

# **Multifunctional Enzymes in Natural Product Biosynthesis**

Kumulative Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

**Zhiyong Yin**

aus Anqing

Bonn 2025

Die vorliegende Arbeit wurde in der Zeit vom 01.02.2020 bis zum 31.01.2025 am Kekulé-Institut für Organische Chemie und Biochemie der Rheinischen Friedrich-Wilhelms-Universität Bonn unter der Leitung von Prof. Dr. Jeroen S. Dickschat angefertigt.

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

Gutachter/Betreuer: Prof. Dr. Jeroen S. Dickschat

Gutachter: Prof. Dr. Arne Lützen

Tag der Promotion: 30.05.2025

Erscheinungsjahr: 2025

## Acknowledgement

First of all, I would like to sincerely thank my supervisor, Prof. Dr. Jeroen S. Dickschat, for his guidance, support, and invaluable insights throughout the course of my thesis work. Your dedication and commitment have been a constant source of inspiration, and I deeply appreciate the time and effort you have invested in helping me grow both academically and personally. Your constructive feedback and thoughtful challenges have pushed me to expand my horizons and reach higher standards. Furthermore, I am especially grateful for your patience and for the motivation you provided, which extended beyond the academic part—thank you for also encouraging me to explore long-distance cycling. This work would not have been possible without your continuous encouragement and belief in me.

I would like to thank Prof. Dr. Arne Lützen for spending time reading and refereeing my thesis, as well as Prof. Dr. Olav Schiemann and PD Dr. Gregor Hagelüken for joining the thesis examination committee.

I would like to extend my deepest gratitude to my labmates, with whom I spent much of my time and shared many unforgettable moments throughout my PhD journey. Special thanks go to Dr. Anwei Hou and Dr. Zhiyang Quan, whose guidance at the start of my PhD was invaluable. I am also profoundly thankful to Dr. Binbin Gu, Dr. Linfu Liang, Dr. Lukas Lauterbach, Dr. Houchao Xu, Dr. Anuj Chhalodia, Heng Li, Kizerbo A. Taizoumbe, Dr. Georges B. Tabekoueng, Alain Tumma, Min Wang, Kexin Yang, Dr. Geng Li, Dr. Fulong Lin, Neran Reuber, and Alina Oseguera Jaramillo for fostering such a supportive and productive lab environment.

In addition, I would like to express my thanks to the NMR department, HPLC pool, and Mass department for their invaluable assistance in measuring numerous samples. I am also deeply grateful to the technical staff for their support and help throughout my research.

Finally, I would like to express my heartfelt gratitude to my parents, my sisters, and my brothers for their unconditional love and support, shaping me into who I am today. My deepest thanks go to Shiyang Zhai, Jingya Wang, Luoxia Shi and Qiong Gu for their help and continuous encouragement throughout this journey. I am also sincerely grateful to my friends for their companionship and support. This thesis would not have been possible without every one of you.

## **Preamble**

This cumulative doctoral thesis, titled “Multifunctional Enzymes in Natural Product Biosynthesis,” consists of seven chapters focusing on multifunctional enzymes in polyketide and terpene biosynthesis. Chapter 1 provides a detailed introduction to the polyketide synthases in polyketide biosynthesis, along with the enzymes involved in terpene biosynthesis. Chapters 2–6 each contain a brief introduction and summary of the associated publications. Chapter 7 offers a comprehensive summary of the work presented in this thesis and discusses how the reported scientific advances could influence future research. The corresponding publications are included in Appendices A–E.

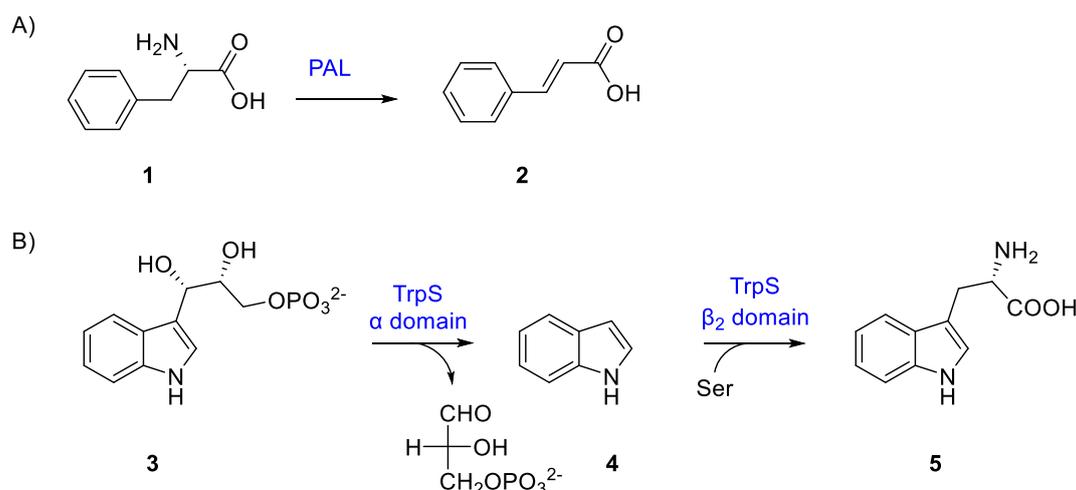
## Table of contents

|  |    |
|--|----|
| 1. Introduction .....  | 1  |
| 1.1 Polyketide biosynthesis.....   | 4  |
| 1.1.1 Phosphopantetheinyl transferases and carrier proteins .....  | 12 |
| 1.1.2 Acyltransferases .....   | 14 |
| 1.1.3 Ketosynthases .....  | 16 |
| 1.1.4 Optional modification domains .....  | 19 |
| 1.2 Non-ribosomal peptide biosynthesis .....   | 23 |
| 1.3 Multifunctional enzymes in terpene biosynthesis .....  | 24 |
| 1.4 Conclusion .....   | 30 |
| 2. A Clickable Coenzyme A Derived Probe for Investigating<br>Phosphopantetheinyl Transferase Activity in Natural Product Biosynthesis..... | 33 |
| 3. Substrate Specificity of a Ketosynthase Domain Involved in Bacillaene<br>Biosynthesis .....   | 39 |
| 4. An Isotopic Probe to Follow the Stereochemical Course of Dehydratase<br>Reactions in Polyketide and Fatty Acid Biosynthesis.....        | 45 |
| 5. The Substrate Scope of Dehydratases in Antibiotic Biosynthesis and their<br>Application in Kinetic Resolutions .....                    | 51 |
| 6. Enantioselective Synthesis of All Stereoisomers of Geosmin and of<br>Biosynthetically Related Natural Products.....                     | 57 |
| 7. Summary and outlook.....  | 63 |
| 8. References.....   | 69 |
| 9. Appendices A–E.....   | 80 |

# Chapter 1

## 1 Introduction

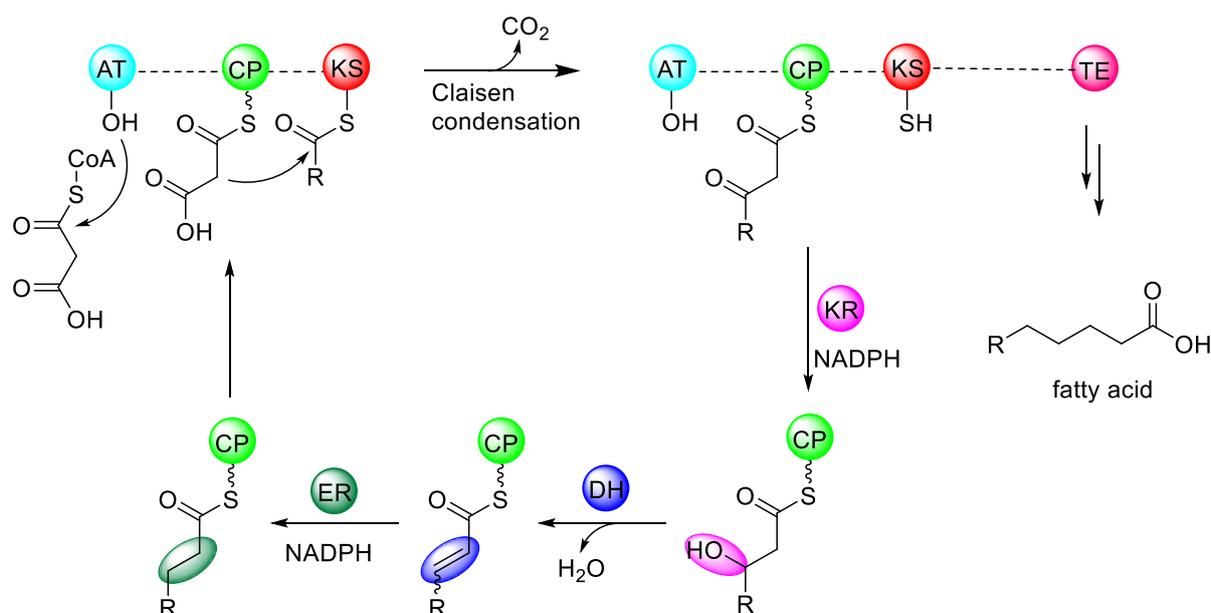
Enzymes are biocatalysts that play an important role in biochemical reactions. They can be divided into two categories based on their catalytic properties, namely monofunctional enzymes and multifunctional enzymes <sup>[1,2]</sup>. Monofunctional enzymes usually catalyze one step reactions. For example, phenylalanine ammonia-lyase (PAL) catalyzes the deamination of phenylalanine (**1**), producing *trans*-cinnamic acid (**2**) in the phenylpropanoid biosynthetic pathway (Scheme 1A) <sup>[3,4]</sup>. In contrast, multifunctional enzymes have the ability to catalyze multiple step reactions, either within a single active site or through different domains in a single polypeptide chain <sup>[2]</sup>. Multifunctional enzymes are significant in the biosynthesis of natural products, such as fatty acid synthases (FASs), polyketide synthases (PKSs), which will be discussed in this thesis.



**Scheme 1.** A) The formation of *trans*-cinnamic acid (**2**) is catalyzed by monofunctional enzyme phenylalanine ammonia-lyase (PAL) from L-phenylalanine (**1**). B) The multifunctional enzyme tryptophan synthase catalyzes the final two reactions in the L-tryptophan biosynthetic pathway.

Multifunctional enzymes are found in both primary and secondary metabolic pathways, such as tryptophan synthase (TrpS) and fatty acid synthase in primary metabolism. Tryptophan synthase is a multifunctional enzyme that catalyzes the final two reactions in the L-tryptophan biosynthetic pathway (Scheme 1B). It comprises a two-domain complex organized into an  $\alpha_2\beta_2$  subunit structure, where the  $\alpha$  and  $\beta$

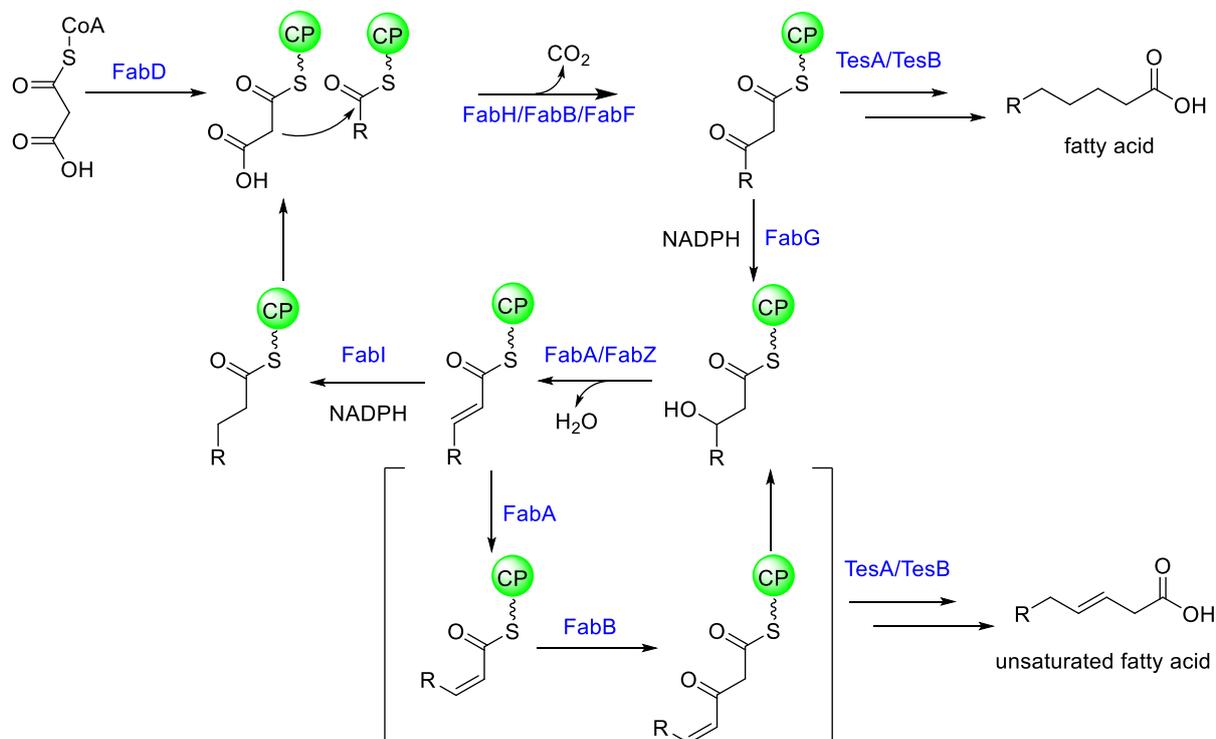
subunits perform distinct but interconnected catalytic functions. The  $\alpha$  subunit catalyzes the cleavage of indole-3-glycerol phosphate (**3**), referred to as the  $\alpha$  reaction. The  $\beta$  subunit, typically forming a dimer with two pyridoxal phosphate cofactors, catalyzes the synthesis of L-tryptophan (**5**) from indole (**4**) and L-serine, known as the  $\beta$  reaction. Together, the  $\alpha_2\beta_2$  complex catalyzes the essential reactions, with enhanced catalytic efficiency [5].



**Scheme 2.** Type I fatty acid biosynthesis. Fatty acid was synthesized by multifunctional enzyme fatty acid synthase through condensation, reduction, dehydration, and a subsequent reduction.

Another example of a multifunctional enzyme in natural product primary metabolic pathways is type I fatty acid synthase (FAS), which is found in fungi and animals and catalyzes a series of reactions to synthesize long-chain fatty acids from acetyl-CoA and malonyl-CoA (Scheme 2). The crystal structures of FAS has also been determined and this gives a deeper understanding of fatty acid biosynthesis [6-9]. In the first step of fatty acid biosynthesis, an acyltransferase transfers the acetyl group of acetyl-CoA to the acyl carrier protein (ACP). Then the starting unit is transferred to the ketosynthase (KS). Simultaneously, acyltransferase (AT) loads malonyl-CoA onto the ACP. The KS then catalyzes Claisen condensation, forming a  $\beta$ -ketoacyl-ACP intermediate. Subsequent modification of the  $\beta$ -ketoacyl-ACP intermediate involves a sequence of reactions: ketoreduction which is catalyzed by a NADPH-dependent ketoreductase (KR), dehydration with a dehydratase (DH), and enoyl reduction by a

NADPH-dependent enoyl reductase (ER), yielding a saturated acyl chain elongated by two carbons. These reactions repeat until the fatty acid reaches 16–18 carbons, after which it is released from ACP, typically by a thioesterase (TE) [10]. This well-organized enzymatic machinery efficiently produces fatty acids as essential cellular components.



**Scheme 3.** Type II fatty acid biosynthesis. Fatty acid was synthesized by different monofunctional enzymes from fatty acid biosynthetic pathway

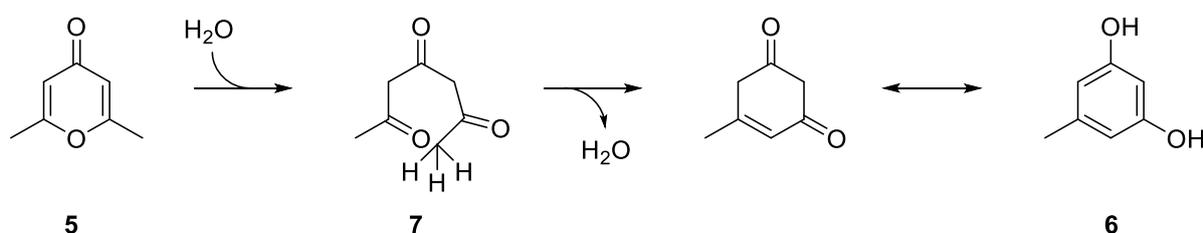
Unlike the multifunctional complexes found in type I fatty acid systems, type II fatty acid biosynthesis is catalyzed by a series of distinct monofunctional enzymes and widely present in bacteria, mitochondria, and plants. Type II fatty acid synthases are capable of producing a variety of fatty acid products, including saturated, unsaturated, branched-chain, and hydroxylated fatty acids (Scheme 3). The *E. coli* type II FAS cycle is divided into two modules: initiation and elongation. In the initiation module, the acetyl-CoA is transferred to ACP by ACP transacylase (FabD) [11].  $\beta$ -Ketoacyl-ACP synthase III (FabH) then catalyzes the first condensation between acetyl-CoA and malonyl-ACP [12]. In the elongation module, the acyl chain gets elongated through repeated cycles of condensation (FabB/FabF) [13], reduction (FabG) [14], dehydration (FabA/FabZ) [15], and enoyl reduction (FabI) [16]. A unique feature of type II FAS is the ability to produce unsaturated fatty acids, facilitated by FabA, which not only dehydrates  $\beta$ -hydroxyacyl-ACP but also catalyzes isomerization to form

intermediates suitable for further elongation. TesA (thioesterase I) <sup>[17]</sup> and TesB (thioesterase II) <sup>[18]</sup> in *E. coli* are responsible for hydrolyzing both acyl-CoA and acyl-ACP substrates. The ACP-bound fatty acids generated through this pathway serve as precursors for essential biomolecules, including acyl phosphates, phospholipids, lipoic acid, and quorum-sensing molecules, highlighting the central role of type II FAS in metabolic processes <sup>[10]</sup>.

Multifunctional enzymes are essential not only in primary metabolic pathways but also in secondary metabolic pathways of natural products, such as polyketides, non-ribosomal peptides, and terpenoids.

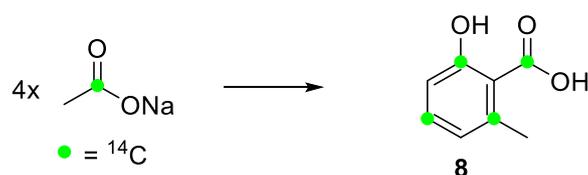
### 1.1 Polyketide biosynthesis

The exploration of polyketides began in the late 19th century, marked by a significant breakthrough from Collie at the University College London. In 1893, while investigating the structure of dehydroacetic acid (**5**), Collie produced orcinol (**6**), by boiling of **5** with barium hydroxide and then treating it with acid. To confirm the structure of **6**, Collie proposed a possible intermediate **7**, which can be formed from dehydroacetic acid (**5**) by the addition of water (Scheme 4). This early discovery led to deeper studies of polyketide structures, biosynthesis, and various applications in pharmaceuticals and industry <sup>[19,20]</sup>.



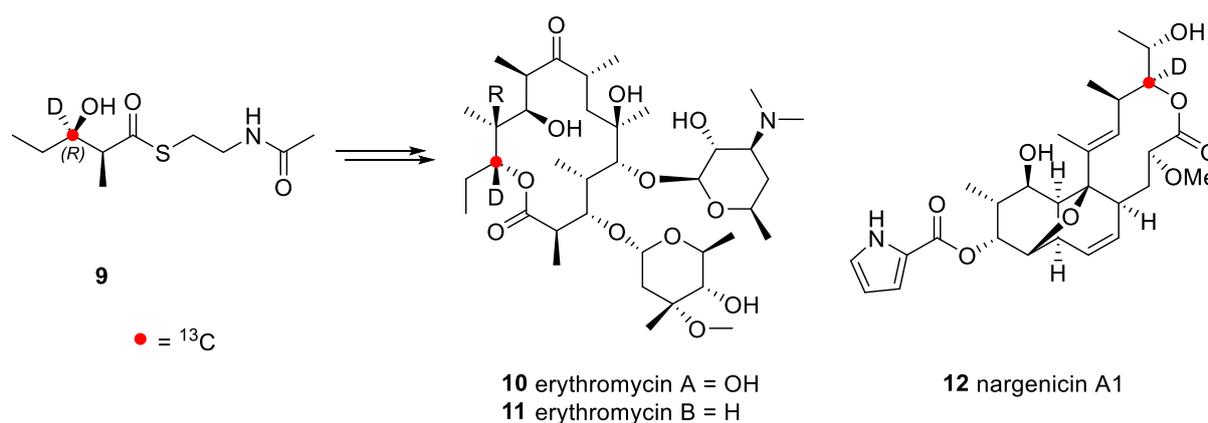
**Scheme 4.** Collie's proposed synthesis of orcinol (**6**) from dehydroacetic acid (**5**).

In the 1950s, Birch proposed that polyketides are generated from acetic acid units through repeated condensation reactions via head-to-tail linkage. This hypothesis was confirmed by feeding isotopically labeled sodium [1-<sup>14</sup>C]acetate in studies investigating the biosynthesis of polyketide 6-methylsalicylic acid (**8**) (Scheme 5) <sup>[21]</sup>.



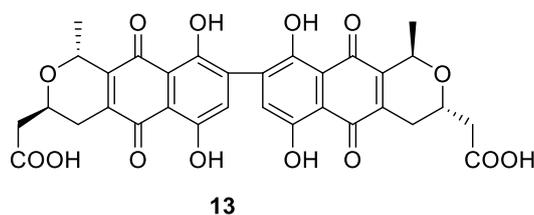
**Scheme 5.** 6-Methylsalicylic acid (**8**) is formed from four acetate units.

With the development of modern analytical chemistry, nuclear magnetic resonance (NMR) and mass spectroscopy (MS) are widely used in polyketide biosynthesis research. For example, Ott used (2*S*,3*R*)-(3-<sup>2</sup>H,3-<sup>13</sup>C)-2-methyl-3-hydroxypentanoyl N-acetylcysteamine (NAC) thioester (**9**) to investigate the biosynthesis of erythromycin A (**10**) and B (**11**) and nargenicin A1 (**12**). The <sup>2</sup>H and <sup>13</sup>C isotopically labeled compounds were retained in the metabolites following the incorporation of the isotopically labeled NAC thioester precursor into erythromycin A (**10**), erythromycin B (**11**) and nargenicin A1 (**12**). The results showed that erythromycin and nargenicin were synthesized directly from the common precursor (2*S*,3*R*)-(3-<sup>2</sup>H,3-<sup>13</sup>C)-2-methyl-3-hydroxypentanoyl NAC-thioester (**9**), without oxidation to the 3-ketoester. This indicated a stepwise polyketide biosynthesis, where each building block is immediately reduced, rather than constructing a polyketone chain before the reduction (Scheme 6) [22].



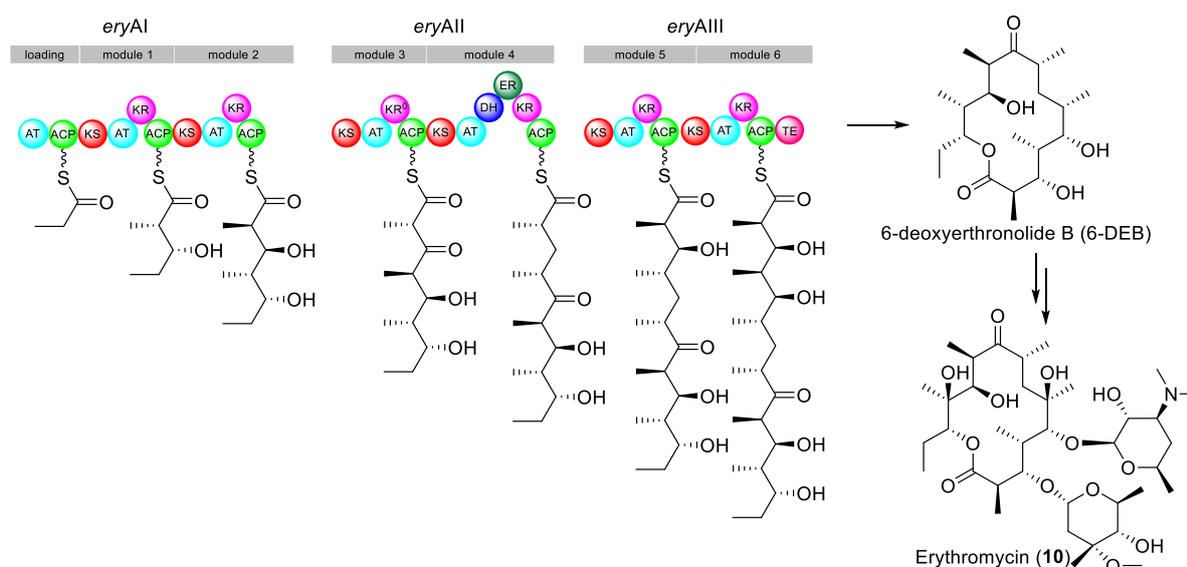
**Scheme 6.** Incorporation of NAC thioester intermediate **9** into erythromycin A (**10**), erythromycin B (**11**) and nargenicin A<sub>1</sub> (**12**).

Advances in genetic techniques during the 1980s, especially the work of Hopwood, led to the discovery of polyketide genes responsible for polyketide biosynthesis. Using actinorhodin (**13**) as a model, Hopwood identified the genes encoding the enzymes for its production (Scheme 7) [23].



### Scheme 7. Structure of actinorhodin (13).

The genes involved in erythromycin biosynthesis were identified independently by Leadlay and Katz in 1990 [24,25]. These findings revealed three large genes (*eryAI*, *eryAII* and *eryAIII*) that encode the deoxyerythronolide-B-synthases (DEBSs) 1, 2 and 3 (Scheme 8). These enzymes are responsible for producing the intermediate 6-deoxyerythronolide B (6-dEB) of erythromycin (10). By comparing their domains to those in fatty acid synthases, Leadlay and Katz confirmed their role in building erythromycin [24-26]. The DEBSs are multifunctional and catalyze the steps needed to produce erythromycin (10) [27]. Building on the foundational work of these pioneers in polyketide biosynthesis and the advances in new technologies, more and more polyketide synthases have been discovered and extensively studied, leading to a deeper understanding of their roles and mechanisms in polyketide biosynthesis.

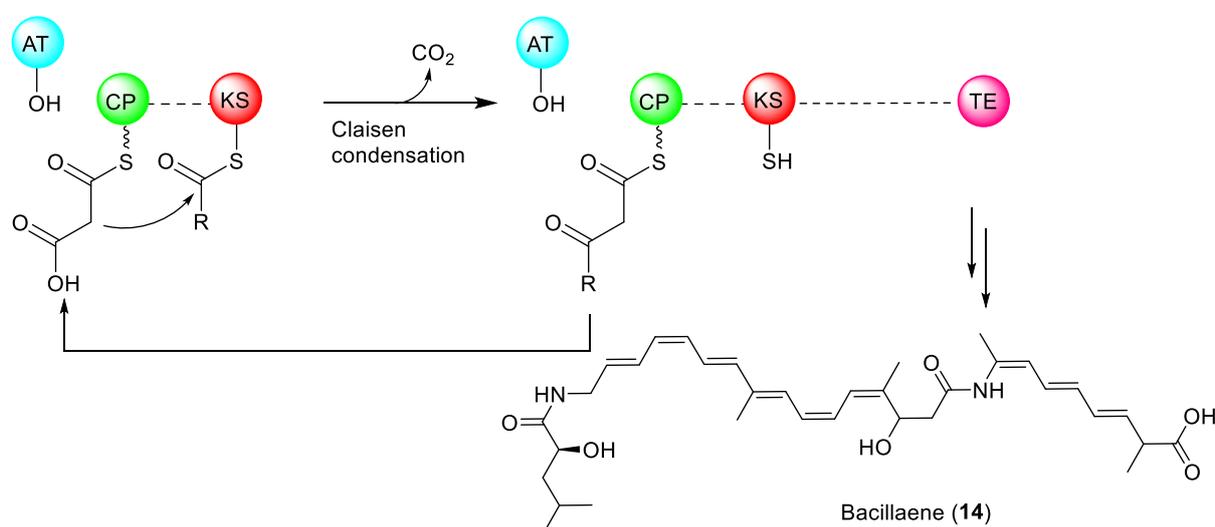


### Scheme 8. The DEBSs synthesize 6-deoxyerythronolide B, the precursor of erythromycin A.

Polyketide synthases are classified into three types: type I PKSs, type II PKSs and type III PKSs (Table 1) [28]. Type I PKSs contain multiple enzymatic domains and are further subdivided into non-iterative PKSs and iterative PKSs. Non-iterative type I PKSs, also known as modular PKSs, are prevalent in bacteria. These modular PKSs work in a sequential manner, with each module catalyzing one elongation cycle before transferring the growing polyketide chain to the next module [28].

