, A.; Taylor, S. E.; Forcat, S.; Allemann, R. K. *Chem. Commun.* **2003**, 2162.
- (47) Faraldos, J. A.; Antonczak, A. K.; González, V.; Fullerton, R.; Tippmann, E. M.; Allemann, R. K. *J. Am. Chem. Soc.* **2011**, *133*, 13906.
- (48) Burkhardt, I.; Kreuzenbeck, N.; Beemelmans, C.; Dickschat, J. S. *Org. Biomol. Chem.* **2019**, *17*, 3348.
- (49) Janke, R.; Görner, C.; Hirte, M.; Brück, T.; Loll, B. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2014**, *70*, 1528.
- (50) Lesburg, C. A.; Zhai, G.; Cane, D. E.; Christianson, D. W. *Science* **1997**, *277*, 1820.
- (51) Gutta, P.; Tantillo, D. J. *J. Am. Chem. Soc.* **2006**, *128*, 6172.
- (52) Zu, L.; Xu, M.; Lodewyk, M. W.; Cane, D. E.; Peters, R. J.; Tantillo, D. J. *J. Am. Chem. Soc.* **2012**, *134*, 11369.
- (53) Lauterbach, L.; Goldfuss, B.; Dickschat, J. S. *Angew. Chem. Int. Ed.* **2020**, *59*, 11943.
- (54) Caruthers, J. M.; Kang, I.; Rynkiewicz, M. J.; Cane, D. E.; Christianson, D. W. *J. Biol. Chem.* **2000**, *275*, 25533.
- (55) Forcat, S.; Allemann, R. K. *Chem. Commun.* **2004**, 2094.
- (56) Faraldos, J. A.; González, V.; Senske, M.; Allemann, R. K. *Org. Biomol. Chem.* **2011**, *9*, 6920.
- (57) Calvert, M. J.; Ashton, P. R.; Allemann, R. K. *J. Am. Chem. Soc.* **2002**, *124*, 11636.
- (58) Calvert, M. J.; Taylor, S. E.; Allemann, R. K. *Chem. Commun.* **2002**, 2384.
- (59) Deligeorgopoulou, A.; Allemann, R. K. *Biochemistry* **2003**, *42*, 7741.
- (60) Chen, M.; Chou, W. K. W.; Al-Lami, N.; Faraldos, J. A.; Allemann, R. K.; Cane, D. E.; Christianson, D. W. *Biochemistry* **2016**, *55*, 2864.
- (61) Lauterbach, L.; Hou, A.; Dickschat, J. S. *Chem. Eur. J.* **2021**, *27*, 7923.

## Publications

- (1) H. Xu, G. Schnakenburg, B. Goldfuss, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2023**, 62, e202306429; *Angew. Chem.* **2023**, 135, e202306429.
- (2) H. Xu, B. Goldfuss, J. S. Dickschat, *Org. Lett.* **2023**, 25, 3330-3334.
- (3) H. Xu, L. Lauterbach, B. Goldfuss, G. Schnakenburg, J. S. Dickschat, *Nat. Chem.* **2023**, 15, accepted (doi: 10.1038/s41557-023-01223-z).
- (4) T. Lou, A. Li, H. Xu, J. Pan, B. Xing, R. Wu, J. S. Dickschat, D. Yang, M. Ma, *J. Am. Chem. Soc.* **2023**, 145, 8474-8485.
- (5) H. Xu, J. S. Dickschat, *ChemBioChem* **2023**, 24, e202300101.
- (6) H. Xu, J. S. Dickschat, *Beilstein J. Org. Chem.* **2023**, 19, 186-203.
- (7) D. Baer, B. Konetschny, A. Kulik, H. Xu, D. Paccagnella, P. Beller, N. Ziemert, J. S. Dickschat, B. Gust, *Microb. Cell Fact.* **2022**, 21, 232.
- (8) H. Xu, A. Wochele, M. Luo, G. Schnakenburg, Y. Sun, H. Brötz-Oesterhelt, J. S. Dickschat, *Beilstein J. Org. Chem.* **2022**, 18, 1159-1165.
- (9) B. Xing, H. Xu, A. Li, T. Lou, M. Xu, K. Wang, Z. Xu, J. S. Dickschat, D. Yang, M. Ma, *Angew. Chem. Int. Ed.* **2022**, 61, e202209785; *Angew. Chem.* **2022**, 134, e202209785.
- (10) H. Xu, J. S. Dickschat, *Chem. Eur. J.* **2022**, 28, e202200405.
- (11) Y.-H. Wang, H. Xu, J. Zou, X.-B. Chen, Y.-Q. Zhuang, W.-L. Liu, E. Celik, G.-D. Chen, D. Hu, H. Gao, R. Wu, P.-H. Sun, J. S. Dickschat, *Nat. Catal.* **2022**, 5, 128-135.
- (12) H. Xu, N. Lackus, T. G. Köllner, J. S. Dickschat, *Org. Lett.* **2022**, 24, 587-591.
- (13) H. Xu, B. Goldfuss, G. Schnakenburg, J. S. Dickschat, *Beilstein J. Org. Chem.* **2022**, 18, 13-24.
- (14) H. Xu, J. S. Dickschat, *Synthesis* **2022**, 54, 1551-1565.
- (15) H. Xu, C. Schotte, R. J. Cox, J. S. Dickschat, *Org. Biomol. Chem.* **2021**, 19, 8482-8486.
- (16) M. Luo, H. Xu, Y. Dong, K. Shen, J. Lu, Z. Yin, M. Qi, G. Sun, L. Tang, J. Xiang, Z. Deng, J. S. Dickschat, Y. Sun, *Angew. Chem. Int. Ed.* **2021**, 60, 19139-19143; *Angew. Chem.* **2021**, 133, 19288-19292.
- (17) H. Xu, B. Goldfuss, J. S. Dickschat, *Chem. Eur. J.* **2021**, 27, 9758-9762.