**Table 1.** Classification of PKSs

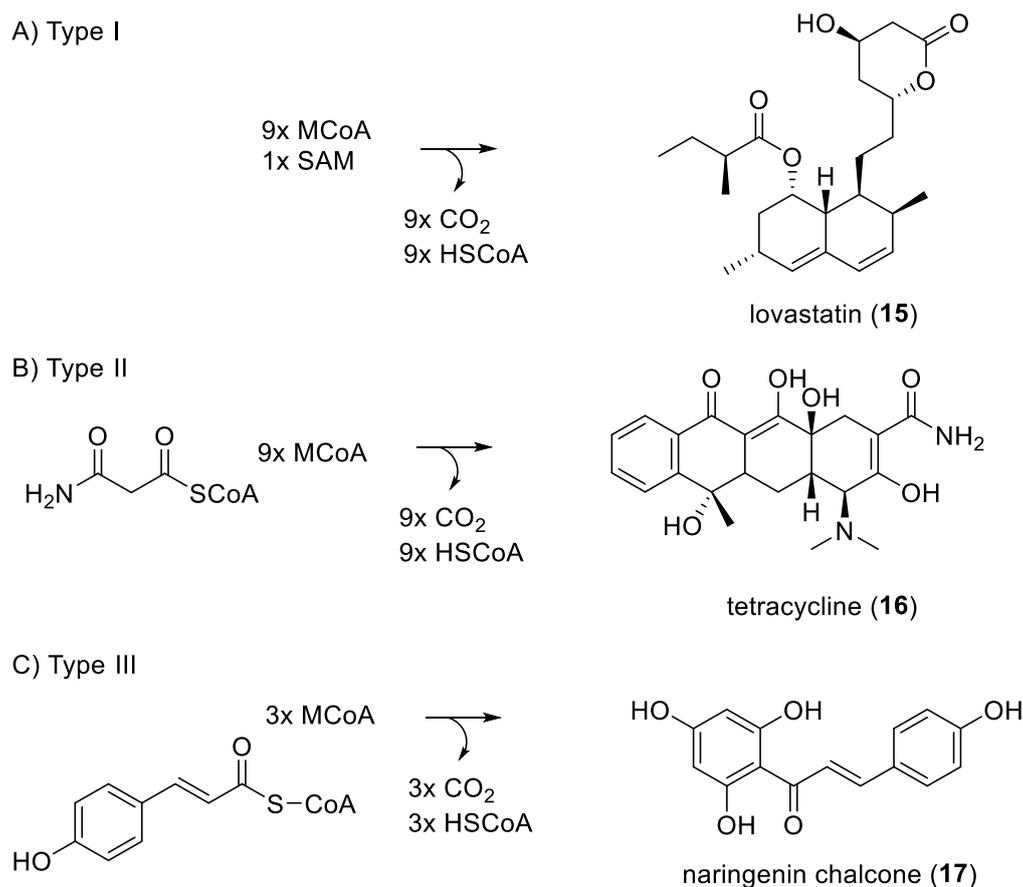
|            | Non-iterative PKSs  |                                | Iterative PKSs |   |             |
|------------|---------------------|--------------------------------|----------------|---|-------------|
|            | Type I PKSs         |                                | Type II PKSs   | Type III PKSs                             |             |
|            | <i>cis</i> -AT PKSs | <i>trans</i> -AT PKSs          |                |   |             |
| Extenders  | various             | malonyl-CoA                    | malonyl-CoA    | malonyl-CoA                               | malonyl-CoA |
| Substrates | ACP-bound           | ACP-bound                      | ACP-bound      | ACP-bound                                 | CoA-bound   |
| Organisms  | bacteria            | mainly fungi,<br>also bacteria | only bacteria  | mainly plants,<br>also bacteria,<br>fungi |             |

**Scheme 9.** The *trans*-AT polyketide biosynthesis cycle.

The non-iterative type I PKSs can be classified into *cis*-AT PKSs and *trans*-AT PKSs [29,30]. In *cis*-AT PKSs, the KS, AT and ACP domains are covalently linked within the same polypeptide chain. The DEBSs, which are required for erythromycin biosynthesis, can be regarded as an example of a non-iterative *cis*-AT type I PKSs (Scheme 8) [24-27]. In contrast, Bacillaene (**14**) is an antibiotic produced by *Bacillus* spp. strains, synthesized via a *trans*-AT PKSs system (Scheme 9). The biosynthetic gene cluster of bacillaene (**14**) from *B. subtilis* 168 was the first reported *trans*-AT PKSs system [31-34]. *Trans*-AT PKSs have a standalone AT domain that transfers the extender unit onto the ACP domain, which is often shared among multiple modules (Scheme 9). Typically, the extender building block is malonyl-CoA, and incorporation of methyl groups into the carbon skeleton is facilitated by SAM (S-

adenosylmethionine)-dependent methyltransferases [29,30]. These characteristics highlight the unique functions of enzymes and substrate specificity of *trans*-AT PKSs in polyketide biosynthesis.

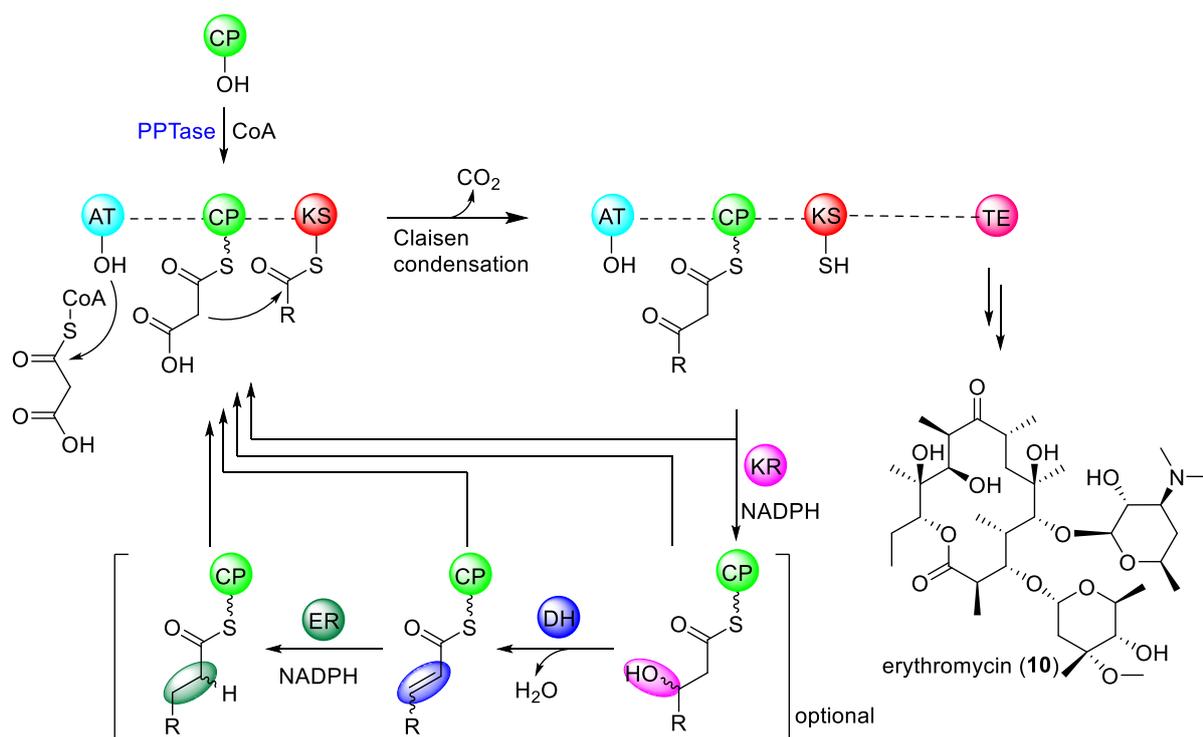
Iterative type I PKSs, found mainly in fungi and some bacteria, catalyze multiple chain elongation cycles using the same set of domains [28]. An example of iterative type I PKSs is lovastatin (**15**) synthase (Scheme 10A) [35].



**Scheme 10.** Examples of iterative PKSs. A) lovastatin (**15**), B) tetracycline (**16**), C) naringenin chalcone (**17**).

Type I polyketide synthases (PKSs) are classic examples of multifunctional enzymes in natural product biosynthesis [36]. They are responsible for synthesizing a variety of important natural products, including antibiotics like erythromycin (a polyketide, **10**) [37]. Within the assembly line of polyketide (Scheme 11), the malonyl/acyl transferase (MAT/AT) domain can select and transfer an acyl group from coenzyme A (CoA) to the  $\alpha$ -hydroxyacyl residue within the polyketide synthase. The acyl carrier protein (ACP) domain, carrying the elongation group, undergoes a crucial chemical transformation known as Claisen condensation. This reaction happens with the

ketosynthase (KS) domain. In the Claisen condensation process, the ACP-bound elongation group reacts with the KS-bound polyketide chain. This reaction results in the formation of a  $\beta$ -keto thioester intermediate, accompanied by the elimination of a unit of carbon dioxide ( $\text{CO}_2$ ). As a result of this reaction, the polyketide chain moves outward by one position, extending its length by two carbons. At the same time, the elongation group from the ACP becomes the newly bound group at the terminal end of the polyketide chain [28,29,38-44].



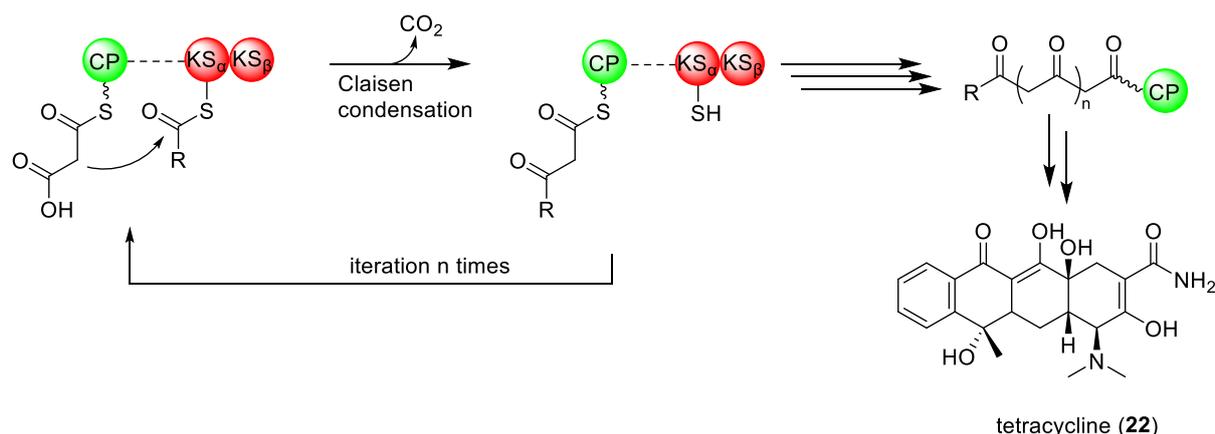
**Scheme 11.** Polyketide biosynthesis cycle.

An important point of this process is the ACP domain, which initially exists in an inactive *apo*-ACP form. It is activated by a phosphopantetheinyl transferase (PPTase), which converts the *apo*-ACP into the active *holo*-ACP form by attaching a phosphopantetheinyl group. The activated ACP (*holo*-ACP) can bind and transport the growing polyketide chain to the domains which are responsible for subsequent enzymatic steps [45,46].

Following the elongation stage, the polyketide chain may undergo additional modifications facilitated by optional domains within the polyketide synthase. These modifications diversify the chemical structure of the polyketide product, introducing specific functional groups and stereochemical features. One of such modifications is carried out by the ketoreductase (KR) domain, which targets the  $\beta$ -keto group within

the polyketide chain. The KR domain, with the cofactor NADPH, catalyzes the reduction of the  $\beta$ -keto group, introducing a hydroxyl group into the polyketide chain. This reduction step is crucial for the generation of structural complexity and can influence the stereochemistry (*R* or *S*) of the hydroxyl groups in polyketides [41]. Subsequently, the polyketide chain may encounter the dehydratase (DH) domain. The DH domain induces the elimination of water from the hydroxyl group within the intermediate, leading to the formation of an  $\alpha,\beta$ -unsaturated alkene intermediate. This transformation can influence the double bond configuration of the polyketide [42]. Moreover, the enoylreductase (ER) domain facilitates subsequent reduction of the polyketide chain. Through a reduction, the ER domain converts the  $\alpha,\beta$ -double bond into a single bond using NADPH as a cofactor [43]. This cycle of chain elongation and modification repeats across modules, with each cycle adding specific structural features to the polyketide. The sequential and modular properties of polyketide biosynthesis enables precise control and diversification of polyketide structures [28,29,38-44].

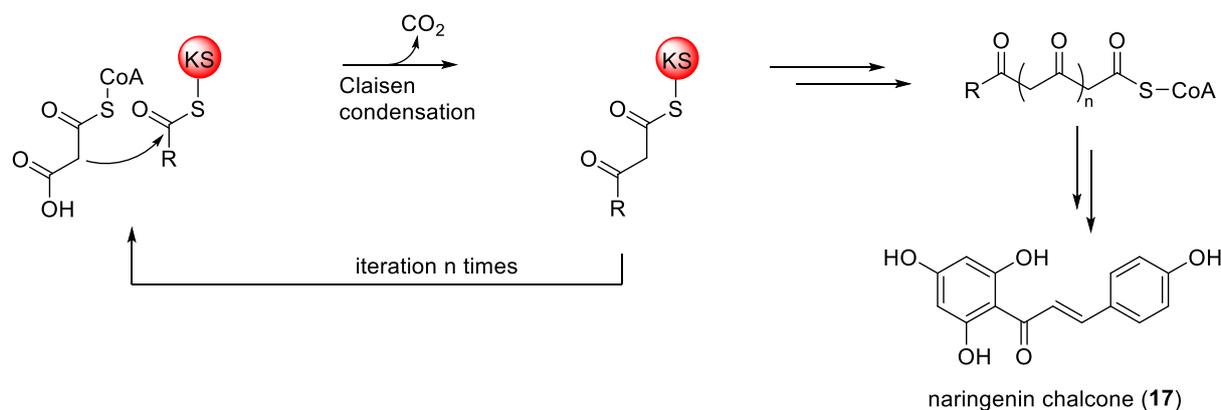
In the final stage, the completed polyketide is released from the polyketide synthase by a thioesterase (TE) domain. The TE domain is important in this process by catalyzing the hydrolysis or macrocyclization of the polyketide, leading to the release of the mature linear or cyclic product [47]. This well-organized assembly process produces a wide variety of polyketide structures with important biological properties.



**Scheme 12.** Type II polyketide biosynthesis cycle.

Type II PKSs are only found in bacteria and work in an iterative mechanism. Both structurally and functionally, they are similar to type II fatty acid synthases which are found in bacteria and plants, consisting of discrete, individually functioning enzymes. The biosynthetic pathway of various type II polyketide gene clusters, as revealed by

their genomic sequences, is defined by a core set of essential enzymes that are iteratively used and encoded by separate genes. Two ketosynthase units ( $\text{KS}_\alpha$  and  $\text{KS}_\beta$ ) and one ACP, which acts as an anchor for the elongating polyketide chain, are usually present in this type of PKSs. The  $\text{KS}_\alpha$  and  $\text{KS}_\beta$  subunits form a heterodimer that facilitates iterative condensation reactions, using malonyl-CoA as an extender unit to elongate the polyketide chain. The  $\text{KS}_\alpha$  subunit is responsible for the catalytic activity, while the  $\text{KS}_\beta$  subunit, also known as the chain length factor (CLF), determines the final length of the polyketide chain (Scheme 12). These enzymes often show a conventional  $\text{KS}_\alpha/\text{KS}_\beta/\text{ACP}$  structure and are clustered together. Type II PKSs consist of different monofunctional enzymes as opposed to type I PKSs, which are made up of a single multifunctional polypeptide. Despite being limited to using only malonyl-CoA as an extender unit, type II polyketide synthases can produce a wide variety of complex polyketide structures. This structural diversity contributes significantly to their ecological and medicinal importance [48]. A representative example of type II PKSs is tetracycline (**16**) synthase (Scheme 10B) [49].



**Scheme 13.** Type III polyketide biosynthesis cycle.

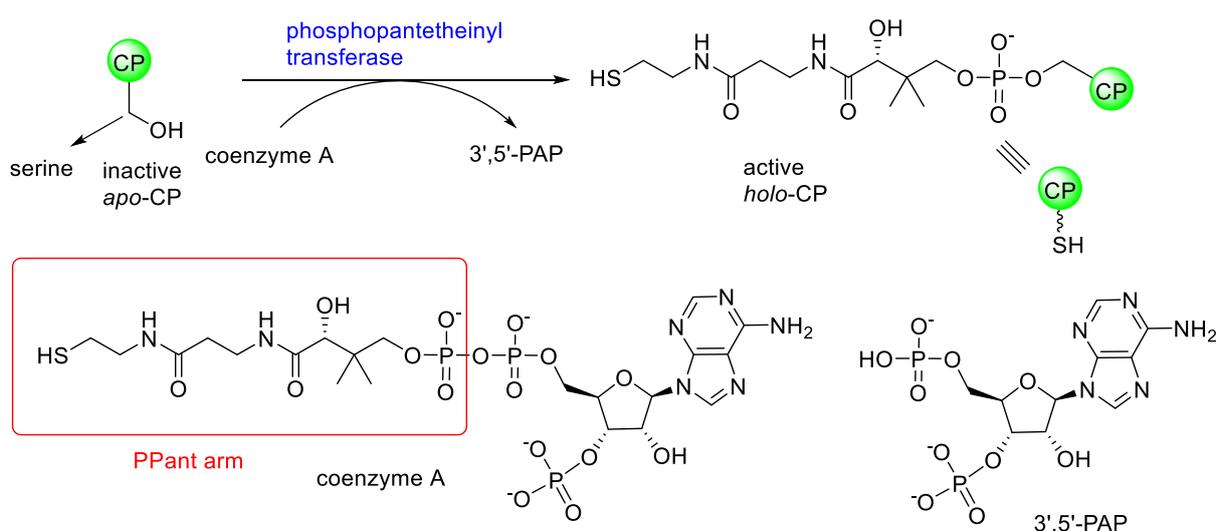
Type III PKSs are monofunctional enzymes, primarily found in plants, also in some fungi and bacteria. Type III PKSs are small homodimeric proteins, approximately 40 kDa in size. They use free malonyl-CoA substrates for chain elongation, unlike type I and type II PKSs, which require an ACP to bind the substrates. The first step in the type III PKSs reaction is the loading of an initial molecule onto the active site cysteine. Subsequently, there are repeated decarboxylative Claisen condensations with malonyl-CoA, culminating in the cyclization of the polyketide intermediate that is attached to the enzyme (Scheme 13). The size and shape of the active site cavity are crucial for the starting molecules, the length of the polyketide chain, and the pathways for the folding and cyclization of the final product [50]. Chalcone synthase

(CHS), a type III PKSs, catalyzes the formation of naringenin chalcone (**17**), a key flavonoid intermediate (Scheme 10C) [51,52].

Despite differences in the structures of different types of PKSs, all PKSs rely on Claisen condensation reactions to assemble polyketides from simple acyl-CoA precursors. The multifunctionality of type I PKSs represents a significant evolutionary advantage, as it integrates multiple catalytic activities into a single enzyme, thereby simplifying the assembly of polyketides. The monofunctional enzymes in type II and type III PKSs provide flexibility for pathway reconfiguration.

### 1.1.1 Phosphopantetheinyl transferases and carrier proteins

Phosphopantetheinyl transferases (PPTases) and carrier proteins are required in multifunctional type I PKSs and FASs, where they facilitate the transfer of acyl groups. However, type II PKSs and FASs, which consist of monofunctional enzymes, also require PPTases and carrier proteins for the transfer of acyl groups [10,48].



**Scheme 14.** Phosphopantetheinyl transferases activate carrier proteins. PPTases transfer the PPant arm from CoA to a conserved serine residue of each CP domain, converting inactive *apo*-CP into active *holo*-CP through phosphoester linkages.

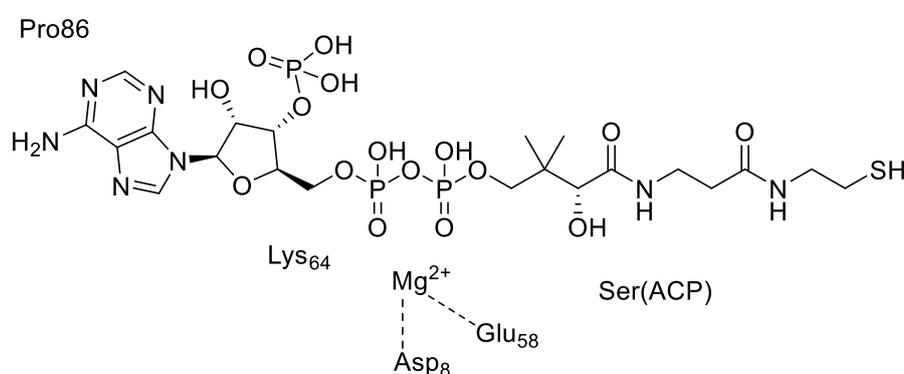
Carrier proteins (CPs) are basic domains in multifunctional FASs, type I PKSs, and NRPSs and represent a class of small proteins that normally have 70 to 100 amino acids. Typically, CPs are responsible for transporting intermediates between different domains. In fatty acid, polyketide and non-ribosomal peptide biosynthetic pathways, the carrier protein utilizes a 4'-phosphopantetheine (PPant) arm, which is approximately 20 Å long, to anchor elongating intermediates through a reactive thioester bond during a series of enzymatic reactions. Before CPs can function, they

must be activated from their *apo* form to the *holo* form, a process mediated by PPTases. PPTases transfer and covalently attach the PPant arm from CoA to a conserved serine residue displayed in a GIDS motif of each CP domain. Through the formation of a phosphoester linkage, the inactive *apo*-CP is converted into the active *holo*-CP (Scheme 14) [45].

CPs can be classified based on the type of intermediate they carry: ACPs, which can carry acyl groups in polyketide biosynthesis, and PCPs, which are responsible for carrying peptidyl groups in non-ribosomal peptide biosynthesis. Unlike the simple classification of CPs, PPTases are generally divided into three classes. This classification is based on their substrate specificities, three-dimensional structures, and conserved amino acid sequences [45,53]. The first type of PPTase is the AcpS type, which consists of about 120 amino acids and forms a homo-trimeric quaternary structure with active sites located at each homotypic interface [54]. AcpS specifically acts on ACPs involved in fatty acid biosynthesis (AcpP) [55]. The Sfp-type PPTases, pseudo-homodimers with roughly 240 amino acids that resemble two AcpS monomers and contain a single active site at the pseudo-dimer interface, make up the second class of PPTases [53,56]. The PPTase Sfp was originally discovered in *Bacillus subtilis* as the enzyme that modifies surfactin synthase. It is known for its remarkable substrate promiscuity, allowing it to activate CPs in both polyketide and non-ribosomal peptide biosynthetic pathways [57]. Additionally, sfp type PPTases can be classified into two types: WxxKEA and FxxKES type, based on the conserved amino acids. The third type of PPTases (type I integrated PPTase) can be found in fungal fatty acid synthases. It is special as it is incorporated as a domain at the C terminus of the fatty acid synthases. In addition, the N-terminus of fatty acid synthases, where the carrier protein is located, is autophosphorylated by this type of PPTases. They post-translationally modify *apo*-ACPs before the assembly of the megasynthases, ensuring that the carrier proteins are properly activated for their role in fatty acid biosynthesis [45,58].

Typically, PPTases specifically activate their natural CP partners [59]. However, there are cases in which a single PPTase can activate various CPs and utilize different CoA analogs [53,57,60]. This indicates that both protein–protein interactions and protein–substrate interactions are crucial in the activation of CPs. To gain a deeper understanding, cocrystal structures of PPTases and CPs were explored. The first co-

crystal structure of a PPTase (AcpS from *B. subtilis*) and its associated ACP has unveiled the protein-protein interactions involved. According to this model, the active site is located at the protein-protein interface. AcpS binds with CoA and  $Mg^{2+}$ , then transfers the PPant arm to the ACP (Scheme 15). During the process of the activation of an ACP by AcpS, two acidic residues, Asp8 and Glu58, chelate and stabilize the  $Mg^{2+}$  ion. In addition, Glu58 facilitates the deprotonation of the conserved serine residue on ACP. Moreover, the opposite of the AcpS monomer stabilizes the adenine base of CoA through hydrophobic interactions with residue Pro86. Lys64 is oriented towards the diphosphate group of CoA and likely donates a proton to aid in the PPant arm transfer stage (Scheme 15) [54].



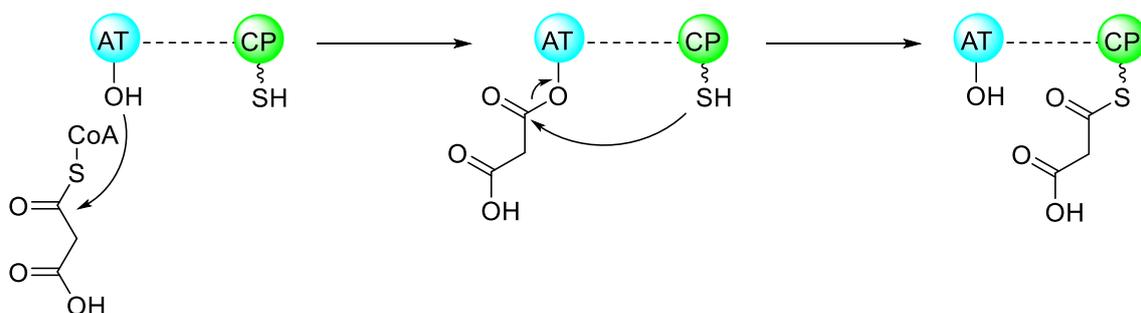
**Scheme 15.** Key residues in the activation of CPs.

In *sfp*-type PPTases, a conserved glutamate residue (Glu151 in *sfp*) was also found to deprotonate the conserved serine in ACPs. There are two acidic residues (Asp107 and Glu109) that coordinate the  $Mg^{2+}$  ion in *sfp*-type PPTases and three acidic residues (E1817, D1772, E1774) in type I integrated PPTases [45,61,62]. Mutagenesis studies have also revealed that AcpS can activate PCPs, if the PCP sequence becomes more similar to the ACP sequence [63].

In addition to the natural substrate CoA, some PPTases can modify CPs with various CoA analogues, where the PPant arm carries non-natural molecules [64,65] and fluorescent probes [66,67]. Based on the ability of some PPTases to use non-natural CoA analogues, as part of this thesis, a method was developed to monitor the activation of CPs by PPTases using a modified HS-CoA with an alkyne group, which allows for Cu-catalyzed alkyne-azide cycloaddition with a fluorescent tag. This provides a potential tool for studying PPTase-CP interactions and will be detailed in Chapter 2.

### 1.1.2 Acyltransferases

Acyltransferases play an important role in initiating polyketide assembly within polyketide biosynthetic pathways. In *cis*-AT type I PKSs, acyltransferases are part of multifunctional enzymes, whereas in systems like type II PKSs and *trans*-AT PKSs, they exist as monofunctional enzymes that work independently within the biosynthetic pathway. Comprising approximately 350 residues, the AT domain catalyzes the first essential step by selecting and transferring activated building blocks to the corresponding ACP domain [68,69]. As described above, in *cis*-AT PKSs, the AT domain is linked within multifunctional PKSs domains. However, in *trans*-AT PKSs, the AT domain exists as a standalone monofunctional enzyme [29,30]. There are two clades of AT domains in *cis*-AT PKSs, the first clade of AT domains utilizes malonyl-CoA as the substrate, and the second clade can use methylmalonyl-CoA and some rarer substrates [40,70,71]. The two most commonly used extender units, malonyl-CoA and methylmalonyl-CoA, can be distinguished by characteristic motifs in the AT domain. Specifically, malonyl-CoA is recognized by the HAFH motif, while methylmalonyl-CoA is identified by the YASH motif, although VASH can occasionally substitute for YASH. These amino acid sequences play a critical role in determining the substrate specificity of AT domains by enabling selective recognition and binding of the extender unit [72,73].

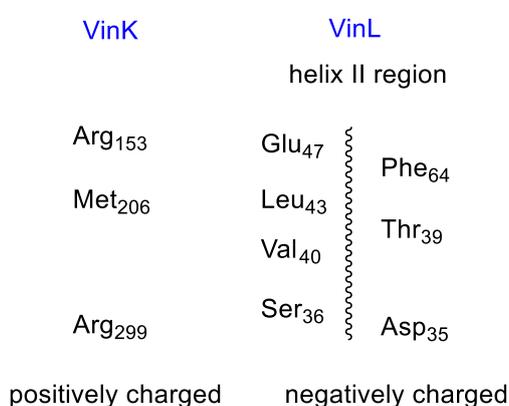


**Scheme 16.** Catalytic mechanism of acyltransferase. AT domains act in a ping-pong mechanism: first, the AT domain becomes self-acylated to form acyl-O-AT. Then, acyl groups are transferred to *holo*-ACP, forming acyl-S-ACP via nucleophilic attack by the thiol group of ACP domain.

The catalytic mechanism of AT domains follows a ping-pong mechanism that consists of two steps (Scheme 16). In the first step, which is also known as self-acylation, the electrophilic (methyl)malonyl group is loaded to the active site serine and forms acyl-O-AT intermediates. The second step, called *trans*-acylation, involves the acyl units

from the first step that are transferred to a *holo*-ACP to form an acyl-S-ACP. This results in the formation of an acyl-enzyme intermediate that undergoes nucleophilic attack by the thiol residue of the pantetheinyl side chain of the ACP [74-78]. The active site key residue serine in the GHSxG motif is very important for the AT domain [79]. Acyltransferases are responsible for substrate recognition in polyketide synthases, and through enzyme-directed mutagenesis, different substrates can be introduced to produce analogs of natural products [80,81].

Although there are several cases, in which loading different acyl units on natural AT domains could produce structurally diverse polyketides, substrate specificity is still a main problem for acyltransferase research [69,81-86]. Many polyketide synthases rely on protein-substrate interactions and protein-protein interactions between ACP and AT domains to select and install extender units [87,88]. The crystal structure of the AT-ACP complex (VinK-VinL) provides insights into how acyltransferases recognize ACPs. Specifically, the structure reveals interactions between certain residues of the AT domain VinK and the ACP VinL (Scheme 17). Key residues in VinK, such as positively charged Arg153 and Arg299, in addition to Met206, interact with the negatively charged helix II region of VinL [87]. Arginine or lysine residues are generally found at the positively charged surface close to the substrate-binding tunnel opening in the acyl-CoA-specific AT domain [89]. Met206 in VinK interacts with hydrophobic residues of VinL. According to the structure, the negatively charged helix II region on ACP is essential for AT domain recognition. This demonstrates how the AT domain utilizes distinct structural features to recognize its corresponding ACP, highlighting the importance of specific protein-protein interactions in the substrate recognition and binding process [87].

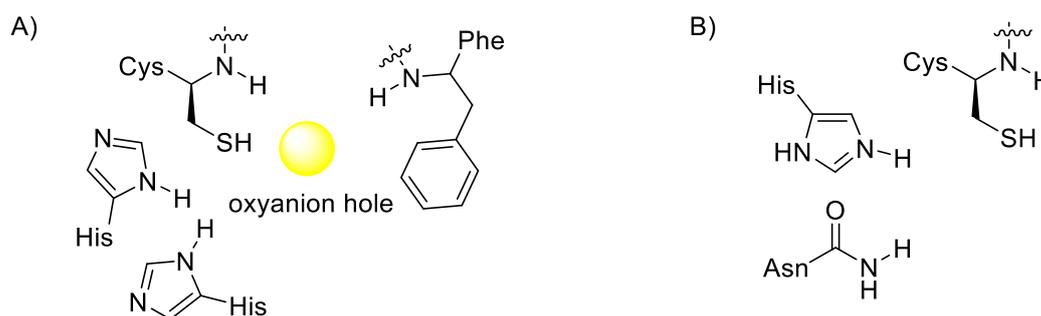


**Scheme 17.** Key residues in the AT-ACP complex (VinK-VinL).

### 1.1.3 Ketosynthases

Ketosynthases are found in both multifunctional and monofunctional systems involved in the biosynthesis of fatty acids and polyketides, where they facilitate the formation of new carbon–carbon bonds through decarboxylative Claisen condensation. This reaction occurs between acyl-ACP (or acyl-CoA) and malonyl-ACP (or malonyl-CoA), resulting in the production of a 3-ketoacyl-ACP (or CoA) and CO<sub>2</sub>, thereby extending the chain [90].

A catalytic triad of a cysteine and two histidine residues defines the active site of KS domains. Additionally, there is a conserved phenylalanine residue which can create an oxyanion hole, when combined with the amides in the catalytic cysteine backbone. This oxyanion hole is crucial for stabilizing intermediates during the enzyme reactions (Scheme 18A) [91-93]. Unlike the Cys-His-His catalytic triad found in type I and II PKSs, type III PKSs have a distinct Cys-His-Asn catalytic triad, with the cysteine residue serving as the nucleophile (Scheme 18B) [50].

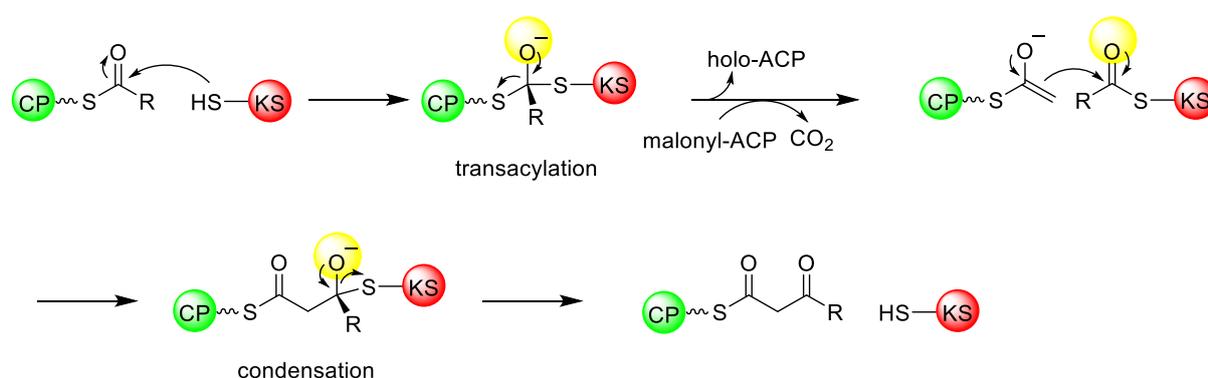


**Scheme 18.** The active site of KS domains. A) Cys-His-His catalytic triad and oxyanion hole in type I and II PKSs; B) Cys-His-Asn catalytic triad in type III PKSs.

In polyketide biosynthesis, the KS domain functions in a ping-pong mechanism, encompassing two main steps: transacylation and condensation. In the first step, an acyl intermediate is transferred from the ACP to the active site cysteine residue of the KS domain. When the acyl intermediate interacts with the cysteine thiol, a tetrahedral intermediate forms, which is stabilized by an oxyanion hole. This oxyanion hole is a structural feature that stabilizes the negative charge on the carbonyl oxygen, ensuring the intermediate remains stable and the reaction proceeds efficiently (Scheme 19) [90-94].

Following the formation of the thioester-linked acyl intermediate, the second condensation step starts. This involves the decarboxylative attack by a malonyl-ACP

group (or malonyl-CoA). Two histidine residues within the active site play crucial roles during this step. They facilitate the decarboxylation of the malonyl group, generating a highly reactive enolate intermediate. This enolate then attacks the thioester-bound acyl group, leading to a nucleophilic acyl substitution reaction. The attack by the enolate results in the formation of a new carbon-carbon bond, effectively elongating the acyl chain. Throughout this process, the oxyanion hole keeps stabilizing the negatively charged intermediates, ensuring that the reaction proceeds smoothly and that the product is generated effectively. The final product of this condensation is a  $\beta$ -ketoacyl intermediate (Scheme 19) [90-94].



**Scheme 19.** The KS domain works via a ping-pong mechanism with two main steps: transacylation and condensation. In the transacylation step, an acyl intermediate is transferred from the ACP to the KS domain cysteine, forming a stabilized tetrahedral intermediate with an oxyanion hole. During the condensation step, malonyl-ACP or malonyl-CoA decarboxylates to form an enolate, which attacks the acyl intermediate, producing a  $\beta$ -ketoacyl intermediate.

In addition to their natural substrates, some KS domains can accept a broad range of substrates [95,96], which is beneficial for engineering approaches to gain enzymatic access to novel molecules [97,98]. For example, the BaeJ-KS2 domain involved in bacillaene (**20**) biosynthesis can accept unnatural substrates [98]. Based on this properties, a method was developed as part of the experimental work to this thesis by using <sup>13</sup>C-labeled substrate surrogates in conjunction with <sup>13</sup>C-NMR to investigate substrate acceptance by BaeJ-KS2 domain of the bacillaene PKS, which will be described in detail in Chapter 3. The KS domain also plays a role in controlling chain length in polyketide and fatty acid biosynthesis pathways. Based on the analysis of various KS domain crystal structures and substrate-bound KS domain structures, the KS domain pocket is regarded as controlling the product structure through physical

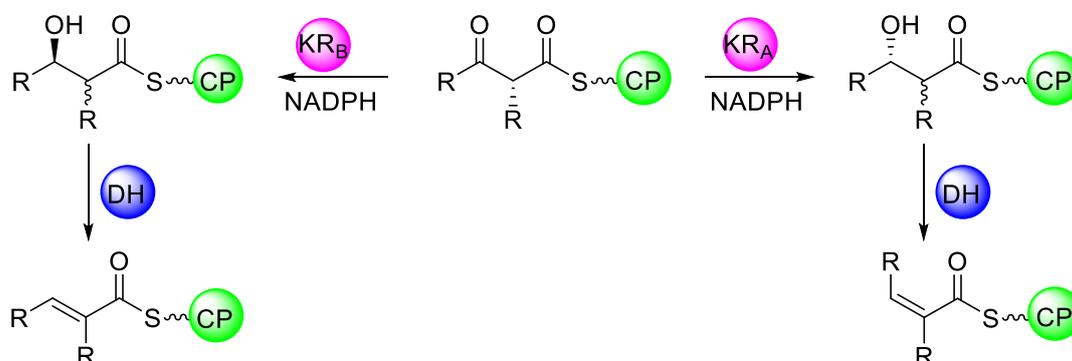
restrictions, polar interactions, and Coulomb repulsion [91]. By some mutational experiments, the conserved phenylalanine acts as a gatekeeper for the substrate during the catalytic step, because it can control malonyl-ACP substrates binding and form the catalytic oxyanion hole [91,99-101]. Similar to PPTase-CP interactions, the interactions between KS domains and CPs are also very important for the accurate assembly of the final product in fatty acid and polyketide biosynthetic pathways. According to some crystal structures of KS-ACP complexes, KS domains usually have a surface around the active site entrance known as the "positive patch", which is rich in arginine and lysine residues. To ensure proper substrate delivery and processing, this positively charged region facilitates interactions with the negatively charged residues on ACP [91,102-104].

#### 1.1.4 Optional modification domains

During polyketide biosynthetic pathways, the  $\beta$ -keto group formed after each carbon chain elongation step can be modified by three optional domains: the KR domain, the DH domain, and the ER domain, which produce hydroxyl, alkenyl or alkyl moieties, respectively [105]. Unlike type I PKSs, which are multifunctional enzymes that may perform multiple steps in a biosynthetic pathway within a single polypeptide chain, type II PKSs consist of separate, monofunctional enzymes. Each domain (KR, DH, ER) is a standalone, monofunctional protein that carries out a specific catalytic task during the biosynthesis of the polyketide chain. These enzymes work together in a coordinated manner, with each enzyme performing a single function in the overall pathway.

Ketoreductases are enzymes that can stereoselectively reduce a  $\beta$ -keto group to a secondary 3-hydroxyl group, utilizing NADPH as a cofactor. Their ability to catalyze reactions with high stereospecificity makes them important in the biosynthesis of complex molecules. KR domains can be classified into two main types based on the stereochemistry of the resulting alcohols: type A and type B (Scheme 20). Type A KR domains, characterized by a conserved W141 residue, reduce  $\beta$ -ketones to form 3L-configured alcohols and are further divided into type A1, which produces 2D,3L-configured products (for R=Me), and type A2, which generates 2L,3L-configured products (for R=Me) and includes an additional conserved residue, H146. Conversely, type B KR domains are distinguished by the conserved residues LDD (93-95), P144, and N148, and can produce 3D-configured alcohols. Type B KR domains can also be divided into

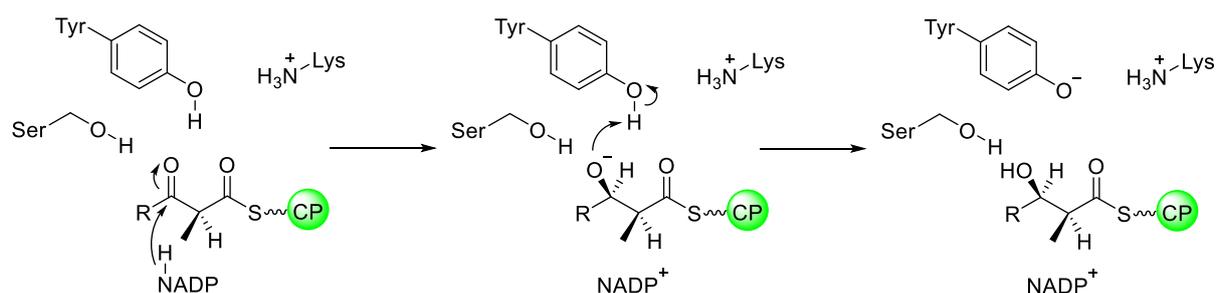
type B1 and type B2. Type B1 KRs produce 2D,3D-configured products (for R=Me), while type B2 KRs, featuring an extra conserved P151 residue, result in 2L,3D-configured products (for R=Me) (Table 2) [106,107].



**Scheme 20.** Ketoreductase (KR) and Dehydratase (DH) catalytic mechanisms. KRs reduce  $\beta$ -keto groups to 3-hydroxyl groups with NADPH, categorized into type A (3L alcohol) and type B (3D alcohol). DHs then form *cis* or *trans* double bond intermediates, respectively (R=Me or H).

**Table 2.** KR types and corresponding products.

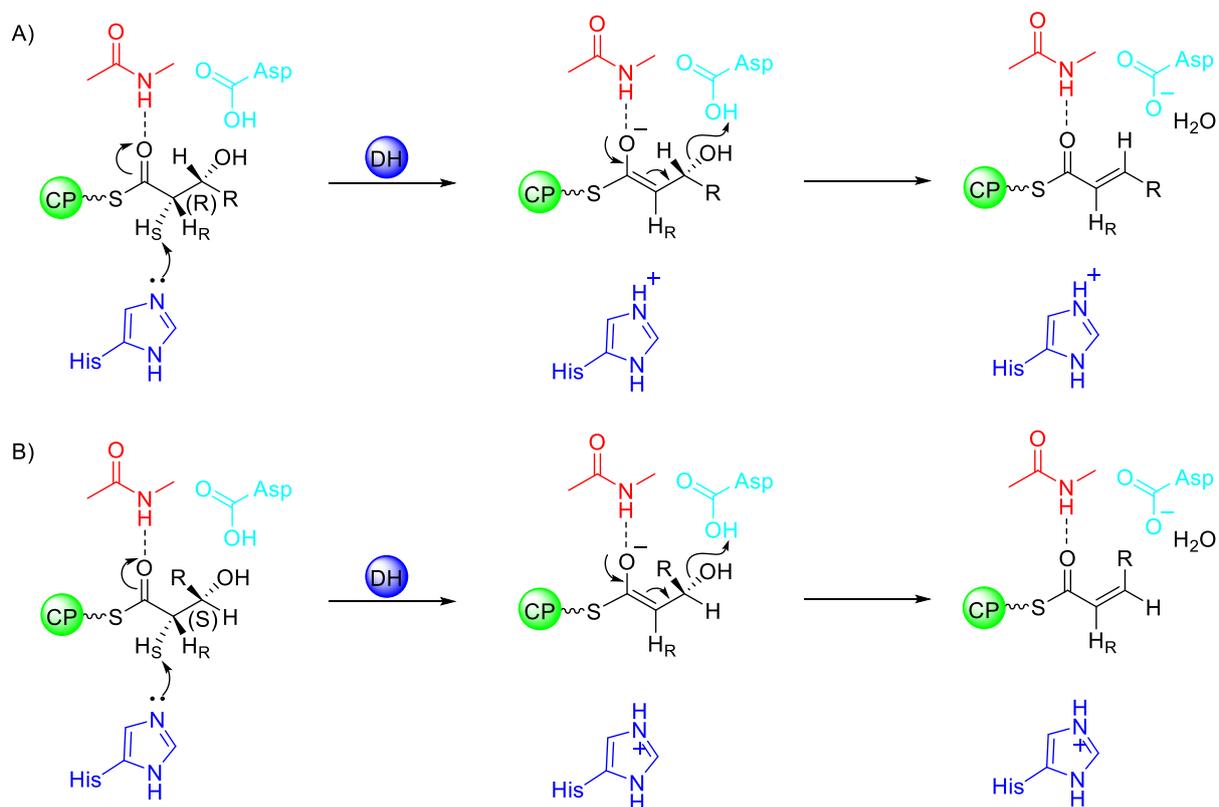
| KR                 | Type A1 | Type A2    | Type B1                | Type B2                |
|--------------------|---------|------------|------------------------|------------------------|
| conserved residues | W141    | W141, H146 | LDD(93-95), P144, N148 | LDD(93-95), P144, N148 |
| products           | 2D,3L   | 2L,3L      | 2D,3D                  | 2L,3D                  |



**Scheme 21.** Reduction catalyzed by the KR. The  $\beta$ -keto group of the diketide intermediate bound by Ser and Tyr, is attacked by the NADPH-derived hydride. The oxygen accepts a proton from Tyr, resulting in a  $\beta$ -hydroxyl intermediate.

In a comparison of several crystal structures of KRs in fatty acid synthases and polyketide biosynthesis modules, it is found that a bigger C-terminal catalytic subdomain (KRc) and a N-terminal structural subdomain (KRn) make up each KR.

Despite lacking catalytic residues, KRc stabilize the KRc subdomain. Several residues are crucial in the KR catalytic mechanism. Tyrosine and serine interact with the  $\beta$ -keto group, with tyrosine donating a proton to the carbonyl oxygen and NADPH transferring a hydride to the  $\beta$ -carbon. A nearby lysine can lower the pKa of the tyrosine (Scheme 21) [108-110].



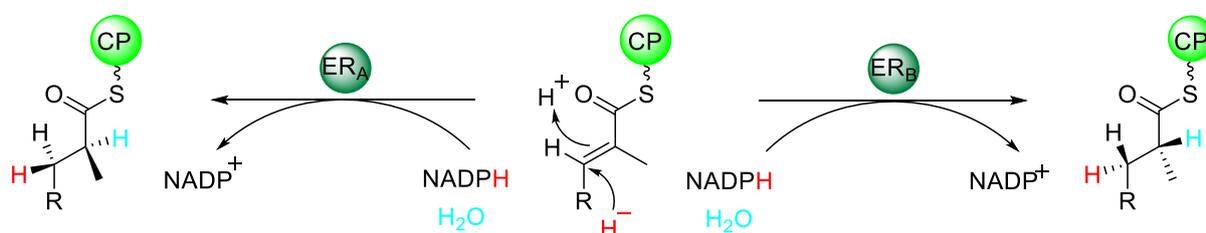
**Scheme 22.** Dehydratases (DHs) catalyze water eliminations to form double bonds in polyketide. (A) DHs act on (3*R*)-hydroxyacyl intermediates to form *trans* double bonds via *syn*-elimination, and (B) on (3*S*)-hydroxyacyl intermediates to produce *cis* double bonds.

Dehydratases catalyze the elimination of water to produce olefinic double bonds (Scheme 20). In polyketide and fatty acid biosynthetic pathways, DH domains catalyze the *syn*-elimination of (3*S*)-hydroxyacyl intermediates which are from A-type KR domains to produce *cis*  $\Delta^2$ -acyl double bonds, and (3*R*)-hydroxyacyl intermediates which are from B-type KR domains to get *trans*  $\Delta^2$ -acyl double bonds. The structural analysis of the DH domain in the curacin synthase reveals the catalytic mechanism (Scheme 22). Firstly, substrate activation begins with a hydrogen bridge between the amide NH of a conserved Pro–Gly/Ala/Ser motif in DH domains and the thioester carbonyl group in the intermediate. This stabilizes the substrate, enabling a

conserved histidine residue to abstract a proton from the  $\alpha$ -position. Subsequently, a conserved aspartate protonates the  $\beta$ -hydroxyl group, facilitating the *syn*-elimination of water and the formation of a double bond. The geometry of this double bond, whether *cis* or *trans*, depends on the stereochemistry of the initial 3-hydroxy group, highlighting the critical interactions between the DH and KR domains within polyketide biosynthesis pathways [105,107,111].

To gain potential insights into the DH domain, isotopically labeled probes were developed to track the stereochemistry of the double bond catalyzed by the DH domain. This is part of this thesis experimental work and will be discussed in detail in Chapter 4. Additionally, several substrates were synthesized to explore the substrate scope of various DH domains in this thesis, which will be introduced in Chapter 5.

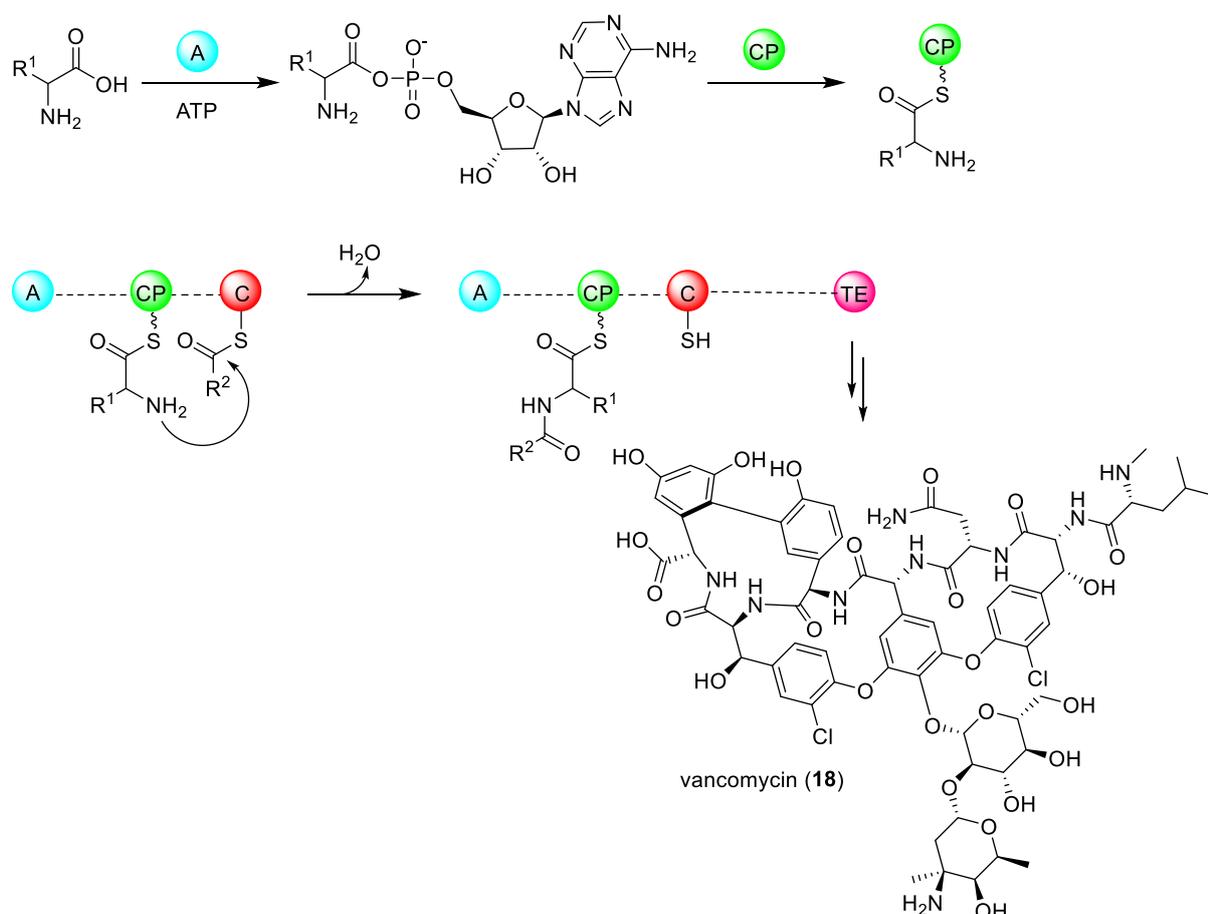
Enoyl reductases in polyketide synthases are responsible for the stereospecific reduction of  $\alpha,\beta$ -unsaturated intermediates using NADPH as a cofactor. During this reduction, a hydride from NADPH is added to the unsaturated thioester intermediate by Michael addition. This is followed by stereospecific protonation at the  $\alpha$ -position of  $\alpha,\beta$ -unsaturated intermediates, which also determines the configuration at  $C_\alpha$ . Similar to KR domains, ER domains in PKSs can also be divided into type A and type B. Type A ERs, characterized by a conserved Y52 residue, can produce L-configured  $\alpha$ -substituted intermediates from  $\alpha,\beta$ -unsaturated intermediates, while type B ERs with a conserved lysine at the corresponding position of Y52 will generate D-configured  $\alpha$ -substituted intermediates (Scheme 23) [107,112]. Both type ERs share a conserved HAAAGGVGMA sequence that was identified as the NADPH binding site [112,113]. The proton donor in the reductase reaction is likely a water molecular, with structural elements ensuring proper positioning for catalysis (Scheme 23) [114].



**Scheme 23.** Enoyl reductases reduce  $\alpha,\beta$ -unsaturated intermediates using NADPH. Type A ERs yield L- $\alpha$ -intermediates, while type B ERs produce D- $\alpha$ -intermediates.

## 1.2 Non-ribosomal peptide biosynthesis

Non-ribosomal peptide synthases (NRPSs) and polyketide synthases exhibit high similarities in their domain functions and biosynthetic processes. However, unlike PKSs, which can include monofunctional enzymes such as those in type II and type III PKSs [48,50], NRPSs are always multifunctional enzymes. This allows NRPSs to carry out all steps of complex biosynthetic pathways within a single protein, making them highly efficient in producing diverse bioactive compounds such as vancomycin (**18**) [115-122].



**Scheme 24.** Non-ribosomal peptide assembly line.

The core domains of typical NRPS modules include three domains: the condensation (C) domain, the adenylation (A) domain, and the peptidyl carrier protein (PCP) domain, also known as the thiolation (T) domain (Scheme 24). These domains are responsible for peptide bond formation. The biosynthesis of non-ribosomal peptide starts with the A domain, which selects and activates amino acids by forming aminoacyl-adenylates in an ATP-dependent reaction. These activated amino acids are subsequently transferred to the thiol group of a phosphopantetheinyl arm on the

PCP domain, forming thioester-linked intermediates. The C domain then catalyzes peptide bond formation between the activated amino acid on the PCP domain and the growing peptide chain, tethered to the PCP domain of the preceding module [115-122].

In addition to the core domains, NRPS modules also include optional domains that introduce structural diversity. For example, epimerization (E) domains catalyze the conversion of L-amino acids to D-amino acids, while N-methyltransferase (N-MT) domains introduce methyl groups to nitrogen atoms within amino acids. Certain C domains may be replaced by Cy domains, which mediate both condensation and cyclization of serine, cysteine, or threonine residues. The final peptide is typically released by a thioesterase (TE) domain, which cleaves the peptide via hydrolysis or facilitates cyclization to form cyclic peptides [115-122].