- (18) N. D. Lackus, J. Morawetz, H. Xu, J. Gershenzon, J. S. Dickschat, T. G. Köllner, *Molecules* **2021**, *26*, 555.
- (19) H. Xu, J. Rinkel, J. S. Dickschat, *Org. Chem. Front.* **2021**, *8*, 1177-1184.
- (20) H. Xu, J. Rinkel, X. Chen, T. G. Köllner, F. Chen, J. S. Dickschat, *Org. Biomol. Chem.* **2021**, *19*, 370-374.
- (21) H. Xu, J. S. Dickschat, *ChemBioChem* **2021**, *22*, 850-854.
- (22) H. Xu, J. S. Dickschat, *Chem. Eur. J.* **2020**, *26*, 17318-17341.

## Abstract

This doctoral dissertation consists of 4 reviews and 16 research articles that present mechanistic studies on terpene biosynthesis. Specifically, sesquiterpenes biosynthetically derived from germacrene A, hedycaryol and germacrene B are summarised in three reviews. All compounds are described together with their structures, natural sources, determination of their absolute configurations, biosynthetic pathways and NMR data. For the original research, the studied terpenes included the new compounds isoishwarane, (1*S*,5*S*,7*R*,10*R*)-guaia-4(15)-en-11-ol, kitaviridene and the sesterviridenes, besides known hedycaryol, (4*S*,7*R*)-germacra-(1(10)*E*,5*E*)-dien-11-ol, (1*S*,7*R*,10*R*)-guaia-4-en-11-ol, selina-4(15),7(11)-diene, patchoulol,  $\alpha$ -humulene, spiroviolene and the non-canonical compounds sodorifen and geosmin. For all these compounds the biosynthesis was extensively studied through conversion of isotopically labelled precursors with the respective terpene synthases, in conjunction with DFT calculations or QM/MM MD simulations, culminating in the proposal of rational biosynthetic pathways for each investigated compound.

The fourth review article summarises site-directed mutagenesis studies on terpene synthases, contributing to a deeper understanding of these enzymes with respect to the functions of residues and motifs in the active site cavity. Two experimental studies regarding the structure-based site-directed mutagenesis were conducted on the diterpene synthase for cattleyene (CyS) and three sesquiterpene synthases for presilphiperfolan-8 $\beta$ -ol (BcBOT), protoillud-6-ene (DbPROS) and longiborneol (CLM1), respectively, demonstrating the potential of enzyme engineering to expand the chemical space that can be reached with terpene synthase catalysis.

Taken together, the results presented in this thesis significantly expand our knowledge on the complex mechanisms of terpene synthase mediated cyclisation reactions and the structural requirements for efficient catalysis by this remarkable class of biocatalysts.