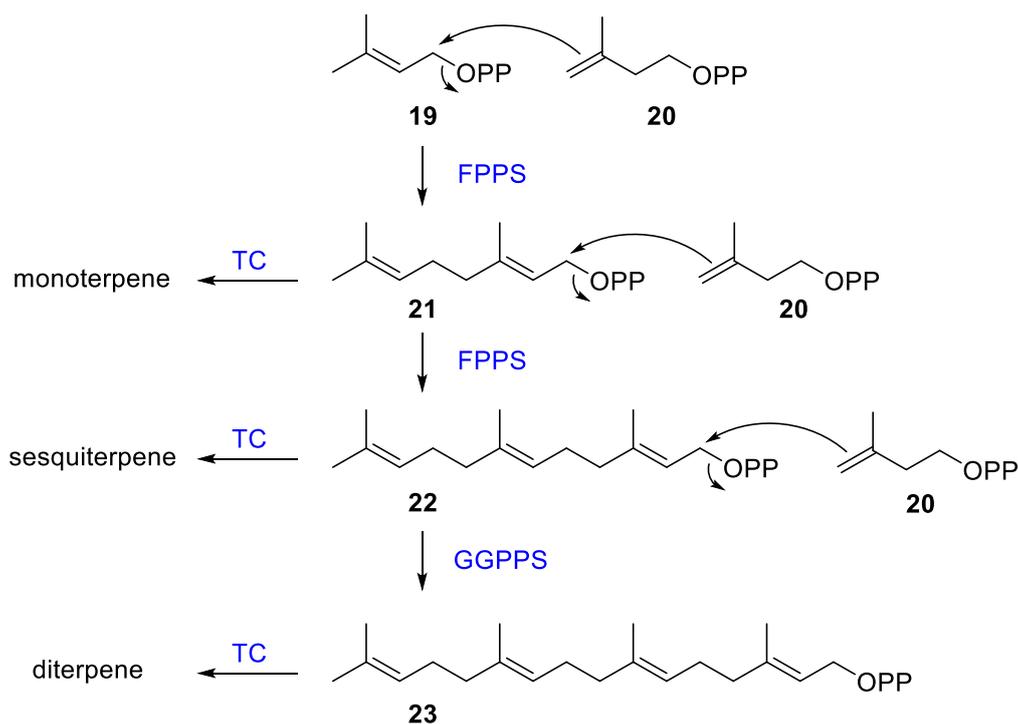
Non-ribosomal peptide biosynthesis has been extensively studied through in vitro systems and structural analyses, providing a deep understanding of its mechanisms. This knowledge has enabled bioengineering approaches to create novel non-ribosomal peptides. Strategies include swapping A domains or A-T bidomains, replacing complete modules, altering binding pocket amino acids to modify substrate specificity, and introducing new domains or modules. These modifications offer significant potential for generating diverse non-ribosomal peptides with tailored biological activities [123].

### **1.3 Multifunctional enzymes in terpene biosynthesis**

In addition to polyketide and non-ribosomal peptide biosynthetic pathways, multifunctional enzymes also play significant roles in terpene biosynthesis. Terpenes are a large group of natural products known for their diverse structures and wide biological activities. Terpene biosynthesis typically make use of monofunctional enzyme, but multifunctional enzyme is also known in some terpene cases, such as geosmin synthase, copalyl diphosphate synthase (CDS) and abietadiene synthase (AS) [124-131].

There are two basic units in the terpene biosynthetic pathway: dimethylallyl diphosphate (DMAPP, **19**) and isopentenyl diphosphate (IPP, **20**) that are produced through the mevalonate (MVA) or the methylerythritol 4-phosphate (MEP) pathway [132,133]. For terpene biosynthesis, the first steps in the production of almost all natural

terpenes are usually the same. To generate acyclic isoprenoid compounds such as geranyl diphosphate (GPP, **21**), farnesyl diphosphate (FPP, **22**), or geranylgeranyl diphosphate (GGPP, **23**), dimethylallyl diphosphate (DMAPP, **19**) is first coupled head-to-tail with one or more isopentenyl diphosphate (IPP, **20**) units via prenyltransferases (PTs). Terpene synthases (TSs) then use these acyclic isoprenoid diphosphates as substrates to convert them into different hydrocarbon products, such as monoterpenes, sesquiterpenes and diterpenes (Scheme 25) [124,134,135].

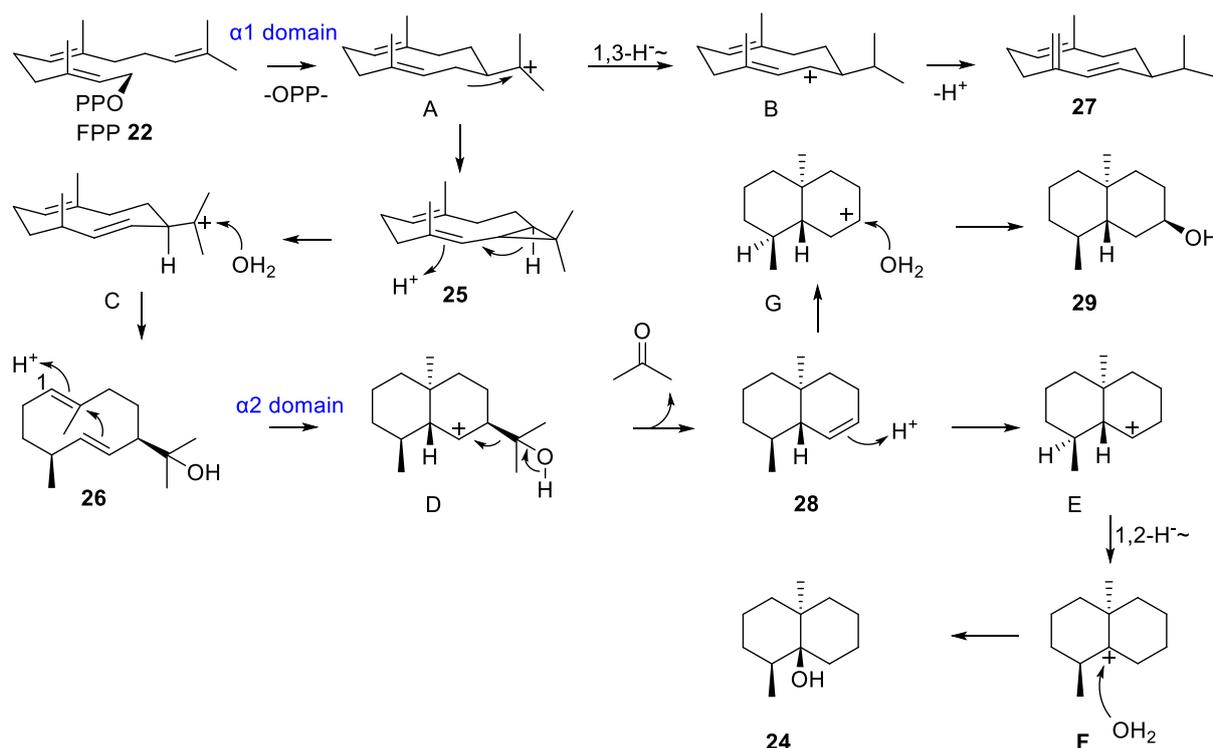


**Scheme 25.** Biosynthesis of terpenes: DMAPP (**19**) reacts with IPP (**20**) to form GPP (**21**) for monoterpene production. GPP is elongated to FPP (**22**) by FPP synthase (FPPS), leading to sesquiterpenes. Further elongation to GGPP (**23**) by GGPP synthase (GGPPS) yields diterpenes.

A distinctive feature of the reactions catalyzed by terpene cyclases is the formation of carbocationic intermediates. Terpene cyclases can be divided into two classes based on their mechanism of carbocation formation: class I and class II. In class I terpene cyclases, carbocations are formed through the metal-triggered dissociation of isoprenoid diphosphates, using conserved DDXXD and NSE/DTE motifs to coordinate  $Mg^{2+}$ , which is essential for the reaction. This process produces an allyl cation and inorganic pyrophosphate (PPi). However, class II terpene cyclases form carbocations by protonating the terminal double bond or epoxide moiety of

isoprenoids, with a central aspartate in the conserved DXDD motif serving as the acid catalyst [124,136].

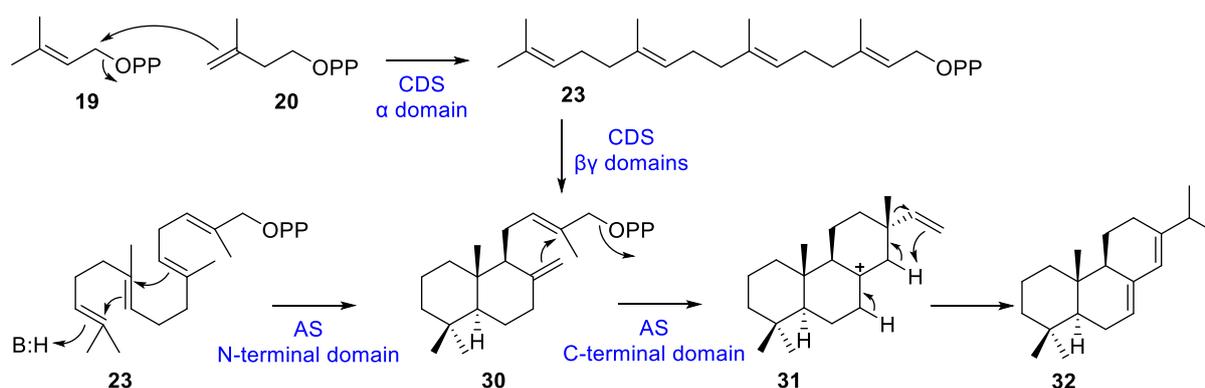
Multifunctional terpene synthases are often present as bifunctional terpene synthases (BTSs), which combine two distinct catalytic activities within a single enzyme [124,137]. BTSs are primarily categorized into two types based on their structural characteristics. Type I BTSs feature an  $\alpha\alpha$  domain architecture, where each  $\alpha$  domain performs a distinct role in the biosynthetic process [124,137]. For example, geosmin synthase from *Streptomyces coelicolor*, which has an  $\alpha\alpha$  domain architecture, catalyzes a class I cyclization of FPP to (1(10)*E*,5*E*)-germacradien-11-ol (**26**) in the N-terminal  $\alpha_1$  domain and a transannulation–fragmentation reaction in the C-terminal  $\alpha_2$  domain to yield geosmin (**24**) and acetone (Scheme 26) [125-126]. In comparison, type II BTSs have an  $\alpha\beta\gamma$  domain architecture, as exemplified by copalyl diphosphate synthase and abietadiene synthase [128-131].



**Scheme 26.** Geosmin biosynthesis pathway. Geosmin biosynthesis starts with FPP cyclizing to the germacradienyl cation (**A**), which is deprotonated to isolepidozene (**25**). Reprotonation, ring opening and attack of water lead to (1(10)*E*,5*E*)-germacradien-11-ol (**26**). Further reprotonation and ring closure produce an octalin intermediate (**28**), which undergoes reprotonation to form isogeosmin (**29**), or reprotonation, hydride shifts and water capture to yield geosmin (**24**). Germacrene D (**27**) forms via a 1,3-hydride shift and deprotonation from **A**.

Geosmin synthase uses its N-terminal  $\alpha$  domain to catalyze the class I cyclization of farnesyl diphosphate (FPP, **22**), producing the intermediate isolepidozene (**25**). Following a reprotonation and ring opening to **C**, capture with water leads to (1(10)*E*,5*E*)-germacradien-11-ol (**26**). A side product of the geosmin synthase N-terminal  $\alpha$  domain is germacrene D (**27**), which is formed through a 1,3-hydride shift from **A** to **B**, followed by deprotonation. The C-terminal  $\alpha$  domain of geosmin synthase then catalyzes the conversion of intermediate **26** into octalin (**28**) and acetone through a reprotonation and ring closure to **D**, which involves a retro-Prins fragmentation. A reprotonation of **28** results in **E**, followed by a 1,2-hydride shift to **F**, and capture with water yields geosmin (**24**). However, during the investigation of geosmin synthase, a new enzyme product, isogeosmin (**29**), was also found in this thesis. The formation of **29** can be explained by an alternative reprotonation of **28** to **G** and subsequent capture of water (Scheme 26) [125-127]. This will be discussed in detail in Chapter 6.

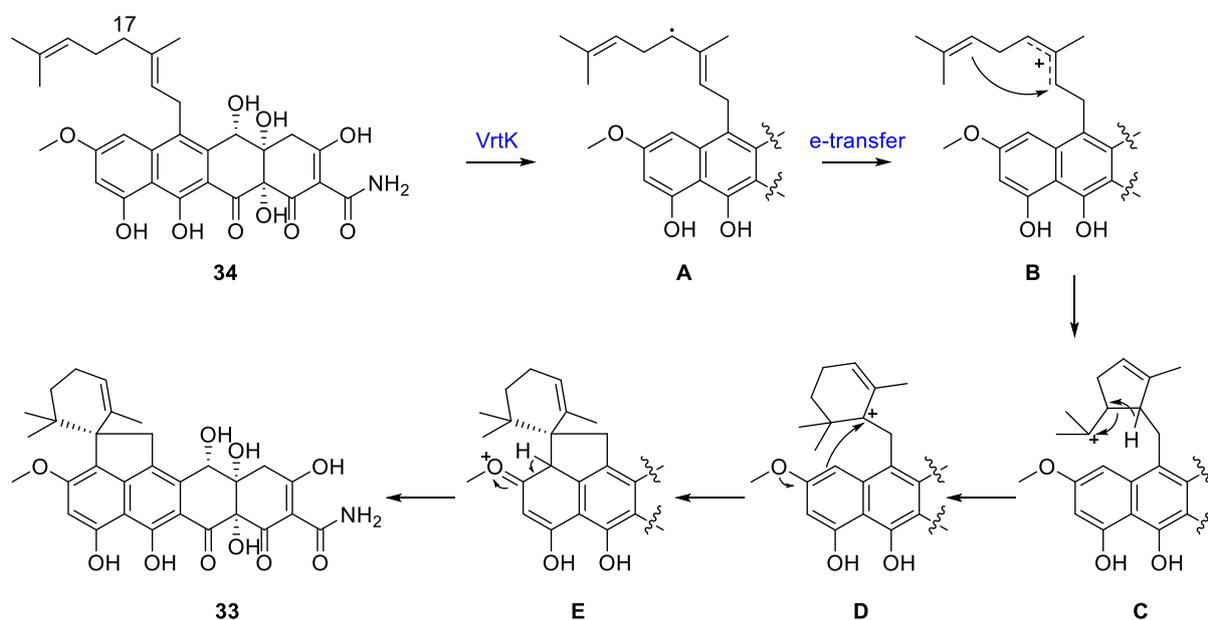
Copalyl diphosphate synthase (CDS) is a type II BTSs from *Penicillium verruculosum*. In this enzyme, the C-terminal  $\alpha$  domain catalyzes the production of geranylgeranyl diphosphate (GGPP, **23**), which is subsequently cyclized by the N-terminal  $\beta\gamma$  domains, functioning as a class II cyclase, to form copalyl diphosphate (**30**) (Scheme 27) [128-130].



**Scheme 27.** Two type II BTSs: CDS and AS. The  $\alpha$  domain of CDS catalyzes the conversion of DMAPP (**19**) and IPP (**20**) to GGPP (**23**), while the  $\beta\gamma$  domains catalyze the conversion of GGPP (**23**) into copalyl diphosphate (**30**). Similarly, in AS, the N-terminal domain converts GGPP (**23**) into copalyl diphosphate (**30**), and the C-terminal domain further catalyzes its transformation into abietadiene (**32**).

Abietadiene synthase (AS) is a bifunctional terpene synthase. It catalyzes the conversion of geranylgeranyl diphosphate (GGPP, **23**) into abietadiene (**32**) through

two distinct reactions. The N terminal domain initiates cyclization by protonating GGPP to form (+)-copalyl diphosphate, which then transfers to the C terminal domain. There, it undergoes ionization and further cyclization to produce (-)-abietadiene (**32**) (Scheme 27) [131].

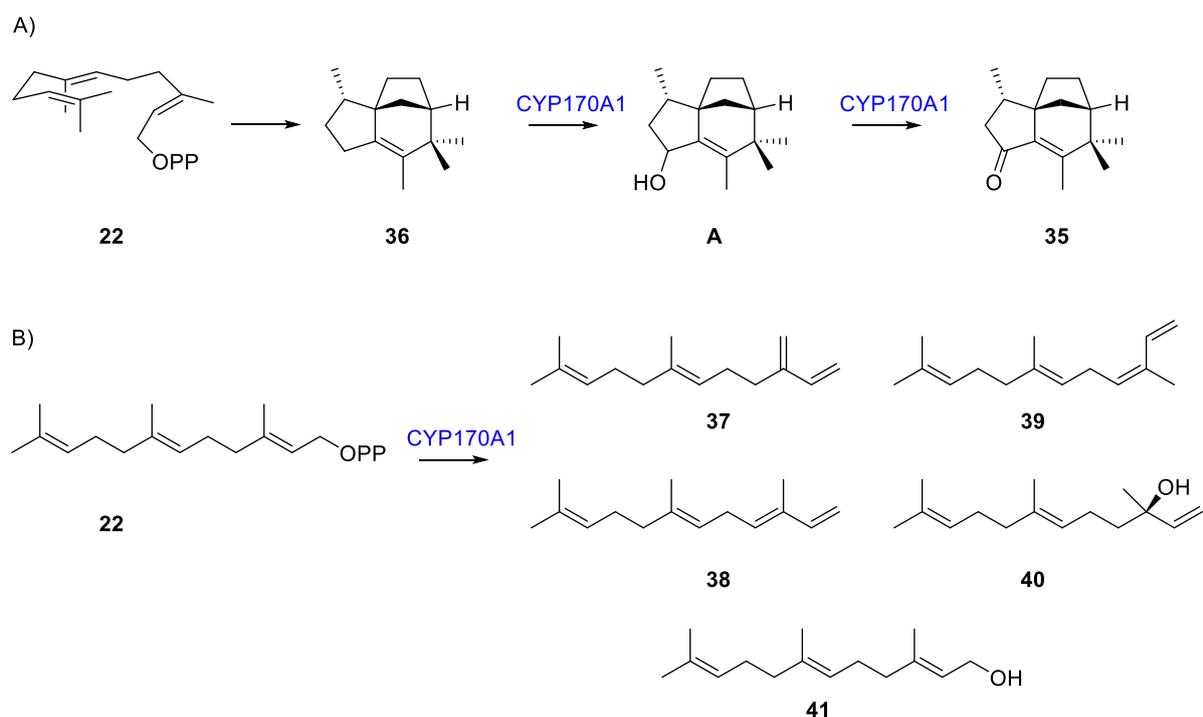


**Scheme 28.** The cyclization of pre-viridicatumtoxin (**34**) by cytochrome P450 enzyme VrtK in viridicatumtoxin (**33**) biosynthesis. A carbocation is proposed to be generated by oxidation of an allylic carbon (C17), corresponding to C4 of geranyl moiety.

The biosynthesis of viridicatumtoxin (**33**), a tetracycline-like meroterpenoid antibiotic produced by *Penicillium aethiopicum*, involves a remarkable bifunctional enzyme, VrtK. This enzyme combines the catalytic capabilities of a terpene cyclase and a cytochrome P450 (CYP450), reflecting the evolutionary versatility of CYP450 enzymes in natural product biosynthesis. VrtK plays a crucial role in the construction of viridicatumtoxin complex structure by catalyzing two key transformations: the cyclization of the geranyl side chain and a Friedel–Crafts alkylation on the naphthacene core (Scheme 28). Unlike typical terpene cyclizations mediated by dedicated terpene synthases, viridicatumtoxin biosynthesis lacks identifiable terpene synthase genes within its gene cluster, suggesting an alternative enzymatic strategy. VrtK initiates its dual function by generating a carbocation intermediate, likely through a hydrogen abstraction and electron transfer mechanism characteristic of P450-mediated pathways. This carbocation undergoes cyclization, followed by 1,3-hydride shifts, leading to a stabilized tertiary carbocation. This reactive intermediate facilitates

the Friedel–Crafts alkylation, completing the structural elaboration of viridicatumtoxin [138,139].

There is another cytochrome P450 enzyme, CYP170A1, which is involved in the biosynthesis of albaflavenone (**35**). As a monooxygenase, CYP170A1 catalyzes the oxidation of (+)-*epi*-isozizaene (**36**) into albaflavenone (**35**). Surprisingly, it also exhibits a terpene synthase activity, enabling it to directly convert farnesyl diphosphate (FPP, **22**) into farnesene isomers (*E*- $\beta$ -farnesene (**37**), (3*E*,6*E*)- $\alpha$ -farnesene (**38**), (3*Z*,6*E*)- $\alpha$ -farnesene (**39**), nerolidol (**40**), and farnesol (**41**) (Scheme 29) [140-143]. This terpene synthase activity relies on a distinct active site within the P450 structure, which contains Mg<sup>2+</sup>-binding DDxD and DTE motifs [142]. The two different activities of CYP170A1, which include oxidation and cyclization reactions, represent a rare example of a monooxygenase acting as a terpene synthase, showcasing its remarkable multifunctionality in natural product biosynthesis.



**Scheme 29.** The monooxygenase activity of CYP170A1. A) Biosynthesis of albaflavenone (**35**) from FPP (**22**). B) Farnesene synthase activity of CYP170A1.

## 1.4 Conclusion

Multifunctional enzymes present both clear advantages and significant challenges in natural product biosynthesis. On the one hand, their ability to catalyze multiple steps within a single protein ensures efficient coordination of reactions, reducing the loss of intermediates and improving the productivity. A key feature of these enzymes is substrate channeling, where intermediates are transferred directly between active sites. This transfer can occur through well-defined tunnels, where a physical pathway exists between the sites, or through proximity-driven mechanisms, where the enzymes are positioned close together, allowing the intermediates to move between active sites without a specific tunnel. For instance, in tryptophan synthase, indole is transferred from the  $\alpha$ -subunit to the  $\beta$ -subunit through a substrate tunnel, ensuring precise and efficient catalysis <sup>[144]</sup>. Similarly, in multifunctional enzymes like glutamate synthase <sup>[145]</sup> and formylglycinamide ribonucleotide amidotransferase involved in purine nucleotide biosynthesis <sup>[146]</sup>, a substrate tunnel facilitates the transport of ammonia. In these enzymes, one domain releases ammonia through the hydrolysis of glutamine. The ammonia is then transferred to another domain, where it is incorporated into a substrate molecule.

In contrast, fusicoccadiene synthase (PaFS) demonstrates a different mechanism of substrate channeling. The bifunctional enzyme PaFS contains a prenyltransferase domain responsible for the biosynthesis of geranylgeranyl diphosphate (GGPP, **8**) and a terpene synthase domain that catalyzes the cyclization of GGPP into fusicoccadiene. In PaFS, substrate channeling occurs not through a physical tunnel but through the arrangement of multiple terpene synthase domains around the prenyltransferase core. This proximity-based mechanism enables efficient transfer of intermediates, preventing them from diffusing into the surrounding solution <sup>[147]</sup>.

Additionally, the expression of a single multifunctional enzyme simplifies the genetic and metabolic requirements compared to expressing several enzymes. On the other hand, multifunctional enzymes are often large and structurally complex, which can make expression challenging, especially in heterologous systems where folding and stability may be influenced. Moreover, the dependency on all functional domains within a single protein introduces vulnerability: the failure of one domain can disrupt the whole pathway, posing a significant bottleneck in biotechnological applications.

However, the structural complexity of multifunctional enzymes poses challenges,

especially in heterologous systems, where proper folding and stability may be difficult to achieve. Furthermore, their reliance on multiple functional domains within a single protein introduces vulnerabilities; the failure of one domain can disrupt the entire biosynthetic pathway, presenting a significant bottleneck in biotechnological applications.

Multifunctional enzymes are critical in natural product biosynthesis, particularly within polyketide, non-ribosomal peptide and terpene biosynthesis. These enzymes perform multiple catalytic functions, enabling the efficient assembly of complex molecules from simple building blocks. Given the diverse applications of natural products in drug discovery and industry, exploring multifunctional enzymes is essential. The research about multifunctional enzymes in this thesis not only enhances our understanding of natural product biosynthesis but also offers the potential to generate novel compounds, optimizing the use of energy and materials in synthetic biology and biocatalysis.



## Chapter 2

### **A Clickable Coenzyme A Derived Probe for Investigating Phosphopantetheinyl Transferase Activity in Natural Product Biosynthesis**

Zhiyong Yin, Houchao Xu and Jeroen S. Dickschat\*

*Org. Biomol. Chem.* **2024**, 22, 8714–8719.

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

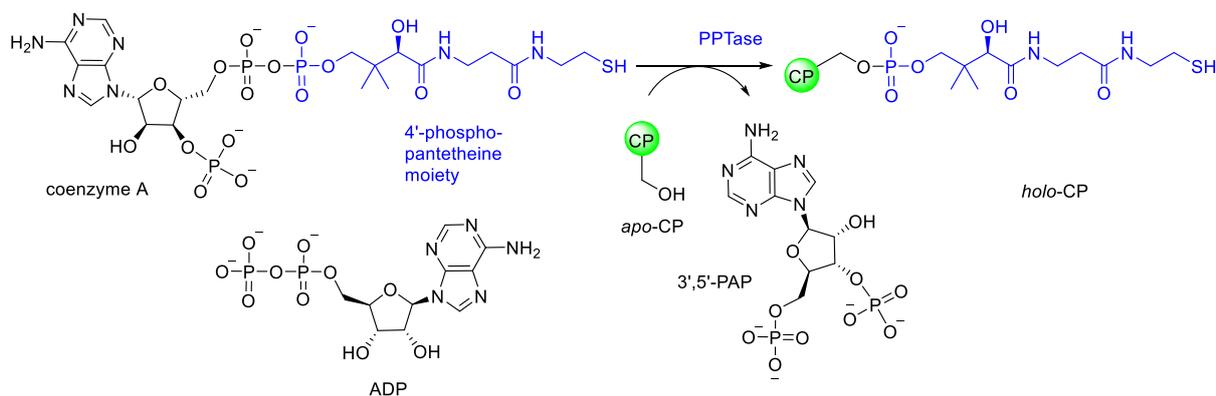
Reprint from *Org. Biomol. Chem.* with kind permission from Royal Society of Chemistry.

The synthesis of compounds and enzyme reactions were performed by me. The enzymes of CoaA, CoaD, and CoaE were provided by Houchao Xu.

The publication “A Clickable Coenzyme A Derived Probe for Investigating Phosphopantetheinyl Transferase Activity in Natural Product Biosynthesis” can be found in Appendix A.

## Introduction

Multifunctional enzymes play a crucial role in the biosynthesis of natural products, particularly in the assembly of complex molecules through coordination of multiple catalytic domains. Natural products are an important source of bioactive compounds, including the antibiotic erythromycin (**10**) [148], which is synthesized by multifunctional polyketide synthases (PKSs), and vancomycin (**18**) [149], produced by multifunctional nonribosomal peptide synthetases (NRPSs). These molecules are assembled through modular processes, with carrier proteins (CPs) responsible for transferring intermediates. Phosphopantetheinyl transferases (PPTases) are essential for activating CPs in the biosynthesis of these natural products [45,46]. PPTases catalyze the attachment of a 4'-phosphopantetheinyl (PPant) group from coenzyme A (HSCoA) to a conserved serine residue on the CPs, converting inactive *apo*-CPs into active *holo*-CPs. This modification introduces a flexible 20 Å PPant arm with a terminal thiol group that can form a thioester bond with intermediates in polyketide and non-ribosomal peptide biosynthetic pathways (Scheme 30) [45].



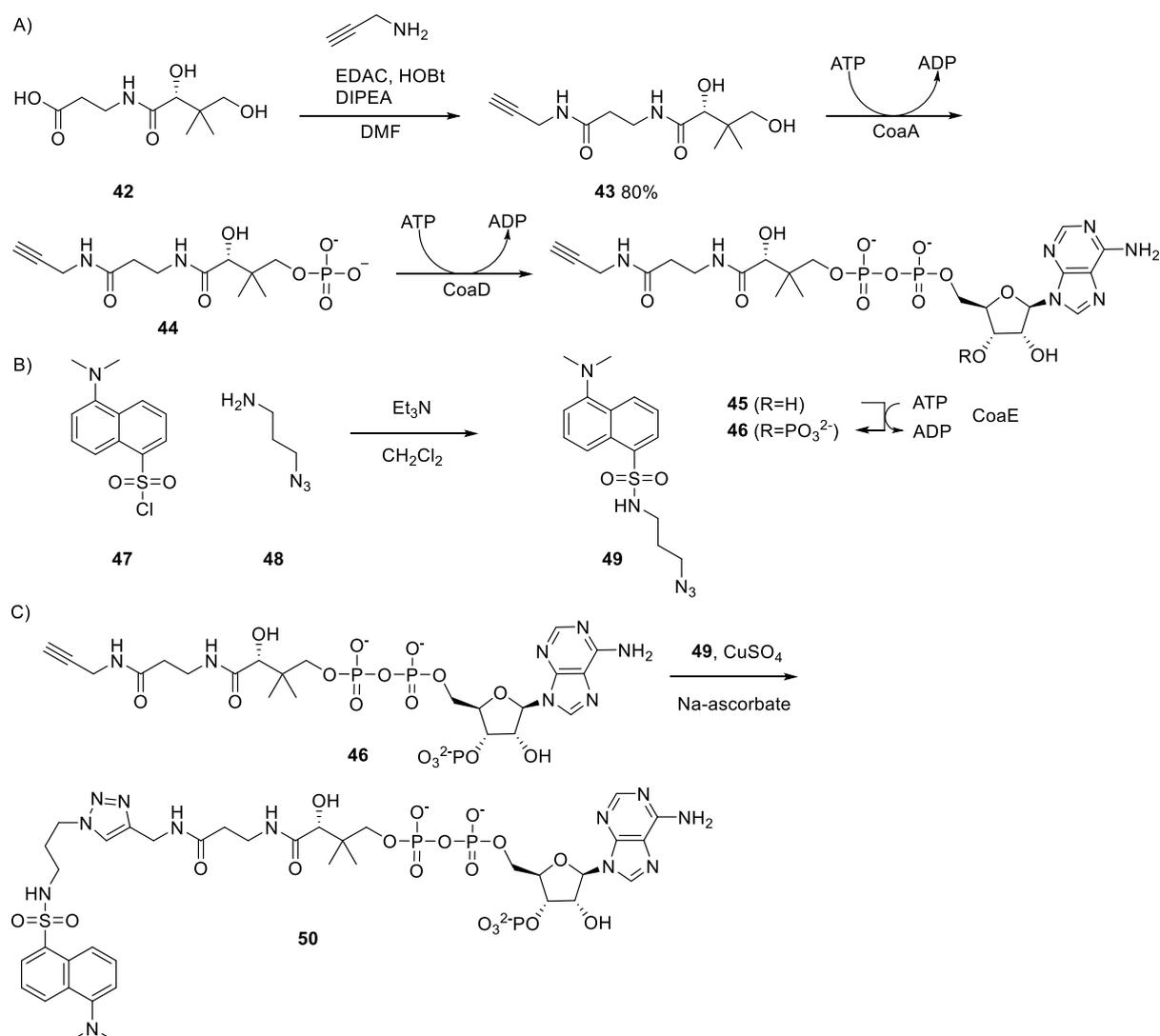
**Scheme 30.** Phosphopantetheinyl transferases activate carrier proteins.

The PPTase Sfp from *Bacillus subtilis*, the enzyme that modifies surfactin synthase [57], is known for its substrate promiscuity and broad use in metabolic engineering. It can modify different CPs and accept a variety of functionalized thioesters, including acetyl-S-CoA and malonyl-S-CoA. This flexibility allows for the attachment of non-natural substrates in mechanistic studies of polyketide and non-ribosomal peptide synthases [66,150-152]. Monitoring the PPTase-catalyzed phosphopantetheinylation of CPs is essential for understanding polyketide and non-ribosomal peptide biosynthesis. Traditional methods, such as the application of radiolabeled HSCoA [46,150] and HPLC-MS analysis [152-154], face challenges related to safety and cost. An

alternative approach using fluorescence labeling has been developed, but it relies on the promiscuous PPTase Sfp [66]. In this study, a new method was developed using an alkyne-functionalized HSCoA derivative, accepted by various PPTases, enabling fluorescence tagging via Cu-catalyzed alkyne-azide cycloaddition (CuAAC).

## Summary

To investigate the activities of PPTases in natural products biosynthesis, the clickable HSCoA derivative **46** was synthesized through a combination of chemical synthesis and enzymatic transformation.

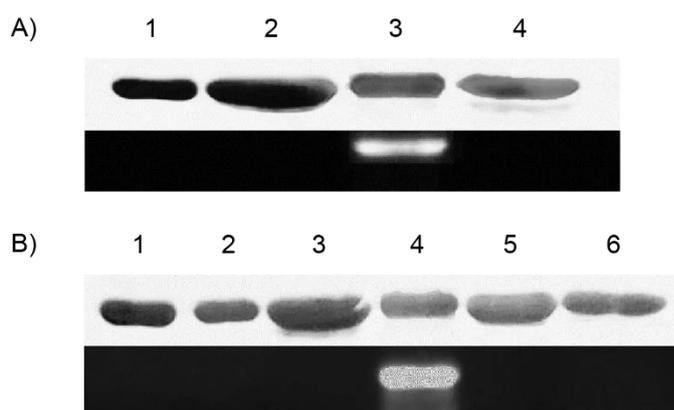


**Scheme 31.** Synthesis of A) the HSCoA derivative **46**, B) the fluorescent probe **49**, and C) the HSCoA derivative **50**.

Starting from pantothenoate (**42**), coupling with propargylamine yielded an intermediate **43** (Scheme 31A). This intermediate **43** was then converted into the HSCoA derivative **46** using the recombinant enzymes CoaA, CoaD, and CoaE from

the *E. coli* coenzyme A biosynthetic pathway (Scheme 31B) [155]. The final product **46** was used as a substrate for various CPs and PPTases. Additionally, the fluorescent probe **49** was synthesized through the amidation of **47** with **48**. Further synthesis also involved the CuAAC reaction between the HSCoA derivative **46** and the probe **49** to obtain a cycloadduct **50** (Scheme 31C).

The initial experiments aimed to investigate the activation of CPs using the promiscuous PPTase Sfp with the recombinant CP (AntG) which is from the antimycin biosynthetic pathway [156]. Although the incubation with the HSCoA analog **46** did not show obvious changes (Figure 1A, lane 2), the successful attachment of a PPant arm analog was confirmed through a CuAAC reaction with azide **49**, resulting in a fluorescent band on the SDS-PAGE (Figure 1A, lane 3). This method can be used for monitoring the activation of CPs by PPTases with HSCoA analog **46**. However, HSCoA analog **50** with substantial structural changes was not accepted by the enzyme system (Figure 1A, lane 4).



**Figure 1.** SDS-PAGE fluorescence assay for the activation of AntG by Sfp. A) Lane 1: purified AntG; lane 2: purified Sfp + AntG + **46**; lane 3: purified Sfp + AntG + **46** + **49** after CuAAC; lane 4: purified Sfp + AntG + **50** after CuAAC. B) Lane 1: purified AntG; lane 2: purified AntG-S34A; lane 3: Sfp + AntG + **46**; lane 4: Sfp + AntG + **46** + **49** after CuAAC; lane 5: Sfp + AntG-S34A + **46**; lane 6: Sfp + AntG-S34A + **46** + **49** after CuAAC. The top rows show the protein gel after coomassie brilliant blue staining, the bottom rows show the same gel before staining under UV irradiation ( $\lambda = 360$  nm).

To confirm substrate attachment to the conserved serine of AntG and to rule out non-covalent interactions, the AntG-S34A variant was obtained through site-directed mutagenesis. Incubation with Sfp and the HSCoA analog **46** showed a fluorescent

band only for wild-type AntG (Figure 1B, lane 4), not for the S34A variant (Figure 1B, lane 6). These results highlight the benefits of the HSCoA analog **46** in monitoring AntG activation by Sfp and its attachment to the highly conserved Ser34.

Following the establishment of the fluorescence assay, additional experiments were conducted to assess the activation of various CPs by Sfp, including AcpP from *E. coli* [157] and JadC from *Streptomyces venezuelae* [158,159]. Both AcpP and JadC were activated with the HSCoA analog **46**, resulting in fluorescent bands after the cycloaddition of azide **49**. Activation was also observed for AntG, AcpP and JadC with KirP [154] and two PPTases from *Streptomyces avermitilis* (accession number BAC69459) and *S. venezuelae* (CCA59476), while other tested PPTases, including EntD [160,161] and several others, failed to activate any of the three CPs.

This study demonstrates the utility of clickable coenzyme A-derived probes for understanding PPTase-mediated activation of CPs in polyketide and non-ribosomal peptide biosynthesis. The activation of various CPs by this clickable HSCoA derivative **46**, including by two previously uncharacterized PPTases, demonstrates its potential in investigating PPTase-CP interactions and broadens the range of promiscuous PPTases beyond Sfp.



## Chapter 3

### Substrate Specificity of a Ketosynthase Domain Involved in Bacillaene Biosynthesis

Zhiyong Yin and Jeroen S. Dickschat\*

*Beilstein J. Org. Chem.* **2024**, *5*, 734–740.

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

Reprint from *Beilstein J. Org. Chem.* with kind permission from Beilstein Institute.

The publication “Substrate Specificity of a Ketosynthase Domain Involved in Bacillaene Biosynthesis” can be found in Appendix B.

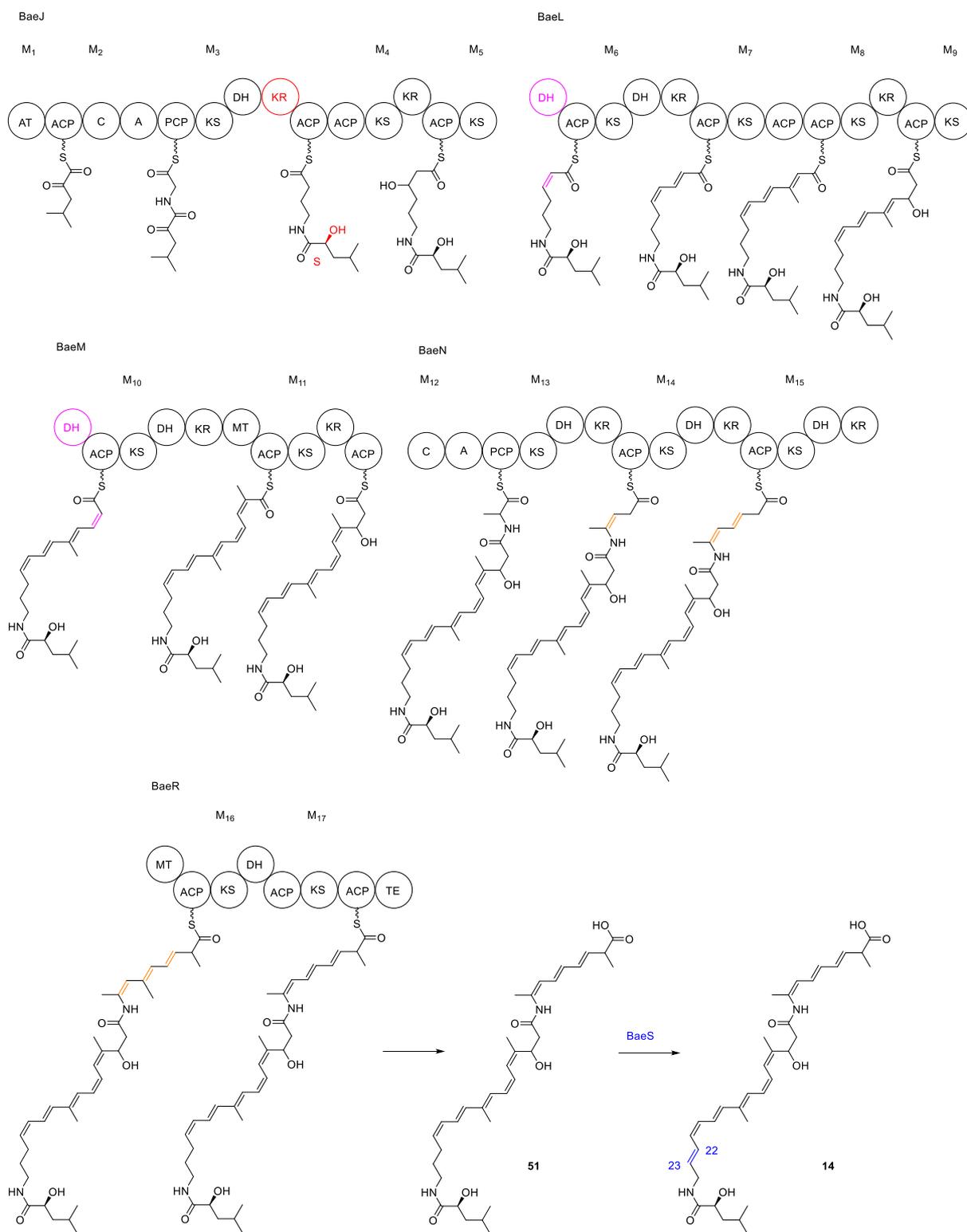
## Introduction

*Trans*-AT polyketide synthases (PKSs) are a class of multifunctional non-iterative type I PKSs. Unlike *cis*-AT PKSs, which contain an integral acyltransferase (AT) domain in each module, *trans*-AT PKSs have independent AT domains to provide extender units for chain elongation [30]. Bacillaene (**14**), an antibiotic isolated from *Bacillus subtilis*, is produced by a multifunctional *trans*-AT PKS-NRPS hybrid system [31,32]. The bacillaene PKS-NRPS features several irregular characteristics, including the absence of a ketoreductase (KR) domain in module 1 and a KR domain in module 3 that performs dual functions: it reduces the  $\beta$ -ketoacyl intermediate during chain elongation and also reduces the  $\alpha$ -ketoisocaproate starter unit to introduce an *S* configured stereocenter (Scheme 32) [162,163].

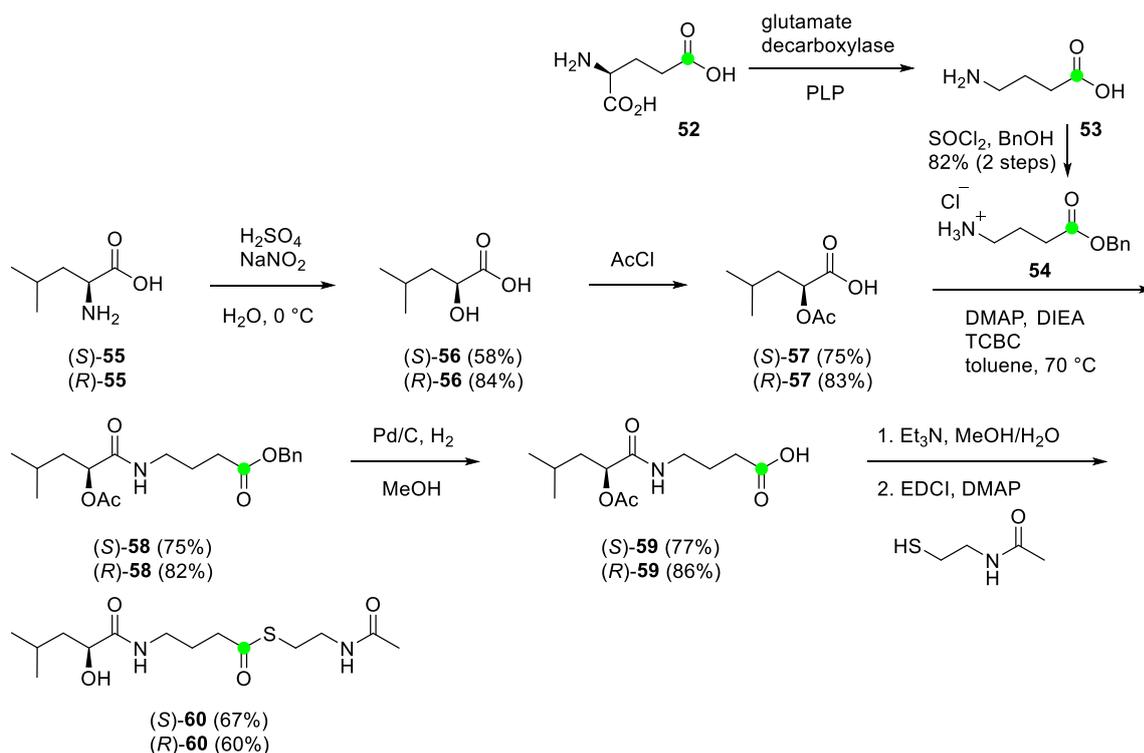
Ketosynthase (KS) domains in *trans*-AT PKSs are responsible for catalyzing the condensation of polyketide intermediates. A phylogenetic analysis of KS domains from *trans*-AT PKSs indicates that these domains form a distinct clade when the structures of their processed substrates are similar. This characteristic can be used for the prediction of KS domain substrate specificity [164]. Although the phylogenetic analysis can provide insights into KS substrate specificity, experimental approaches are still important. In this chapter, the substrate specificity of BaeJ-KS2, a KS domain from module 4 of the bacillaene PKSs, was investigated using a new method involving  $^{13}\text{C}$ -labeled substrate surrogates and  $^{13}\text{C}$ -NMR spectroscopy. This work provides new insights into the substrate processing capabilities of KS domains involved in complex natural product biosynthesis.

## Summary

To explore the substrate specificity of BaeJ-KS2,  $^{13}\text{C}$ -labeled (*S*)-**60** was synthesized as a mimic of the ACP-bound intermediate. Starting with (5- $^{13}\text{C}$ )glutamate (**52**), the glutamate decarboxylase from *Escherichia coli* was cloned and expressed, enabling the conversion of glutamate to  $\gamma$ -aminobutyric acid (**53**). The crude product of enzyme reaction,  $\gamma$ -aminobutyric acid (**53**) was used for an esterification with benzylalcohol to yield the  $^{13}\text{C}$ -labeled ester (**54**), which was then coupled with (*S*)-**57** by using Yamaguchi conditions to produce the amide (*S*)-**58**. After deprotection and esterification with *N*-acetylcysteamine, the desired *N*-acetylcysteamine thioester (SNAC ester) (*S*)-**60** was obtained. The same process yielded (*R*)-**60** from D-leucine (Scheme 33).



**Scheme 32.** Biosynthetic model for bacillaene (**14**). M<sub>1</sub> – M<sub>17</sub> indicate modules 1 – 17. A = adenylation domain, ACP = acyl carrier protein, AT = acyltransferase, C = condensation domain, DH = dehydratase, KR = ketoreductase, KS = ketosynthase, MT = methyltransferase, PCP = peptidyl carrier protein, TE = thioesterase



**Scheme 33:** Synthesis of the BaeJ-KS2 substrate surrogates (S)-**60** and (R)-**60**. Green dots represent  $^{13}\text{C}$ -labeled carbons.

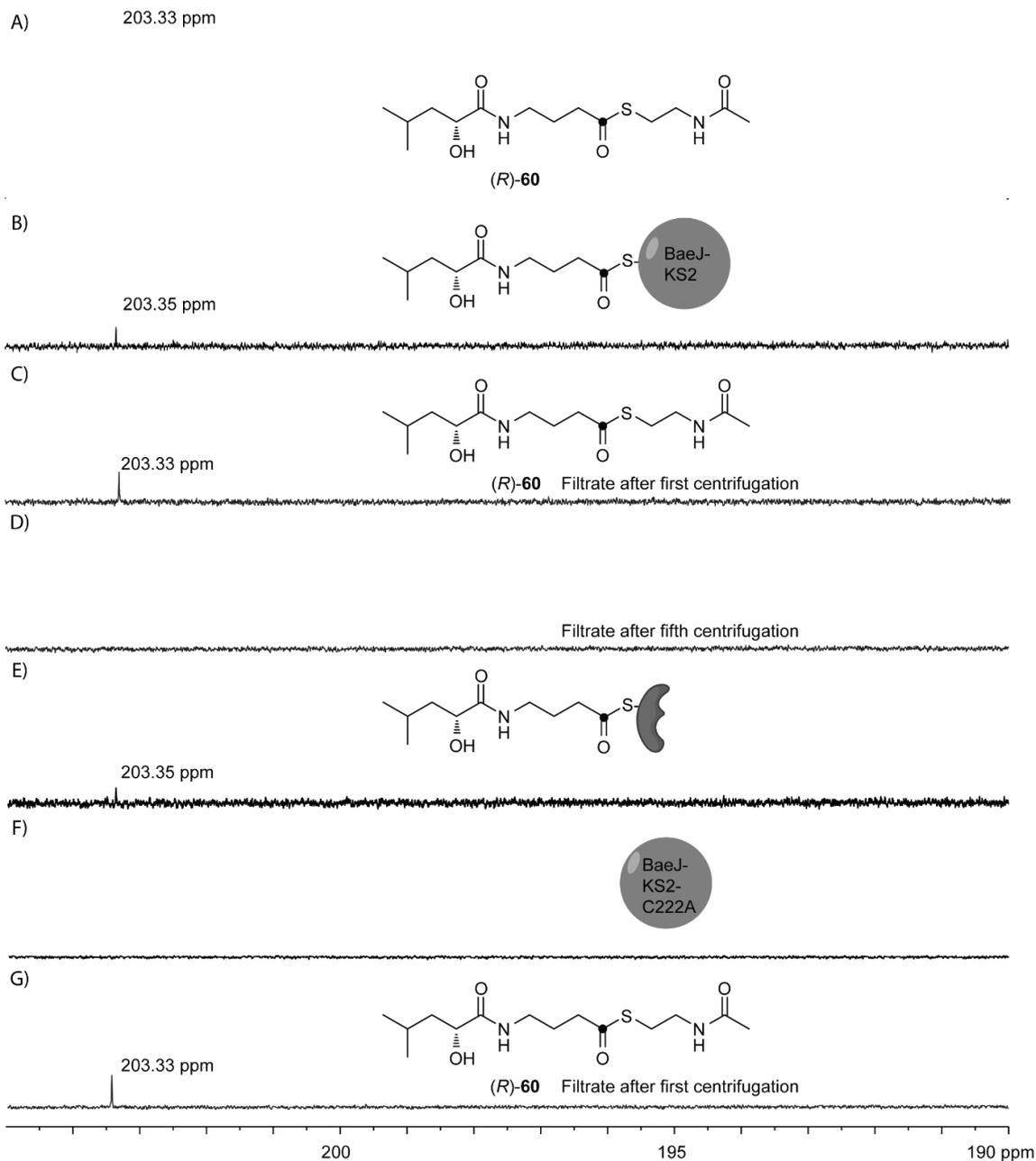
To investigate the function of BaeJ-KS2, the recombinant enzyme was expressed and purified. After incubating the enzyme with  $^{13}\text{C}$ -labeled SNAC derivatives (S)-**60** or (R)-**60** at  $25^\circ\text{C}$  for 30 minutes, unreacted substrate was removed by repeated ultrafiltration. The resulting solutions were analyzed by  $^{13}\text{C}$ -NMR spectroscopy, revealing a small chemical shift difference at  $\delta = 203.35$  ppm ( $\delta = 203.33$  ppm of free **60** in incubation buffer) (Figure 2A and 2B), indicating the presence of enzyme-bound substrate. Analysis of the filtrates confirmed the presence of free **60** after the first centrifugation (Figure 2C), but none after the final round (Figure 2D), indicating effective removal of unbound substrate.

Protein binding of the substrate surrogates (S)-**60** and (R)-**60** was confirmed by digesting BaeJ-KS2 with protease K after buffer exchange. Following digestion and ultrafiltration, the thioester carbonyl signals ( $\delta = 203.35$  ppm) of  $^{13}\text{C}$ -NMR spectroscopy in the filtrates for both substrate enantiomers were detected, indicating successful binding to the enzyme (Figure 2E). However, these results do not clarify whether the binding is covalent or non-covalent.

To further investigate whether **60** binds covalently to BaeJ-KS2, site-directed

mutagenesis was used to replace the conserved Cys222, which is implicated in substrate attachment. After expression and purification of the enzyme variant BaeJ-KS2-C222A, it was incubated with (*S*)-**60** and (*R*)-**60** using the same protocol as before. In contrast to the wild type enzyme, no thioester carbonyl signals were detected in the protein BaeJ-KS2-C222A samples (Figure 2F), while signals for free **60** ( $\delta = 203.33$  ppm) were observed in the filtrates after the first centrifugation step (Figure 2G). These results confirm that the conserved cysteine is the key residue for covalent binding of **60** to BaeJ-KS2.

In summary, a  $^{13}\text{C}$ -NMR-based method was developed to investigate the substrate specificity of KS domains using  $^{13}\text{C}$ -labeled substrate surrogates. Applying this method to BaeJ-KS2 from the bacillaene PKSs revealed that BaeJ-KS2 can tolerate an inversion in the configuration of the  $\alpha$ -hydroxyisocaproate moiety.



**Figure 2.**  $^{13}\text{C}$ -NMR spectra of *(R)*-**60** incubated with BaeJ-KS2 and BaeJ-KS2-C222A. A) Free **60** dissolved in incubation buffer; B) *(R)*-**60** bound to BaeJ-KS2 after incubation and buffer exchange (5 centrifugations); C) the filtrate obtained from the incubation of *(R)*-**60** BaeJ-KS2 containing free *(R)*-**60** (first centrifugation); D) the filtrate from the same experiment containing no *(R)*-**60** (fifth centrifugation); E) the filtrate obtained from the incubation of *(R)*-**60** with BaeJ-KS2 followed by buffer exchange and then digestion with proteinase K; F) *(R)*-**60** is not bound to BaeJ-KS2-C222A after incubation and buffer exchange (5 centrifugations); G) the filtrate obtained from the incubation of *(R)*-**60** BaeJ-KS2-C222 containing free *(R)*-**60** (first centrifugation).

## Chapter 4

# An Isotopic Probe to Follow the Stereochemical Course of Dehydratase Reactions in Polyketide and Fatty Acid Biosynthesis

Zhiyong Yin,<sup>a</sup> Elisa Liebhart,<sup>b</sup> Evi Stegmann,<sup>b</sup> Heike Brötz-Oesterhelt,<sup>b</sup> and Jeroen S. Dickschat<sup>\*a</sup>

*Org. Chem. Front.* **2022**, *9*, 2714–2720.

<sup>a</sup> Kekulé-Institute for Organic Chemistry and Biochemistry, University of Bonn, Gerhard-Domagk-Straße 1, 53121 Bonn, Germany.

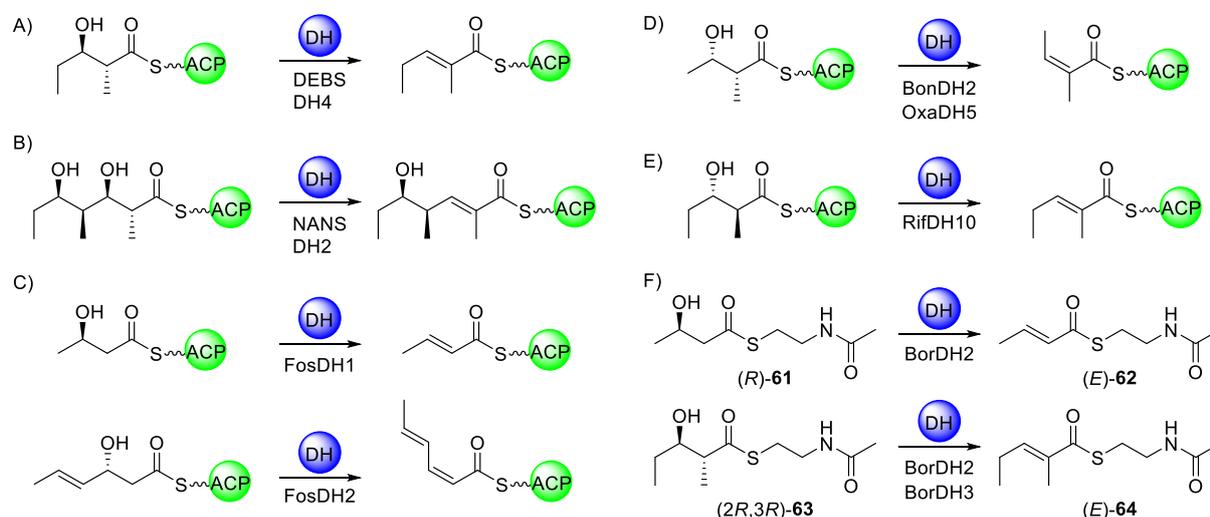
<sup>b</sup> Pharmaceutical Biology, University of Tübingen, Auf der Morgenstelle 8, 72076 Tübingen, Germany.

Reprint from *Org. Chem. Front.* with kind permission from Royal Society of Chemistry. The synthesis of compounds and enzyme reactions were performed by me. The plasmids of ShawDH1 and ShawDH2 were provided by Elisa Liebhart, Evi Stegmann and Heike Brötz-Oesterhelt.

The publication “An Isotopic Probe to Follow the Stereochemical Course of Dehydratase Reactions in Polyketide and Fatty Acid Biosynthesis” can be found in Appendix C.

## Introduction

In type I polyketide synthases (PKSs), multifunctional enzymes coordinate the assembly of polyketide chains through a series of interconnected catalytic domains. Key domains, such as acyl carrier proteins (ACPs), acyl transferases (ATs), and ketosynthases (KSs), work together to select extender units, transfer intermediates, and catalyze the fusion of the growing polyketide chain. Additional domains, including ketoreductases (KRs), dehydratases (DHs), and enoyl reductases (ERs), facilitate modifications, converting  $\beta$ -ketoacyl intermediates into various functionalized products. This modular assembly and subsequent tailoring enable the biosynthesis of diverse polyketide structures with important biological activity [165].



**Scheme 34.** Previous in vitro experiments with dehydratases from different PKSs.

Some bioinformatic studies showed that the stereospecificity of ketoreductase (KR) domains can directly influence the dehydration step catalyzed by DH domains. Type A KRs with a conserved W141 motif can generate a 3D hydroxyl intermediates and continue to install  $Z$ -configured double bond configurations intermediates catalyzed by DH domains. While type B KRs have a conserved LDD (93-95), P144, N148 motif and produce  $E$ -configured double bond configured intermediates with the function of DH domains [106]. However, there is only limited experimental work to directly track the stereochemical course of DH-catalyzed reactions. For instance, DEBS DH4 from erythromycin biosynthesis catalyzes the *syn* dehydration of a  $(2R,3R)$ -hydroxyacyl intermediate, resulting in an  $E$ -configured product (Scheme 34) [166]. Similarly, NANS DH2 from nanchangmycin [167], RifDH10 from rifamycin [168], BorDH2 and BorDH3 from borrelidin [169], and FosDH1 from fostriecin [170] biosynthesis also catalyzes *syn*

dehydrations to yield *E*-configured double bond products. In contrast, other DH domains, such as FosDH2 from fostriecin <sup>[170]</sup>, OxaDH5 from oxazolomycin <sup>[171]</sup> and BonDH2 from the bongkrecic acid <sup>[171]</sup> biosynthetic pathways, have been shown to produce *Z* double bonds by *syn* dehydration (Scheme 34).

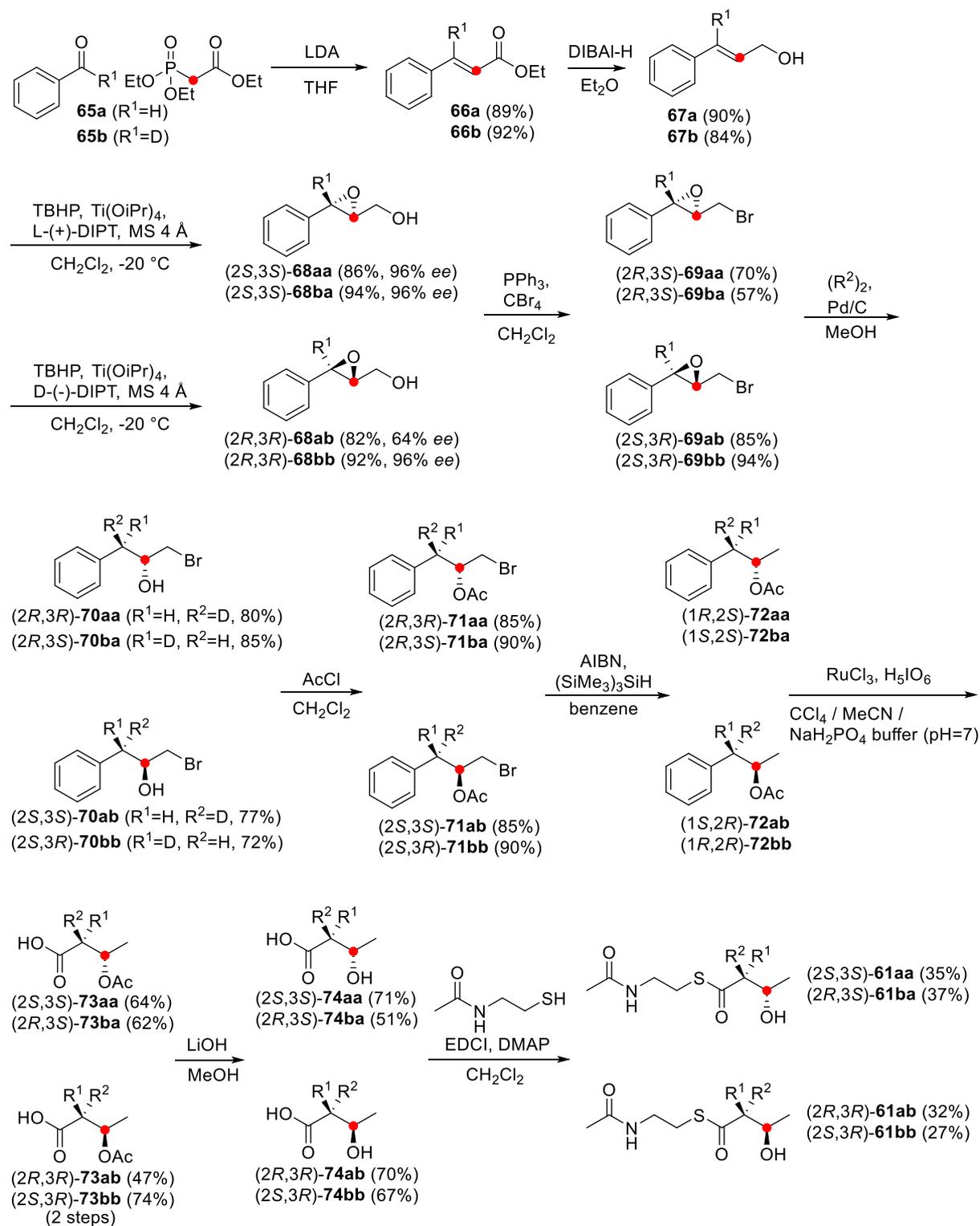
Despite these findings, detailed stereochemical investigations of the dehydration mechanisms of many DH domains are lacking. In this chapter, I developed four isotopically labeled probes that can track the stereochemical course catalyzed by DH domains using <sup>13</sup>C-NMR spectroscopy. These probes were applied to various DH domains which are from polyketide and fatty acid biosynthetic pathways.

## Summary

In polyketide and fatty acid biosynthetic pathways, DH domains can catalyze the elimination of water from  $\beta$ -hydroxyacyl intermediates and introduce *E*- or *Z*-configured double bonds in the final products. To understand the stereochemical course of these reactions, four stereoisomeric and isotopically labeled probes were designed and synthesized. These probes were presented as (2-<sup>2</sup>H,3-<sup>13</sup>C)-3-hydroxybutyric acid, and activated as *N*-acetylcysteamine (SNAC) thioesters, which can be regarded as substitutes of the natural coenzyme A thioesters. These probes were synthesized from non-labeled (**65a**) or deuterated benzaldehyde (**65b**) which can be made from methyl benzoate after DIBAL-D reduction and PCC oxidation. The Sharpless epoxidation is the key step to control the stereochemistry of **68** and will produce non-deuterated **68aa** and deuterated **68ba** with L-(+)-DIPT, or non-deuterated **68ab** and deuterated **68bb** with D-(-)-DIPT <sup>[172]</sup>. The regioselective deuteration is also important to obtain four stereoisomers of **70** with D<sub>2</sub> (non-deuterated **69**) or H<sub>2</sub> (deuterated **69**) and Pd/C <sup>[173]</sup>. These intermediates were further converted into hydroxybutyric acids **74** and the corresponding SNAC ester probes **61** (Scheme 35).

To investigate the sensitivity of these probes, BorDH2 which can convert (3*R*)-configured substrates into *E*-configured products is used as a reference. Incubations of four probes **61** with BorDH2 revealed that only 3*R* configured **61** is accepted by the enzymes, which showed a signal for the labeled carbon to a product in the olefinic region in the <sup>13</sup>C NMR spectrum. After analysis of the signals of the two enzyme crude products obtained from **61ab** and **61bb**, there is an upfield shift (-0.1 ppm) for the <sup>13</sup>C-NMR signal of synthesized standard **62b** (141.03 ppm) compared with the signal from the **61bb** as substrate (141.13 ppm for **62a**). This also confirms a

*syn* elimination for BorDH2 as reported. This indicated that these probes can also be used to distinguish between *syn* and *anti* eliminations of DH domains (Scheme 36).

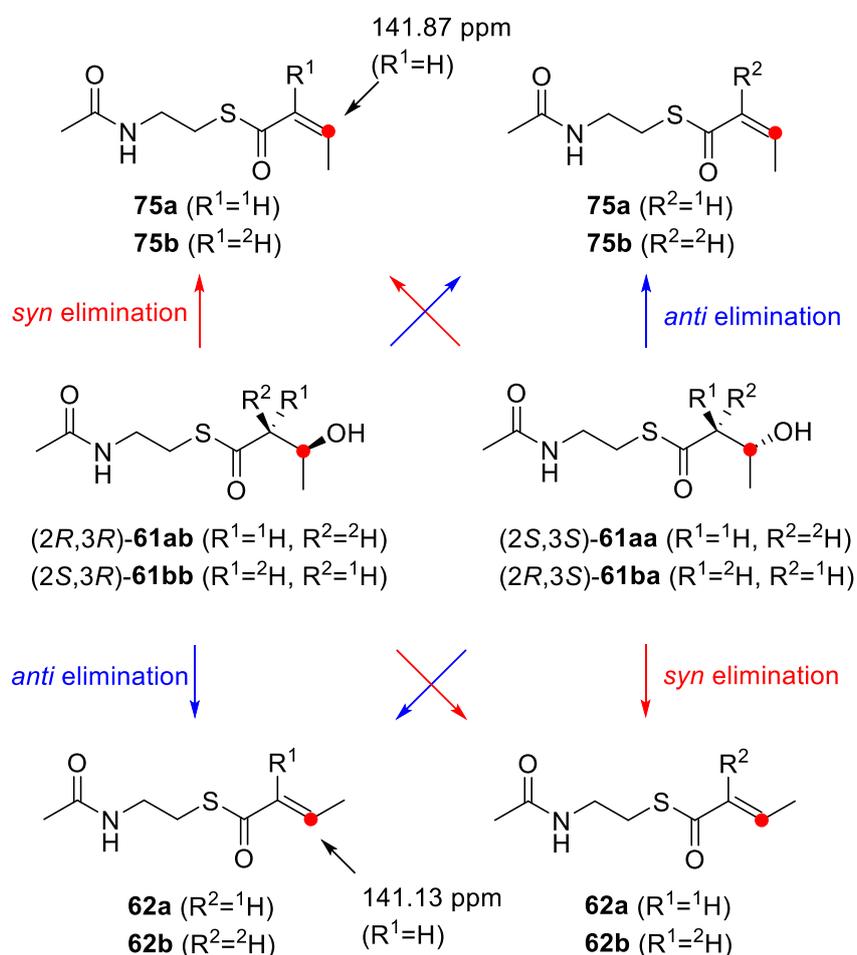


**Scheme 35.** Synthesis of  $^{13}C$ - and  $^2H$ -labeled probes **61**.

The study explored the stereochemical course of various DHs, including BorDH2, BorDH3, and BorDH5 from the borrelidin biosynthetic pathway. Results showed that

BorDH2 and BorDH5 prefer substrates with (2*R*,3*R*) configuration and catalyze with *syn* elimination to produce *E*-configured double bonds. In contrast, BorDH3 can not accept any of the synthesized stereoisomer probes.

For the DHs involved in fostriecin biosynthesis (FosDH1 and FosDH2), *syn* elimination was observed in both cases, and producing *E*-configured double bond products from (3*R*)-configured substrates. This is not in agreement with the previously reported result, which showed that FosDH2 catalyzed the formation of products with an *E*-configured double bond. The results suggest that the formation of *Z*-configured double bond products in the fostriecin biosynthetic pathway requires longer acyl chains or ACP-bound intermediates, and potentially mediated by protein-protein interactions.



**Scheme 36.** Results for enzymatic conversions of the <sup>13</sup>C and <sup>2</sup>H labeled probes **61** with dehydratases.

In addition to these polyketide synthases, the study also investigated the stereochemical course of DHs from other biosynthetic pathways, including LkcB from

lankacidin <sup>[174]</sup>, ShawDH1, ShawDH2 from ADEP1 <sup>[175]</sup> and FabZ from *Escherichia coli* fatty acid synthases <sup>[176]</sup>. All these enzymes were shown to perform *syn* eliminations with 3*R*-configured probes.

In summary, this study highlighted the value of isotopically labeled probes in elucidating the stereochemical pathways of dehydratase reactions in polyketide and fatty acid biosynthesis. The results from various DHs also suggest that protein-protein interactions may be more important in *Z*-configured double bond products formation in polyketide biosynthesis.

## Chapter 5

# The Substrate Scope of Dehydratases in Antibiotic Biosynthesis and their Application in Kinetic Resolutions

Zhiyong Yin,<sup>a</sup> Daniel Bär,<sup>b</sup> Bertolt Gust,<sup>b</sup> and Jeroen S. Dickschat\*<sup>a</sup>  
*Org. Biomol. Chem.* **2022**, *20*, 9103–9107.

<sup>a</sup> Kekulé-Institute for Organic Chemistry and Biochemistry, University of Bonn, Gerhard-Domagk-Straße 1, 53121 Bonn, Germany.

<sup>b</sup> Pharmaceutical Biology, University of Tübingen, Auf der Morgenstelle 8, 72076 Tübingen, Germany.

Reprint from *Org. Biomol. Chem.* with kind permission from Royal Society of Chemistry.

The synthesis of compounds and enzyme reactions were performed by me. The plasmids of Cpz2 was provided by Daniel Bär and Bertolt Gust.

The publication “The Substrate Scope of Dehydratases in Antibiotic Biosynthesis and their Application in Kinetic Resolutions” can be found in Appendix D.

## Introduction

Multifunctional enzymes in type I polyketide synthases (PKSs) integrate several catalytic activities within a single protein, facilitating efficient biosynthesis of polyketide chains. In these enzymes, domains such as acyl carrier proteins (ACPs), acyltransferases (ATs), and ketosynthases (KSs) are essential for chain elongation and substrate loading [28]. The dehydratase (DH) domain, part of this multifunctional architecture, plays a crucial role in introducing double bonds by eliminating a water molecule from hydroxyl intermediates [105]. With advancement in bioinformatic studies in polyketide biosynthesis, these domain arrangement can be used for predicting some of the polyketide structures. Based on our research on DH domains, we also found that the protein-protein interactions are very important for polyketide stereochemical course [177]. For example, FosDH2, a dehydratase in the fostriecin biosynthetic pathway, can catalyze the dehydration to install a *Z*-configured double bond in its product from the natural ACP-bound substrate [170], but will introduce an *E*-configured double bond in the product obtained from a substrate surrogate (*N*-acetylcysteamine thioester, SNAC ester) [177].

Although there are many reports reported about the substrate scope of different polyketide synthase domains [72,178-182], there are still very few cases about the substrate preference of DH domains [181-184]. This chapter focuses on the substrate scope and stereoselectivity of various DH domains involved in antibiotic and fatty acid biosynthetic pathways.

## Summary

Dehydratase (DH domain) in polyketide and fatty acid synthases can catalyze the elimination of water from (3L or 3D)  $\beta$ -hydroxyacyl intermediates which is introduced by the KR domain (type A or type B) and introduce an *E* or *Z*-configured double bond in the final product [105]. In order to investigate the substrate scope of different DH domains, nine stereoisomeric SNAC ester substrates were synthesized. SNAC esters **61**, **76** and **83** were prepared by using commercially available corresponding carboxylic acids. SNAC esters **77–81** were synthesized stereoselectively by utilizing Evans' oxazolidinones in aldol reactions as a key step [185]. SNAC ester **82** was obtained as a racemate. A panel of dehydratases, including BorDH2 and BorDH5 from borrelidin (**84**) biosynthesis [186], FosDH1 and FosDH2 from fostriecin (**85**) biosynthesis [170], RifDH10 from rifamycin (**86**) biosynthesis [187], ShawDH1 and

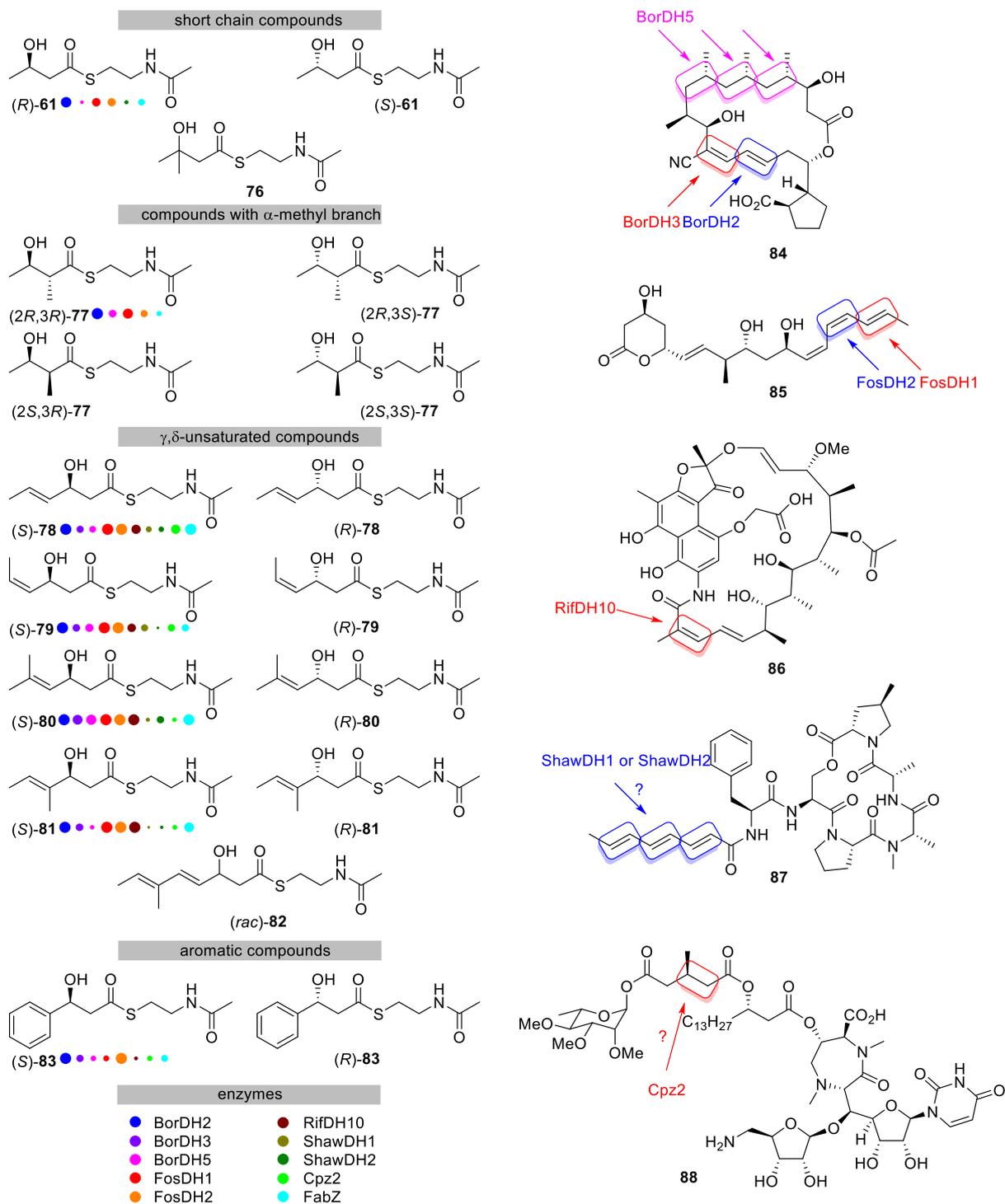
ShawDH2 potentially involved in acyldepsipeptide ADEP1 (**87**) biosynthesis [175], Cpz2 from caprazamycin (**88**) [188], and FabZ from fatty acid biosynthetic pathway [176], were tested with all synthetic 3-hydroxy SNAC esters (Figure 3). A color-coded system was used to indicate enzyme-substrate compatibility, with the size of the dot representing the conversion rate (Figure 3). None of these enzymes accepted substrates with 3L-configured alcohol or the compound with a tertiary alcohol group. The products were measured by <sup>1</sup>H-NMR, comparing spectra with starting materials and synthetic standards.

The screening results revealed that BorDH2, FosDH1, and FabZ were highly effective in converting various 3D-hydroxy SNAC esters, with BorDH2 showing the best performance. Other enzymes like BorDH5, ShawDH2, and Cpz2 showed lower activity or poor substrate acceptance. Overall, BorDH2 was chosen for further study due to its good activity and wide substrate range.

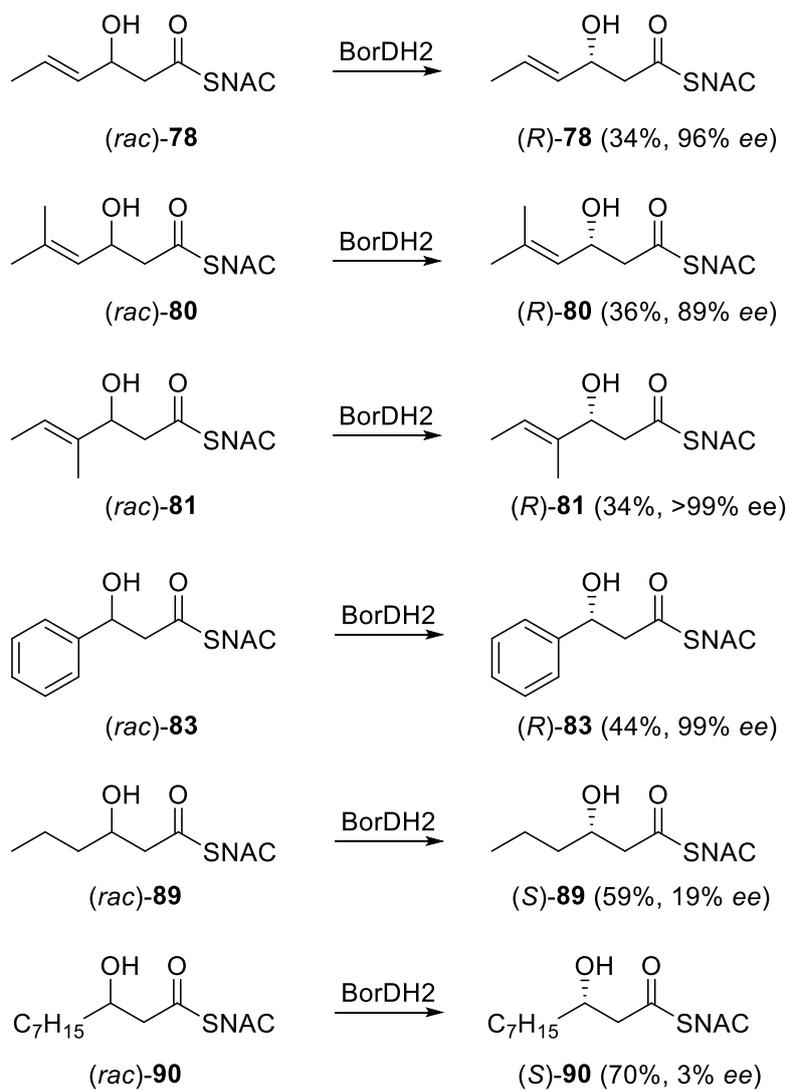
The study assessed the effectivity of BorDH2 on various substrates, revealing full conversion (100%) in multiple trials for the  $\gamma,\delta$ -unsaturated substrates (S)-**78**, (S)-**79**, (S)-**80**, and (S)-**81**, as well as the aromatic compound (S)-**83**. It also performed well with substrates (R)-**61** and (2R,3R)-**77**, showing conversions of 92±4% and 99±1%, respectively. However, the conversion rate for (rac)-**82** was only 33±5%.

Although BorDH2 demonstrated significant potential with enantiomerically pure 3D alcohols, the synthesis of these materials is complex and results in loss of stereochemical information, limiting their practical applications. In this study, BorDH2 was used for enzymatic kinetic resolutions. It was tested with racemic 3-hydroxy-SNAC esters (Scheme 37), yielding high enantiomeric purity for  $\gamma,\delta$ -unsaturated esters (rac)-**78**, (rac)-**80**, (rac)-**81**, and the aromatic compound (rac)-**83**, with enantiomeric excesses ranging from 89% to nearly 100%. However, the reaction with the saturated 3-hydroxyhexanoyl-SNAC (rac)-**89** was less efficient, resulting in a high recovery of the starting material with only 19% ee. Chain elongation substrate 3-hydroxydecanoyl-SNAC (rac)-**90** displayed even poorer results, showing the limitation of this method.

This study demonstrated that BorDH2 exhibits high substrate tolerance and efficiency for kinetic resolutions of 3-hydroxy-SNAC esters, showing particular effectiveness with  $\gamma,\delta$ -unsaturated and aromatic substrates, while less suitable for saturated and longer-chain substrates.



**Figure 3.** Synthetic 3-hydroxy SNAC esters used as DH substrates. The colour code indicates which enzyme can convert which substrate. The size of the coloured dot refers to the conversion rate.



**Scheme 37.** Kinetic resolutions with BorDH2. Dehydration products are not shown.



## Chapter 6

# Enantioselective Synthesis of All Stereoisomers of Geosmin and of Biosynthetically Related Natural Products

Zhiyong Yin,<sup>a</sup> Michael Maczka,<sup>b</sup> Gregor Schnakenburg,<sup>c</sup> Stefan Schulz<sup>b</sup> and Jeroen S. Dickschat<sup>\*a</sup>

*Org. Biomol. Chem.* **2024**, *22*, 5748–5758.

<sup>a</sup> Kekulé-Institute for Organic Chemistry and Biochemistry, University of Bonn, Gerhard-Domagk-Straße 1, 53121 Bonn, Germany.

<sup>b</sup> Institute for Organic Chemistry, TU Braunschweig, Hagenring 30, 38106 Braunschweig, Germany.

<sup>c</sup> Institute for Inorganic Chemistry, University of Bonn, Gerhard-Domagk-Straße 1, 53121 Bonn, Germany.

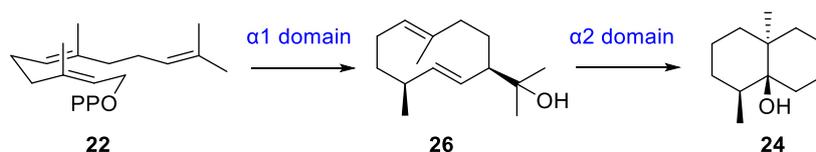
Reprint from *Org. Biomol. Chem.* with kind permission from Royal Society of Chemistry.

The synthesis of compounds and enzyme reactions were performed by me. The investigation of headspace extracts was provided by Michael Maczka and Stefan Schulz. The X-ray structures were measured and elucidated by Gregor Schnakenburg.

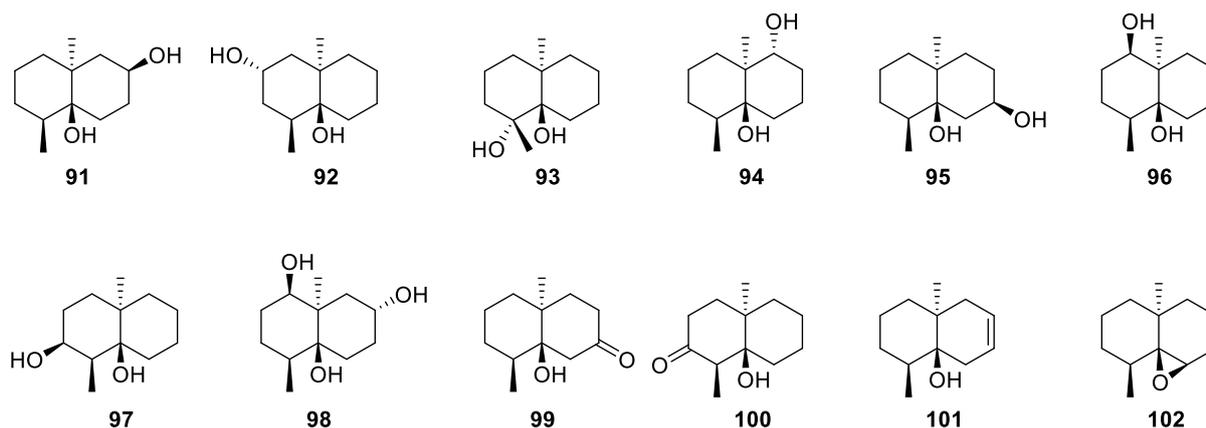
The publication “Enantioselective synthesis of all stereoisomers of geosmin and of biosynthetically related natural products” can be found in Appendix E.

## Introduction

Geosmin synthase is a bifunctional enzyme involved in the biosynthesis of geosmin (**24**), a compound that contributes to the earthy smell of soil [189,190]. Geosmin synthase has two domains: the N-terminal  $\alpha 1$  domain catalyzes the conversion of farnesyl diphosphate (**22**) into (1(10)*E*,5*E*)-germacradien-11-ol (**26**), and the the C-terminal  $\alpha 2$  domain converts (1(10)*E*,5*E*)-germacradien-11-ol (**26**) to geosmin (**24**) (Scheme 38) [125-127,191,192].



**Scheme 38.** Geosmin synthase catalyzes the conversion of FPP (**6**) into (1(10)*E*,5*E*)-germacradien-11-ol (**7**) and geosmin (**5**).

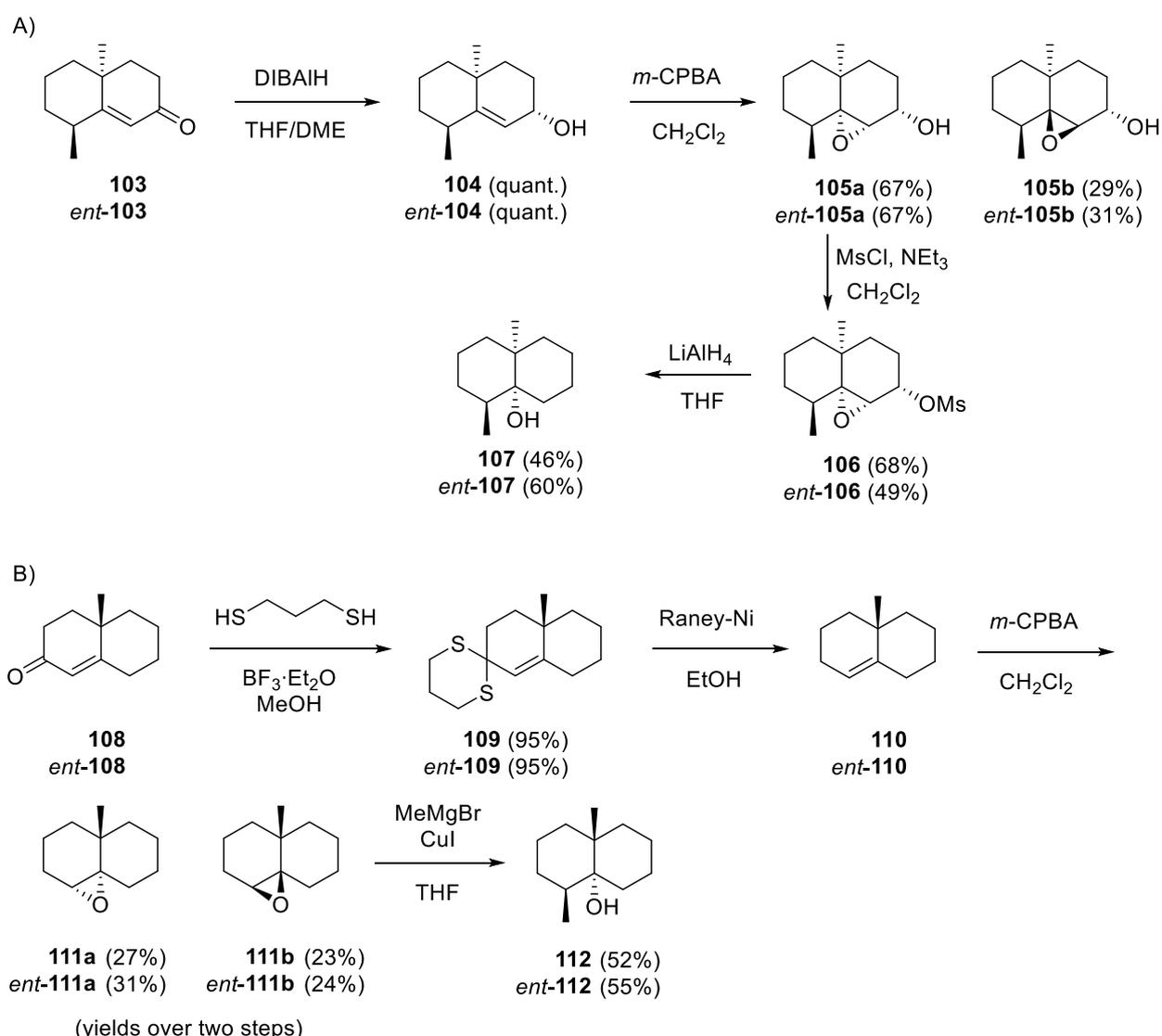


**Scheme 39.** Known naturally occurring derivatives of geosmin.

Several oxidized geosmin derivatives have been identified from various sources. For example, hydroxylated compounds like cybullol (**91**) from *Cyathus bulleri* [193] and different alcohols (**92-98**) have been discovered [194-196]. Ketones (**99** and **100**) are known biodegradation products of geosmin from *Pseudomonas* sp. [197]. Additional derivatives, such as dehydrogeosmin (**101**) [198] and epoxide **102** [199], have been isolated from different species (Scheme 39). Compounds (*rac*)-**91**, (*rac*)-**95**, **96**, and **101**, have been synthesized [199-202], along with efforts focused on the enantioselective synthesis of geosmin (–)-**24** [203-205] and its enantiomer (+)-**24** [202,203,206]. In this chapter, the enantioselective synthesis of all stereoisomers of geosmin (**24**), the geosmin derivative **95**, and the identification of a previously unrecognized side product from geosmin biosynthetic pathway are presented.

## Summary

To synthesize 5-*epi*-geosmin (**107**, Scheme 40A), octalone **103**, a key intermediate in the synthesis of (–)-**103**, was stereoselectively prepared via Robinson annulation using a chiral auxiliary [203]. After repeated recrystallization, enantiomerically pure octalone **103** was obtained and converted into 5-*epi*-geosmin (**107**) through a series of diastereoselective reactions, including reduction to **104**, epoxidation to **105**, mesylation, and subsequent treatment with LiAlH<sub>4</sub> (Scheme 40A) [204]. The study also presents the enantioselective synthesis of *ent*-**107** and provides X-ray structures for **103**, *ent*-**103**, and *ent*-**104**.

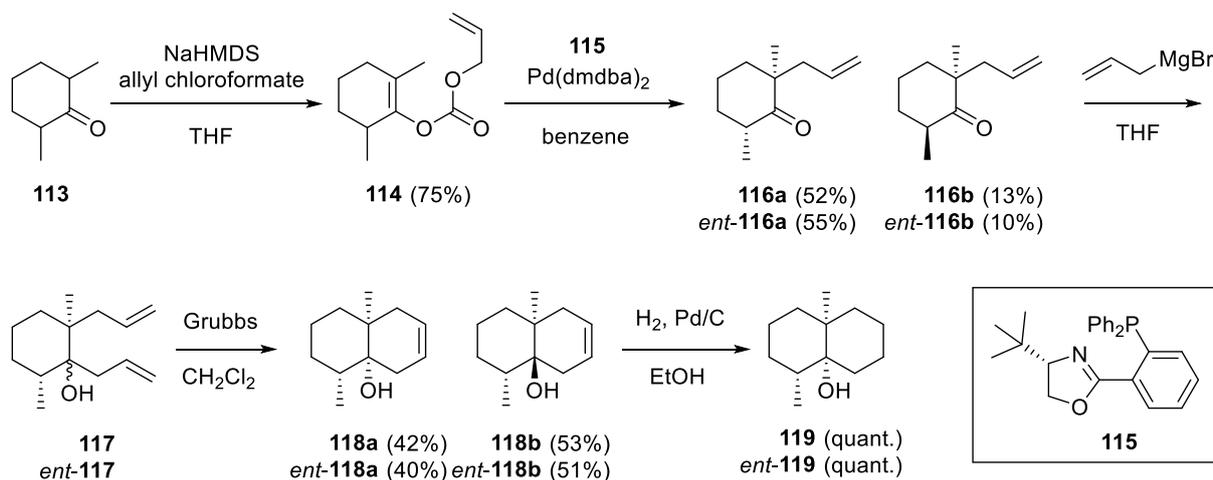


**Scheme 40.** A) Synthesis of 5-*epi*-geosmin (**107**) and *ent*-5-*epi*-geosmin (*ent*-**107**); B) 10-*diepi*-geosmin (**112**) and *ent*-5,10-*diepi*-geosmin (*ent*-**112**).

To synthesize 5,10-*diepi*-geosmin (**112**, Scheme 40B), enantiomerically pure octalone **108**, obtained through the same method with octalone **103** [203], was converted into

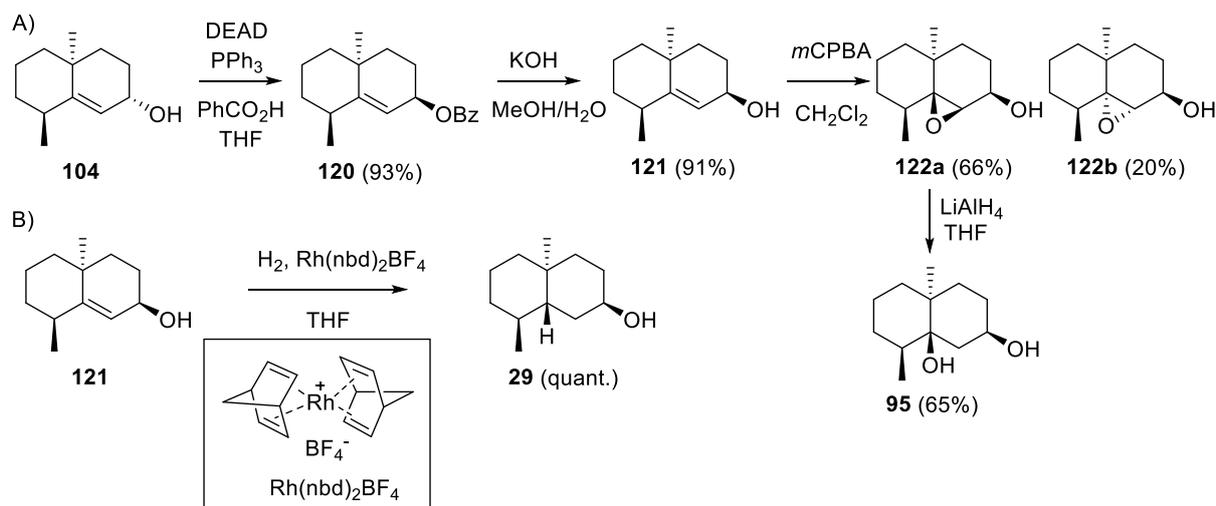
5,10-*diepi*-geosmin (**112**) via a series of steps. First, octalone **108** was converted into dithiane **109** using 1,3-propanedithiol and  $\text{BF}_3 \cdot \text{Et}_2\text{O}$ , followed by defunctionalisation to **110** with Raney nickel. Epoxidation of **110** with *m*-CPBA was performed, followed by the CuI-catalyzed addition of  $\text{MeMgBr}$  to get 5,10-*diepi*-geosmin (**112**). The same route was used to synthesize *ent*-5,10-*diepi*-geosmin (*ent*-**112**) from *ent*-**108**.

The synthesis of 4,5-*diepi*-geosmin (**119**) (Scheme 41) started with 2,6-dimethylcyclohexanone (**113**), which was converted into carbonate ester **114**. Using the Stoltz variation of the Tsuji-Trost allylation<sup>[207]</sup>, **116a** and **116b** were obtained (dr  $\approx$  4:1). The intermediate **116a** was treated with allylmagnesium bromide to produce **117**, followed by cyclization with Grubbs I catalyst to give dehydrogeosmin stereoisomers **118a** and **118b**. Hydrogenation of **118a** yielded the final product **119**. The same route was used to synthesize *ent*-**119**.



**Scheme 41.** Synthesis of 4,5-*diepi*-geosmin (**119**) and *ent*-4,5-*diepi*-geosmin (*ent*-**119**).

The synthesis of the hydroxylated geosmin **95** started from intermediate **104** (Scheme 42A). The Mitsunobu inversion of compound **104** yielded the benzoate ester **120**<sup>[208]</sup>, which was subsequently saponified to produce alcohol **121**. Epoxidation of **121** yielded **122a** and **122b** (dr  $\approx$  3:1), which can be separated chromatographically. The major stereoisomer **122a** was converted to **95** via  $\text{LiAlH}_4$ -mediated epoxide opening. The synthetic **95** exhibited NMR data and optical rotation consistent with those of the natural product, confirming its absolute configuration as (4*S*,5*S*,7*R*,10*S*)-**95**, which was further verified by X-ray analysis.



**Scheme 42.** Synthesis of **95** and the newly identified product **29** of geosmin synthase.

To further investigate bacterial volatiles, an unknown compound was identified in the headspace extracts of *Streptomyces reticuli* [209]. This compound displayed a mass spectrum similar to geosmin and was proposed to be a geosmin isomer. To identify the structure, the geosmin synthase gene from *S. ambofaciens* was cloned and expressed in *E. coli*, revealing the unknown compound as decalol **29**, named isogeosmin. An enantioselective synthesis of **29** confirmed its structure and absolute configuration as (4*S*,5*R*,7*R*,10*S*)-**29** (Scheme 42B), with matching optical rotations between the natural and the synthetic compound.

In summary, this study achieved the enantioselective synthesis of all eight stereoisomers of geosmin, along with the identification and synthesis of isogeosmin (**29**), a side product of geosmin synthase, confirming its structure and absolute configuration. Although ecological roles of geosmin are well documented, the biological functions of its side products, including isogeosmin, remain to be explored.



## Chapter 7

### Summary and outlook

The research presented in this doctoral dissertation mainly focuses on the investigation of multifunctional enzymes in natural product biosynthesis. This work has resulted in five publications, including a wide range of topics within protein-protein and protein-substrate interactions involved in natural product biosynthesis. This study provides valuable insights into tracking biochemical reactions and substrate specificity in multifunctional enzymes within natural product biosynthetic pathways.

The initial study, titled “A Clickable Coenzyme A Derived Probe for Investigating Phosphopantetheinyl Transferase (PPTase) Activity in Natural Product Biosynthesis,” describes a method for monitoring the activation of carrier proteins (CPs) by PPTases. This is achieved by using a modified HSCoA with an alkyne group, which is used for Cu-catalyzed alkyne-azide cycloaddition (CuAAC, “click chemistry”) for the attachment of a fluorescent tag. This method is outlined in Chapter 2 <sup>[210]</sup>. The subsequent study, “Substrate Specificity of a Ketosynthase Domain Involved in Bacillaene Biosynthesis” investigates the substrate specificity of BaeJ-KS2, a ketosynthase domain within the bacillaene biosynthetic pathway. In this study, a new method using <sup>13</sup>C-labeled substrate surrogates, combined with <sup>13</sup>C-NMR analysis, was developed to explore the specificity of ketosynthase domains, as detailed in Chapter 3 <sup>[211]</sup>. A further study, “An Isotopic Probe to Follow the Stereochemical Course of Dehydratase Reactions in Polyketide and Fatty Acid Biosynthesis” details the development of <sup>13</sup>C and <sup>2</sup>H isotopically labeled probes for monitoring the configuration of double bonds produced by dehydratase (DH) domains in polyketide and fatty acid synthases, as presented in Chapter 4 <sup>[177]</sup>. This research raises further questions about the substrate scope of dehydratases. This research continues with, “The Substrate Scope of Dehydratases in Antibiotic Biosynthesis and their Application in Kinetic Resolutions,” explores nine dehydratases involved in secondary metabolites biosynthesis, along with FabZ from fatty acid biosynthesis. A panel of *N*-acetylcysteamine (SNAC) thioesters is used for investigating the substrate scope of these dehydratases in this study. The most efficient enzyme, BorDH2, is applied in kinetic resolutions, as detailed in Chapter 5 <sup>[212]</sup>. Lastly, the final publication, “Enantioselective Synthesis of All Stereoisomers of Geosmin and of Biosynthetically

Related Natural Products,” explores geosmin synthase, a bifunctional terpene synthase. This study not only leads to the discovery of a new natural compound: isogeosmin (**29**), but also discuss the stereoselective synthesis of all stereoisomers of geosmin as discussed in Chapter 6 <sup>[213]</sup>.

A significant part of this research focuses on polyketide synthases, which are multifunctional enzymes and play crucial role in the biosynthesis of natural products. For instance, PPTases can convert CPs into their active form, *holo*-CPs. These *holo*-CPs can be used for carrying elongating chain and extender units within polyketide and non-ribosomal peptide biosynthetic pathways. It is important to monitor the PPTase-catalyzed phosphopantetheinylation of CPs for understanding the biosynthesis of polyketides and non-ribosomal peptides. Although HPLC-MS analysis can be used to track this process <sup>[152-154]</sup>, it requires expensive equipment. In most cases, Sfp, a PPTase from *Bacillus subtilis* known for its remarkable substrate promiscuity and broad applicability <sup>[57]</sup>, is used in engineering to activate CPs. However, to enhance the versatility and efficiency of biocatalytic processes, there is a pressing need to identify more PPTases with characteristics similar to Sfp. This dissertation addressed these problems by synthesizing an HSCoA derivative with an alkyne functional group substituting for the thiol group. These derivatives were used for enzyme reactions, and after the modified alkyne group was attached to CPs, a fluorescence tag could be added through click chemistry. The results of whether PPTases activate CPs can be visualized by performing SDS-PAGE, where activated CPs with fluorescence tags can be observed under UV light. The method provides a simple way to monitor the process of PPTases activating CPs without the need of expensive machines or complex procedures. Using this method, several PPTases that are capable of activating various CPs and have the potential to be used in the engineering of polyketides and non-ribosomal peptides production have been identified.

This study investigated substrate specificity of a ketosynthase (KS) domain in a multifunctional enzyme, bacillaene synthase, is also involved in developing a new method. This approach offers a direct way to determine whether substrates are accepted by enzymes. By utilizing <sup>13</sup>C-labeled substrates in enzyme reactions, it allows the monitoring of products through <sup>13</sup>C NMR spectroscopy. This method leads to the exploration of BaeJKS2 substrate tolerance. The peak shifts of the enzyme

reaction samples in  $^{13}\text{C}$  NMR spectra indicates whether enzymes accept substrates or not.

During the exploration of PPTases-CPs and KS domains, these probes were obtained by combining enzymatic and chemical synthesis, which also serves as strong evidence that enzyme research in natural product biosynthesis is very useful.

Additionally, a considerable part of this dissertation explores the DH domains involved in fatty acid and polyketide biosynthesis. For the investigation of DH domains, the first publication of this part describes the design and synthesis of four stereoisomeric and isotopically labeled probes. After the incubation of labeled substrates with different DH domains, the products were analyzed by  $^{13}\text{C}$  NMR spectroscopy. The differences in  $^{13}\text{C}$  NMR chemical shift make it easy to track the double bond configuration of products. Additionally, these chemical shift differences also indicate which type of elimination occurs during dehydrations, whether *syn*-elimination or *anti*-elimination. Subsequent analysis revealed that the double bond configuration of some enzyme products, which were expected to be *Z*-configured based on the final products, did not match the experimental *Z* configuration. For example, FosDH2 in fostriecin biosynthesis, catalyzed the formation of *Z*-configured double bond products from (3*L*)-configured substrates. However, it is observed FosDH2 catalyzed the formation of *E*-configured double bond products from (3*D*)-configured substrates [170,177]. Based on these results, we proposed that protein-protein interactions is very important within multifunctional enzymes in natural product biosynthesis.

In addition to developing probes for DH domains, the substrate scope of various DH domains was also investigated in Chapter 5. Several sets of substrates were synthesized and tested across different DH domains in polyketide and fatty acid synthases. The results showed that the enzyme BorDH2 exhibits significant activity and selectivity towards various substrates. Furthermore, the application of BorDH2 was explored by incubating it with racemic substrates. The enzyme reaction products were then isolated and analyzed for their enantiomeric purity, providing an effective method for obtaining enantiomerically pure alcohols. These studies on DH domains emphasize the identification of double bond configurations in dehydratase products and the use of dehydratases for producing enantiomerically pure alcohols. These findings not only advance the understanding of DH domains but also highlight

potential applications in enantioselective synthesis.

The challenges of using isotope labeling techniques, e.g. in isotopically labeled probes for investigating KS and DH domains, lie in the multi-step synthesis of labeled compounds, which involves handling highly volatile intermediates and requires expensive substrates.

In addition to exploring polyketide synthases, this dissertation also investigates multifunctional terpene synthases. During the research on geosmin biosynthesis, the bifunctional geosmin synthase which features a unique structure with two domains that serve different functions was studied. This enzyme catalyzes the production of geosmin and isogeosmin (**29**). Isogeosmin was indentified for the first time during geosmin biosynthesis research as explained in Chapter 6. The structure of isogeosmin (**29**) was elucidated using NMR spectroscopy and the absolute configuration was determined through enantioselective total synthesis, comparing the optical rotations the synthetic and the isolated compound. Geosmin is a very interesting compound with a strong earthy odour and widespread occurrence in various organisms. As a result, all the stereoisomers of geosmin were also synthesized. These compounds were synthesized through a multi-step enantioselective process, which has significantly enhanced our understanding of geosmin stereochemistry. When comparing the products of geosmin synthase, derived from FPP, with all the stereoisomers of geosmin, it was found that only geosmin was present in the enzyme reaction products. This finding also demonstrates the specificity of the enzyme in controlling stereochemistry during the reaction. This research significantly improves the understanding the production of valuable terpenes and highlights the potential for engineering new terpene synthase. Moreover, the insights gained contribute to the advancement of natural product discovery.

The implications of these findings in this dissertation are extensive. The probe developed for monitoring PPTase-CP interactions highlights the complexity of polyketide and non-ribosomal peptide synthases and their applications in engineering synthetic biology. The ability of this probe to investigate PPTase activities toward CPs opens a new route for discovering potential PPTases in natural product production. Future research may focus on investigating the protein-protein interactions of PPTase-CP through co-crystallization of protein complexes and engineering complete

polyketide and non-ribosomal peptide pathways. Investigating KS domains by using  $^{13}\text{C}$  labeled substrates could provide a novel approach for exploring the substrate specificity of KS domains in polyketide biosynthesis. Additionally, the research into KS domains offers fundamental knowledge that could benefit the C–C bond formation and engineering of polyketides production. Moreover,  $^{13}\text{C}$  and deuterated labeled probes can easily follow the stereochemical course of DH domains, revealing the configuration of double bonds produced in polyketide biosynthesis. This approach not only provides a detailed understanding of the enzymatic reaction products but also offers broader insights into the mechanism of DH domains, paving the way for future studies in DH domains. The research about DH domains substrate scope can provide essential knowledge for the development of biocatalysts and opens new directions in multifunctional enzyme applications. The discovery of multifunctional terpene synthases indicates the potential existence of new enzymes with unique functions, as well as the possibility of identifying new terpenes.

In the long term, this research improves our understanding of multifunctional enzymes involved in natural product biosynthesis, which is important for enzyme science. Overall, this cumulative thesis enhances our understanding of multifunctional enzymes, emphasizing complex biochemical pathways and their wider applicability. These findings create opportunities for future research by addressing unsolved questions and exploring new areas in the study of multifunctional enzymes.



## References

1. G R. Stark, *Trends Biochem. Sci.* **1977**, *2*, 64–66.
2. X. Y. Cheng, W. J. Huang, S. C. Hu, H. L. Zhang, H. Wang, J. X. Zhang, H. H. Lin, Y. Z. Chen, Q. Zou, Z. L. Ji, *PloS One* **2012**, *7*, e38979.
3. H. Ritter, G. E. Schulz, *Plant Cell* **2004**, *16*, 3426–3436.
4. T. Vogt, *Mol. Plant* **2010**, *3*, 2–20.
5. S. Raboni, S. Bettati, A. Mozzarelli, *Cell. Mol. Life Sci.* **2009**, *66*, 2391–2403.
6. S. W. White, J. Zheng, Y. M. Zhang, C. O. Rock, *Annu. Rev. Biochem.* **2005**, *74*, 791–831.
7. S. Jenni, M. Leibundgut, D. Boehringer, C. Frick, B. Mikolásek, N. Ban, *Science* **2007**, *316*, 254–261.
8. M. Leibundgut, S. Jenni, C. Frick, N. Ban, *Science* **2007**, *316*, 288–290.
9. T. Maier, M. Leibundgut, N. Ban, *Science* **2008**, *321*, 1315–1322.
10. C. S Heil, S. S Wehrheim, K. S. Paithankar, M. Grininger, *ChemBioChem* **2019**, *20*, 2298–2321.
11. K. Magnuson, W. Oh, T. J. Larson, J. E. Cronan, *FEBS Lett.* **1992**, *299*, 262–266.
12. J. T. Tsay, W. Oh, T. J. Larson, S. Jackowski, C. O. Rock, *J. Biol. Chem.* **1992**, *267*, 6807–6814.
13. P. Edwards, J. S. Nelsen, J. G. Metz, K. Dehesh, *FEBS Lett.* **1997**, *402*, 62–66.
14. A. C. Price, Y. M. Zhang, C. O. Rock, S. W. White, *Biochemistry* **2001**, *40*, 12772–12781.
15. R. J. Heath, C. O. Rock, *J. Biol. Chem.* **1996**, *271*, 27795–27801.
16. R. J. Heath, C. O. Rock, *J. Biol. Chem.* **1995**, *270*, 26538–26542.
17. H. Cho, J. E. Cronan Jr, *J. Biol. Chem.* **1993**, *268*, 9238–9245.
18. J. Naggert, M. L. Narasimhan, L. DeVeaux, H. Cho, Z. I. Randhawa, J. E. Cronan Jr, B. N. Green, S. Smith, *J. Biol. Chem.* **1991**, *266*, 11044–11050.
19. J. N. Collie, W. S. Myers, *J. Chem. Soc.* **1893**, *63*, 122–128.
20. J. Staunton, K. J. Weissman, *Nat. Prod. Rep.* **2001**, *18*, 380–416.
21. A. J. Birch, P. A. Massy-Westropp, C. J. Moye, *Aus. J. Chem.* **1955**, *8*, 539–544.
22. D. E. Cane, P. C. Prabhakaran, W. Tan, W. R. Ott, *Tetrahedron Lett.* **1991**, *32*, 5457–5460.
23. F. H. Malpartida, D. A. Hopwood, *Nature* **1984**, *309*, 462–464.
24. J. Cortés, S. F. Haydock, G. A. Roberts, D. J. Bevitt, P. F. Leadlay, *Nature* **1990**, *348*, 176–178.

25. J. S. Tuan, J. M. Weber, M. J. Staver, J. O. Leung, S. Donadio, L. Katz, *Gene* **1990**, *90*, 21–29.
26. S. Donadio, M. J. Staver, J. B. McAlpine, S. J. Swanson, L. Katz, *Science* **1991**, *252*, 675–679.
27. D. J. Bevitt, J. Cortés, S. F. Haydock, P. F. Leadlay, *Eur. J. Biochem.* **1992**, *204*, 39–49.
28. C. Hertweck, *Angew. Chem. Int. Ed.* **2009**, *48*, 4688–4716.
29. J. Piel, *Nat. Prod. Rep.* **2010**, *27*, 996–1047.
30. E. J. Helfrich, J. Piel, *Nat. Prod. Rep.* **2016**, *33*, 231–316.
31. P. S. Patel, S. Huang, S. Fisher, D. Pirnik, C. Aklonis, L. Dean, E. Meyers, P. Fernandes, F. Mayerl, *J. Antibiot.* **1995**, *48*, 997–1003.
32. X. H. Chen, J. Vater, J. Piel, P. Franke, R. Scholz, K. Schneider, A. Koumoutsi, G. Hitzeroth, N. Grammel, A. W. Strittmatter, G. Gottschalk, R. D. Sussmuth, R. Borriss, *J. Bacteriol.* **2006**, *188*, 4024–4036.
33. X. H. Chen, A. Koumoutsi, R. Scholz, A. Eisenreich, K. Schneider, I. Heinemeyer, B. Morgenstern, B. Voss, W. R. Hess, O. Reva, H. Junge, B. Voigt, P. R. Jungblut, J. Vater, R. Sussmuth, H. Liesegang, A. Strittmatter, G. Gottschalk, R. Borriss, *Nat. Biotechnol.* **2007**, *25*, 1007–1014.
34. J. Moldenhauer, X. H. Chen, R. Borriss, J. Piel, *Angew. Chem. Int. Ed.* **2007**, *46*, 8195–8197.
35. S. M. Ma, J. W.-H. Li, J. W. Choi, H. Zhou, K. K. M. Lee, V. A. Moorthie, X. Xie, J. T. Kealey, N. A. Da Silva, J. C. Vederas, Y. Tang, *Science* **2009**, *326*, 589–592.
36. T. R. Fallon, V. V. Shende, I. H. Wierzbicki, R. P. Auber, D. J. Gonzalez, J. H. Wisecaver, B. S. Moore, *Science* **2024**, *385*, 671–678.
37. R. Pieper, C. Kao, C. Khosla, G. Luo, D. E. Cane, *Chem. Soc. Rev.* **1996**, *25*, 297–302.
38. A. Nivina, K. P. Yuet, J. Hsu, C. Khosla, *Chem. Rev.* **2019**, *119*, 12524–12547.
39. G. J. Dodge, F. P. Maloney, J. L. Smith, *Nat. Prod. Rep.* **2018**, *35*, 1082–1096.
40. T. Robbins, Y. C. Liu, D. E. Cane, C. Khosla, *Curr. Opin. Struct. Biol.* **2016**, *41*, 10–18.
41. S. A. Bonnett, J. R. Whicher, K. Papireddy, G. Florova, J. L. Smith, K. A. Reynolds, *Chem. Biol.* **2013**, *20*, 772–783.
42. W. D. Fiers, G. J. Dodge, D. H. Sherman, J. L. Smith, C. C. Aldrich, *J. Am. Chem. Soc.* **2016**, *138*, 16024–16036.

43. D. H. Kwan, P. F. Leadlay, *ACS Chem. Biol.* **2010**, *5*, 829–838.
44. M. L. Adrover-Castellano, J. J. Schmidt, D. H. Sherman, *ChemCatChem* **2021**, *13*, 2095–2116.
45. J. Beld, E. C. Sonnenschein, C. R. Vickery, J. P. Noel, M. D. Burkart, *Nat. Prod. Rep.*, **2014**, *31*, 61–108.
46. R. H. Lambalot, A. M. Gehring, R. S. Flugel, P. Zuber, M. LaCelle, M. A. Marahiel, R. Reid, C. Khosla, C. T. Walsh, *Chem. Biol.* **1996**, *3*, 923–936.
47. R. S. Gokhale, D. Hunziker, D. E. Cane, C. Khosla, *Chem. Biol.* **1999**, *6*, 117–125.
48. C. Hertweck, A. Luzhetskyy, Y. Rebets, A. Bechthold, *Nat. Prod. Rep.* **2007**, *24*, 162–190.
49. J. R. D. McCormick, *Biogenesis of antibiotic substances* **1965**, 73–91.
50. I. Abe, H. Morita, *Nat. Prod. Rep.* **2010**, *27*, 809–838.
51. J. L. Ferrer, J. M. Jez, M. E. Bowman, R. A. Dixon, J. P. Noel, *Nat. Struct. Biol.* **1999**, *6*, 775–784.
52. I. Abe, T. Watanabe, H. Morita, T. Kohno, H. Noguchi, *Org. Lett.* **2006**, *8*, 499–502.
53. J. H. Kim, M. Komatsu, K. Shin-Ya, S. Omura, H. Ikeda, *Proc. Natl. Acad. Sci.* **2018**, *115*, 6828–6833.
54. K. D. Parris, L. Lin, A. Tam, R. Mathew, J. Hixon, M. Stahl, C. C. Fritz, J. Seehra, W. S. Somers, *Structure* **2000**, *8*, 883–895.
55. S. Rafi, P. Novichenok, S. Kolappan, X. Zhang, C. F. Stratton, R. Rawat, C. Kisker, C. Simmerling, P. J. Tonge, *J. Biol. Chem.* **2006**, *281*, 39285–39293.
56. R. Finking, J. Solsbacher, D. Konz, M. Schobert, A. Schäfer, D. Jahn, M. A. Marahiel, *J. Biol. Chem.* **2002**, *277*, 50293–50302.
57. L. E. Quadri, P. H. Weinreb, M. Lei, M. M. Nakano, P. Zuber, C. T. Walsh, *Biochemistry* **1998**, *37*, 1585–1595.
58. F. Fichtlscherer, C. Wellein, M. Mittag, E. Schweizer, *Eur. J. Biochem.* **2000**, *267*, 2666–2671.
59. R. Bunet, R. Riclea, L. Laureti, L. Hotel, C. Paris, J. M. Girardet, D. Spiteller, J. S. Dickschat, P. Leblond, B. Aigle, *PLoS One* **2014**, *9*, e87607.
60. B. Zhang, W. Tian, S. Wang, X. Yan, X. Jia, G. K. Pierens, W. Chen, H. Ma, Z. Deng, X. Qu, *ACS Chem. Biol.* **2017**, *12*, 1732–1736.
61. M. R. Mofid, M. A. Marahiel, R. Ficner, K. Reuter, *Acta Crystallogr. D Biol. Crystallogr.* **1999**, *55*, 1098–1100.
62. P. Johansson, B. Mulinacci, C. Koestler, R. Vollrath, D. Oesterhelt, M. Grininger,

- Struct.* **2009**, *17*, 1063–1074.
63. R. Finking, M. R. Mofid, M.A. Marahiel, *Biochem.* **2004**, *43*, 8946–8956.
64. M. Rothmann, M. Kang, R. Villa, I. Ntai, J. J. La Clair, N. L. Kelleher, E. Chapman, M. D. Burkart, *J. Am. Chem. Soc.* **2013**, *135*, 5962–5965.
65. N.M. Kosa, R.W. Haushalter, A.R. Smith, M. D. Burkart, *Nat. Methods* **2012**, *9*, 981–984.
66. J. J. La Clair, T. L. Foley, T. R. Schegg, C. M. Regan, M. D. Burkart, *Chem. Biol.* **2004**, *11*, 195–201.
67. K. M. Clarke, A. C. Mercer, J. J. La Clair, M. D. Burkart, *J. Am. Chem. Soc.* **2005**, *127*, 11234–11235.
68. D. Schwarzer, M. A. Marahiel, *Naturwissenschaften* **2001**, *88*, 93–101.
69. E. M. Musiol-Kroll, W. Wohlleben, *Antibiotics* **2018**, *7*, 62.
70. H. Jenke-Kodama, A. Sandmann, R. Müller, E. Dittmann, *Mol. Biol. Evol.* **2005**, *22*, 2027–2039.
71. E. E. Drufva, E. G. Hix, C. B. Bailey, *Synth. Syst. Biotechnol.* **2020**, *5*, 62–80.
72. F. Del Vecchio, H. Petkovic, S. G. Kendrew, L. Low, B. Wilkinson, R. Lill, J. Cortes, B. A. Rudd, J. Staunton, P. F. Leadlay. *J. Ind. Microbiol. Biotechnol.* **2003**, *30*, 489–494.
73. G. Yadav, R. S. Gokhale, D. Mohanty. *J. Mol. Biol.* **2003**, *328*, 335–363.
74. B. J. Dunn, D. E. Cane, C. Khosla, *Biochem.* **2013**, *52*, 1839–1841.
75. J. J. Shen, F. Chen, X. X. Wang, X. F. Liu, X. A. Chen, X. M. Mao, Y. Q. Li, *Front. microbiol.* **2018**, *9*, 1840.
76. Y. Y. Wang, L. F. Bai, X. X. Ran, X. H. Jiang, H. Wu, W. Zhang, M. Y. Jin, Y. Q. Li, H. Jiang, *Protein Pept. Lett.* **2015**, *22*, 2–7.
77. Y. Q. Cheng, G. L. Tang, B. Shen, *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 3149–3154.
78. A. F. Marsden, P. Caffrey, J. F. Aparicio, M. S. Loughran, J. Staunton, P. F. Leadlay, *Science* **1994**, *263*, 378–380.
79. C. W. Liew, M. Nilsson, M. W. Chen, H. Sun, T. Cornvik, Z. X. Liang, J. Lescar, *J. Biol. Chem.* **2012**, *287*, 23203–23215.
80. U. Sundermann, K. Bravo-Rodriguez, S. Klopries, S. Kushnir, H. Gomez, E. Sanchez-Garcia, F. Schulz, *ACS Chem. Biol.* **2013**, *8*, 443–450.
81. M. Grininger, L. Buyachuihan, S. Reiners, Y. Zhao, 2024, DOI: [doi.org/10.21203/rs.3.rs-3914462/v1](https://doi.org/10.21203/rs.3.rs-3914462/v1).

82. B. J. Dunn, C. Khosla, *J. R. Soc. Interface*. **2013**, *10*, 20130297.
83. X. Ruan, A. Pereda, D. L. Stassi, D. Zeidner, R. G. Summers, M. Jackson, A. Shivakumar, S. Kakavas, M. J. Staver, S. Donadio, L. Katz, *J. Bacteriol.* **1997**, *179*, 6416–6425.
84. L. Liu, A. Thamchaipenet, H. Fu, M. Betlach, G. Ashley, *J. Am. Chem. Soc.* **1997**, *119*, 10553–10554.
85. D. L. Stassi, S. J. Kakavas, K. A. Reynolds, G. Gunawardana, S. Swanson, D. Zeidner, M. Jackson, H. Liu, A. Buko, L. Katz, *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 7305–7309.
86. I. Koryakina, C. Kasey, J. B. McArthur, A. N. Lowell, J. A. Chemler, S. Li, D. A. Hansen, D. H. Sherman, G. J. Williams, *ACS Chem. Biol.* **2017**, *12*, 114–123.
87. A. Miyanaga, S. Iwasawa, Y. Shinohara, F. Kudo, T. Eguchi, *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 1802–1807.
88. Z. Ye, E. M. Musiol, T. Weber, G. J. Williams, *Chem. Biol.* **2014**, *21*, 636–646.
89. C. Oefner, H. Schulz, A. D'Arcy, G. E. Dale, *Acta Crystallogr. D Biol. Crystallogr.* **2006**, *62*, 613–618.
90. Y. Chen, E. E. Kelly, R. P. Masluk, C. L. Nelson, D. C. Cantu, P. J. Reilly, *Protein Sci.* **2011**, *20*, 1659–1667.
91. A. Chen, Z. Jiang, M. D. Burkart, *Chem. Sci.* **2022**, *13*, 4225–4238.
92. W. G. J. Hol, P. T. van Duijnen, H. J. C. Berendsen, *Nature* **1978**, *273*, 443–446.
93. S. W. White, J. Zheng, Y. M. Zhang, C. O. Rock, *Annu. Rev. Biochem.* **2005**, *74*, 791–831.
94. A. Witkowski, A. K. Joshi, S. Smith, *Biochemistry* **2002**, *41*, 10877–10887.
95. M. Jenner, S. Frank, A. Kampa, C. Kohlhaas, P. Pöplau, G. S. Briggs, J. Piel, N. J. Oldham, *Angew. Chem. Int. Ed.* **2013**, *52*, 1143–1147.
96. C. Kohlhaas, M. Jenner, A. Kampa, G. S. Briggs, J. P. Afonso, J. Piel, N. J. Oldham, *Chem. Sci.* **2013**, *4*, 3212–3217.
97. L. Du, C. Sánchez, B. Shen, *Metab. Eng.* **2001**, *3*, 78–95.
98. D. C. Gay, G. Gay, A. J. Axelrod, M. Jenner, C. Kohlhaas, A. Kampa, N. J. Oldham, J. Piel, A. T. Keatinge-Clay, *Structure* **2014**, *22*, 444–451.
99. W. Lee, B. Engels, *Biochem.* **2014**, *53*, 919–931.
100. J. Wang, S. M. Soisson, K. Young, W. Shoop, S. Kodali, A. Galgoci, R. Painter, G. Parthasarathy, Y. S. Tang, R. Cummings, S. Ha, K. Dorso, M. Motyl, H. Jayasuriya, J. Ondeyka, K. Herath, C. Zhang, L. Hernandez, J. Allocco, Á. Basilio,

- J. R. Tormo, O. Genilloud, F. Vicente, F. Pelaez, L. Colwell, S. H. Lee, B. Michael, T. Felcetto, C. Gill, L. L. Silver, J. D. Hermes, K. Bartizal, J. Barrett, D. Schmatz, J. W. Becker, D. Cully, S. B. Singh, *Nature* **2006**, *441*, 358–361.
101. J. T. Mindrebo, A. Chen, W. E. Kim, R. N. Re, T. D. Davis, J. P. Noel, M. D. Burkart, *ACS Catal.* **2021**, *11*, 6787–6799.
102. A. Bräuer, Q. Zhou, G. L. C. Grammbitter, M. Schmalhofer, M. Ruhl, V. R. I. Kaila, H. B. Bode, M. Groll, *Nat. Chem.* **2020**, *12*, 755–763.
103. J. C. Milligan, D. J. Lee, D. R. Jackson, A. J. Schaub, J. Beld, J. F. Barajas, J. J. Hale, R. Luo, M. D. Burkart, S. C. Tsai, *Nat. Chem. Biol.* **2019**, *15*, 669–671.
104. T. G. Bartholow, T. Sztain, A. Patel, D. J. Lee, M. A. Young, R. Abagyan, M. D. Burkart, *Commun. Biol.* **2021**, *4*, 340.
105. Z. Yin, J. S. Dickschat, *Nat. Prod. Rep.* **2021**, *38*, 1445–1468.
106. P. Caffrey, *ChemBioChem* **2003**, *4*, 654–657.
107. A. T. Keatinge-Clay, *Nat. Prod. Rep.* **2016**, *33*, 141–149.
108. A. T. Keatinge-Clay, R. M. Stroud, *Structure* **2006**, *14*, 737–748.
109. A. T. Keatinge-Clay, *Chem. Biol.* **2007**, *14*, 898–908.
110. J. Zheng, A. T. Keatinge-Clay, *Med. Chem. Commun.* **2013**, *4*, 34–40.
111. D. L. Akey, J. R. Razelun, J. Tehranisa, D. H. Sherman, W. H. Gerwick, J. L. Smith, *Structure* **2010**, *18*, 94–105.
112. D. H. Kwan, Y. Sun, F. Schulz, H. Hong, B. Popovic, J. C. C. Sim-Stark, S. F. Haydock, P. F. Leadlay, *Chem. Biol.* **2008**, *15*, 1231–240.
113. C. M. Amy, A. Witkowski, J. Naggert, B. Williams, Z. Randhawa, S. Smith, *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86*, 3114–3118.
114. D. Khare, W. A. Hale, A. Tripathi, L. Gu, D. H. Sherman, W. H. Gerwick, K. Håkansson, J. L. Smith, *Structure* **2015**, *23*, 2213–2223.
115. A. Koglin, C. T. Walsh, *Nat. Prod. Rep.* **2009**, *26*, 987–1000.
116. T. Stachelhaus, H. D. Mootz, V. Bergendahl, M.A. Marahiel, *J. Biol. Chem.* **1998**, *273*, 22773–22781.
117. R. Finking, M. A. Marahiel, *Annu. Rev. Microbiol.* **2004**, *58*, 453–488.
118. M. A. Fischbach, C. T. Walsh, *Chem. Rev.* **2006**, *106*, 3468–3496.
119. K. J. Weissman, *Nat. Chem. Biol.* **2015**, *11*, 660–670.
120. M. A. Marahiel, *Nat. Prod. Rep.* **2016**, *33*, 136–140.
121. C. T. Walsh, *Nat. Prod. Rep.* **2016**, *33*, 127–135.
122. G. L. Challis, J. H. Naismith, *Curr. Opin. Struct. Biol.* **2004**, *14*, 748–756.

123. K. A. J. Bozhüyük, F. Fleischhacker, A. Linck, F. Wesche, A. Tietze, C. P. Niesert, H. B. Bode, *Nat. Chem.* **2018**, *10*, 275–281.
124. J. L. Faylo, T. A. Ronnebaum, D. W. Christianson, *Acc. Chem. Res.* **2021**, *54*, 3780–3791.
125. J. S. Dickschat, H. B. Bode, T. Mahmud, R. Müller, S. Schulz, *J. Org. Chem.* **2005**, *70*, 5174–5182.
126. J. Jiang, X. He, D. E. Cane, *Nat. Chem. Biol.* **2007**, *3*, 711–715.
127. Z. Yin, M. Maczka, G. Schnakenburg, S. Schulz, J. S. Dickschat, *Org. Biomol. Chem.* **2024**, *22*, 5748–5758.
128. T. Mitsuhashi, M. Okada, I. Abe, *ChemBioChem* **2017**, *18*, 2104–2109.
129. T. A. Ronnebaum, S. A. Eaton, E. A. E. Brackhahn, D. W. Christianson, *Biochemistry* **2021**, *60*, 3162–3172.
130. T. A. Ronnebaum, K. Gupta, D. W. Christianson, *J. Struct. Biol.* **2020**, *210*, 107463.
131. R. J. Peters, M. M. Ravn, R. M. Coates, R. B. Croteau, *J. Am. Chem. Soc.* **2001**, *123*, 8974–8978.
132. M. Rohmer, M. Knani, P. Simonin, B. Sutter, H. Sahm, *Biochem. J.* **1993**, *295*, 517–524.
133. M. Rohmer, *Nat. Prod. Rep.* **1999**, *16*, 565–574.
134. J. S. Dickschat, *Nat. Prod. Rep.* **2016**, *33*, 87–110.
135. Z. Yin, J. S. Dickschat, *Nat. Prod. Rep.* **2023**, *40*, 28–45.
136. D. W. Christianson, *Chem. Rev.* **2017**, *117*, 11570–11648.
137. J. T. Wang, T. T. Shi, L. Ding, J. Xie, P. J. Zhao, *Catalysts* **2023**, *13*, 581.
138. Y. H. Chooi, Y. J. Hong, R. A. Cacho, D. J. Tantillo, Y. Tang, *J. Am. Chem. Soc.* **2013**, *135*, 16805–16808.
139. J. D. Rudolf, C. Y. Chang, *Nat. Prod. Rep.* **2020**, *37*, 425–463.
140. X. Lin, R. Hopson, D. E. Cane, *J. Am. Chem. Soc.* **2006**, *128*, 6022–6023.
141. B. Zhao, X. Lin, L. Lei, D. C. Lamb, S. L. Kelly, M. R. Waterman, D. E. Cane, *J. Biol. Chem.* **2008**, *283*, 8183–8189.
142. B. Zhao, L. Lei, D. G. Vassylyev, X. Lin, D. E. Cane, S. L. Kelly, H. Yuan, D. C. Lamb, M. R. Waterman, *J. Biol. Chem.* **2009**, *284*, 36711–36719.
143. X. Lin, D. E. Cane, *J. Am. Chem. Soc.* **2009**, *131*, 6332–6333.
144. C. C. Hyde, S. A. Ahmed, E. A. Padlan, E. W. Miles, D. R. Davies, *J. Biol. Chem.* **1988**, *263*, 17857–17871.

145. C. Binda, R. T. Bossi, S. Wakatsuki, S. Arzt, A. Coda, B. Curti, M. A. Vanoni, A. Mattevi, *Structure* **2000**, *8*, 1299–1308.
146. R. Anand, A. A. Hoskins, J. Stubbe, S. E. Ealick, *Biochemistry* **2004**, *43*, 10328–10342.
147. 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.
148. J. M. McGuire, R. L. Bunch, R. C. Anderson, H. E. Boaz, E. H. Flynn, H. M. Powell, J. W. Smith, *Antibiot. Chemother.* **1952**, *2*, 281–283.
149. C. Vézina, A. Kudelski, S. N. Sehgal, *J. Antibiot.* **1975**, *28*, 721–726.
150. J. L. Meier, A. C. Mercer, H. Rivera Jr, M. D. Burkart, *J. Am. Chem. Soc.* **2006**, *128*, 12174–12184.
151. P. J. Belshaw, C. T. Walsh, T. Stachelhaus, *Science* **1999**, *284*, 486–489.
152. S. A. Sieber, C. T. Walsh, M. A. Marahiel, *J. Am. Chem. Soc.* **2003**, *125*, 10862–10866.
153. M. Pavlidou, E. K. Pross, E. M. Musiol, A. Kulik, W. Wohlleben, T. Weber, *FEMS Microbiol. Lett.* **2011**, *319*, 26–33.
154. A. M. Gehring, R. H. Lambalot, K. W. Vogel, D. G. Drueckhammer, C. T. Walsh, *Chem. Biol.* **1997**, *4*, 17–24.
155. R. B. Hamed, L. Henry, J. R. Gomez-Castellanos, A. Asghar, J. Brem, T. D. Claridge, C. J. Schofield, *Org. Biomol. Chem.* **2013**, *11*, 8191–8196.
156. R. F. Seipke, J. Barke, C. Brearley, L. Hill, D. W. Yu, R. J. M. Goss, M. I. Hutchings, *PloS One* **2011**, *6*, e22028.
157. M. Rawlings, J. E. Cronan, *J. Biol. Chem.* **1992**, *267*, 5751–5754.
158. L. Han, K. Yang, E. Ramalingam, R. H. Mosher, L. C. Vining, *Microbiology* **1994**, *140*, 3379–3389.
159. L. Wang, J. McVey, L. C. Vining, *Microbiology* **2001**, *147*, 1535–1545.
160. P. E. Coderre, C. F. Earhart, *J. Gen. Microbiol.* **1989**, *135*, 3043–3055.
161. A. M. Gehring, K. A. Bradley, C. T. Walsh, *Biochemistry* **1997**, *36*, 8495–8503.
162. R. A. Butcher, F. C. Schroeder, M. A. Fischbach, P. D. Straight, R. Kolter, C. T.

- Walsh, J. Clardy, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 1506–1509.
163. C. T. Calderone, S. B. Bumpus, N. L. Kelleher, C. T. Walsh, N. A. Magarvey, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 12809–12814.
164. T. Nguyen, K. Ishida, H. Jenke-Kodama, E. Dittmann, C. Gurgui, T. Hochmuth, S. Taudien, M. Platzer, C. Hertweck, J. Piel, *Nat. Biotechnol.* **2008**, *26*, 225–233.
165. K. Weissman, *Beilstein J. Org. Chem.* **2017**, *13*, 348–371.
166. C. R. Valenzano, Y.-O. You, A. Garg, A. Keatinge-Clay, C. Koshla, D. E. Cane, *J. Am. Chem. Soc.* **2010**, *132*, 14697–14699.
167. X. Guo, T. Liu, C. R. Valenzano, Z. Deng, D. E. Cane, *J. Am. Chem. Soc.* **2010**, *132*, 14694–14696.
168. D. Gay, Y.-O. You, A. Keatinge-Clay, D. E. Cane, *Biochemistry* **2013**, *52*, 8916–8928.
169. O. Vergnolle, F. Hahn, A. Baerga-Ortiz, P. F. Leadlay, J. N. Anderson, *ChemBioChem* **2011**, *12*, 1011–1014.
170. D. D. Shah, Y.-O. You, D. E. Cane, *J. Am. Chem. Soc.* **2017**, *139*, 14322–14330.
171. X. Xie, D. E. Cane, *Biochemistry* **2018**, *57*, 3126–3129.
172. T. Katsuki, K. B. Sharpless, *J. Am. Chem. Soc.* **1980**, *102*, 5974–5976.
173. N. Di Blasio, M. T. Lopardo, P. Lupattelli, *Eur. J. Org. Chem.* **2009**, *6*, 938–944.
174. J. S. Dickschat, O. Vergnolle, H. Hong, S. Garner, S. R. Bidgood, H. C. Dooley, Z. Deng, P. F. Leadlay, Y. Sun, *ChemBioChem* **2011**, *12*, 2408–2412.
175. D. Thomy, E. Culp, M. Adamek, E. Y. Cheng, N. Ziemert, G. D. Wright, P. Sass, H. Brötz-Oesterhelt, *Appl. Environ. Microbiol.* **2019**, *85*, e01292–19.
176. S. Mohan, T. M. Kelly, S. S. Eveland, C. R. H. Raetz, M. S. Anderson, *J. Biol. Chem.* **1994**, *269*, 32896–32903.
177. Z. Yin, E. Liebhart, E. Stegmann, H. Brötz-Oesterhelt, J. S. Dickschat *Org. Chem. Front.* **2022**, *9*, 2714–2720.
178. I. Koyakina, J. B. McArthur, M. W. Draelos, G. J. Williams, *Org. Biomol. Chem.* **2013**, *11*, 4449–4458.
179. E. Kalkreuter, K. S. Bingham, A. M. Keeler, A. N. Lowell, J. J. Schmidt, D. H. Sherman, G. J. Williams, *Nat. Commun.* **2021**, *12*, 2193.

180. A. Pinto, M. Wang, M. Horsman, C. N. Boddy, *Org. Lett.* **2012**, *14*, 2278–2281.
181. W. D. Fiers, G. J. Dodge, Y. Li, J. L. Smith, R. A. Fecik, C. C. Aldrich, *Chem. Sci.* **2015**, *6*, 5027–5033.
182. Y. Li, G. J. Dodge, W. D. Fiers, R. A. Fecik, J. L. Smith, C. C. Aldrich, *J. Am. Chem. Soc.* **2015**, *137*, 7003–7006.
183. A. Faille, S. Gavalda, N. Slama, C. Lherbet, L. Maveyraud, V. Guillet, F. Laval, A. Quemard, L. Mourey, J.-D. Pedelacq, *J. Mol. Biol.* **2017**, *429*, 1554–1569.
184. K. Finzel, C. Nguyen, D. R. Jackson, A. Gupta, S. C. Tsai, M. D. Burkhardt, *Chem. Biol.* **2015**, *22*, 1453–1460.
185. D. A. Evans, J. Bartroli, T. L. Shih, *J. Am. Chem. Soc.* **1981**, *103*, 2127–2129.
186. C. Olano, B. Wilkinson, C. Sanchez, S. J. Moss, R. Sheridan, V. Math, A. J. Weston, A. F. Brana, C. J. Martin, M. Oliynyk, C. Mendez, P. F. Leadlay, J. A. Salas, *Chem. Biol.* **2004**, *11*, 87–97.
187. T. Schupp, C. Toupet, N. Engel, S. Goff, *FEMS Microbiol. Lett.* **1998**, *159*, 201–267.
188. L. Kaysser, L. Lutsch, S. Siebenberg, E. Wemakor, B. Kammerer, B. Gust, *J. Biol. Chem.* **2009**, *284*, 14987–14996.
189. N. N. Gerber, H. A. Lechevalier, *Appl. Microbiol.* **1965**, *13*, 935–938.
190. R. G. Buttery, J. A. Garibaldi, *J. Agric. Food Chem.* **1976**, *24*, 1246–1247.
191. X. He, D. E. Cane, *J. Am. Chem. Soc.* **2004**, *126*, 2678–2679.
192. H. Xu, J. S. Dickschat, *ChemBioChem* **2023**, *24*, e202300101.
193. W. A. Ayer, M. G. Paice, *Can. J. Chem.* **1976**, *54*, 910–916.
194. N. Ding, L. Han, Y. Jiang, G. Li, J. Liu, Y. Mu, X. Huang, *Magn. Res. Chem.* **2018**, *56*, 352–359.
195. L. Ding, C. Hertweck, *J. Nat. Prod.* **2020**, *83*, 2207–2211.

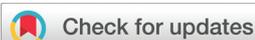
196. N. Ding, Y. Jiang, L. Han, X. Chen, J. Ma, X. Qu, Y. Mu, J. Liu, L. Li, C. Jiang, X. Huang, *J. Nat. Prod.* **2016**, *79*, 799–805.
197. R. W. Eaton, P. Sandusky, *Biodegradation* **2010**, *21*, 71–79.
198. R. Kaiser, C. Nussbaumer, *Helv. Chim. Acta.* **1990**, *73*, 133–139.
199. A. Rieck, N. Bülow, W. A. König, *Phytochemistry* **1995**, *40*, 847–851.
200. W. A. Ayer, L. M. Browne, S. Fung, *Can. J. Chem.* **1976**, *54*, 3276–3282.
201. Q. Liu, L. Han, B. Qin, Y. Mu, P. Guan, S. Wang, X. Huang, *Org. Chem. Front.* **2018**, *5*, 1719–1723.
202. U. Huber, W. Boland, W. A. König, B. Gehrcke, *Helv. Chim. Acta.* **1993**, *76*, 1949–1955.
203. G. Revial, *Tetrahedron Lett.* **1989**, *30*, 4121–4124.
204. A. Saito, A. Tanaka, T. Oritani, *Tetrahedron Asymmetr.* **1996**, *7*, 2923–2928.
205. A. Tanaka, T. Tokuyama, A. Saito, T. Oritani, *Biosci. Biotechnol. Biochem.* **1998**, *62*, 2435–2437.
206. H. J. Swarts, A. A. Haaksma, B. J. M. Jansen, A. de Groot, *Tetrahedron* **1992**, *48*, 5497–5508.
207. D. C. Behenna, B. M. Stoltz, *J. Am. Chem. Soc.* **2004**, *126*, 15044–15045.
208. T. L. Ho, C. Y. Su, *J. Org. Chem.* **2000**, *65*, 3566–3568.
209. C. A. Citron, J. Gleitzmann, G. Laurenzano, R. Pukall, J. S. Dickschat, *ChemBioChem* **2012**, *13*, 202–214.
210. Z. Yin, H. Xu, J. S. Dickschat, *Org. Biomol. Chem.* **2024**, *22*, 8714–8719.
211. Z. Yin, J. S. Dickschat, *Beilstein J. Org. Chem.* **2024**, *5*, 734–740.
212. Z. Yin, D. Bär, B. Gust, J. S. Dickschat, *Org. Biomol. Chem.* **2022**, *20*, 9103–9107.
213. Z. Yin, M. Maczka, Gr. Schnakenburg, S. Schulzb, J. S. Dickschat, *Org. Biomol. Chem.* **2024**, *22*, 5748–5758.

## Appendix A

### **A Clickable Coenzyme A Derived Probe for Investigating Phosphopantetheinyl Transferase Activity in Natural Product Biosynthesis**

*Org. Biomol. Chem.* **2024**, 22, 8714–8719.

[doi.org/10.1039/D4OB01485E](https://doi.org/10.1039/D4OB01485E)



Cite this: *Org. Biomol. Chem.*, 2024, **22**, 8714

Received 10th September 2024,  
Accepted 5th October 2024

DOI: 10.1039/d4ob01485e

rsc.li/obc

## A clickable coenzyme A derived probe for investigating phosphopantetheinyl transferase activity in natural product biosynthesis†

Zhiyong Yin, Houchao Xu and Jeroen S. Dickschat \*

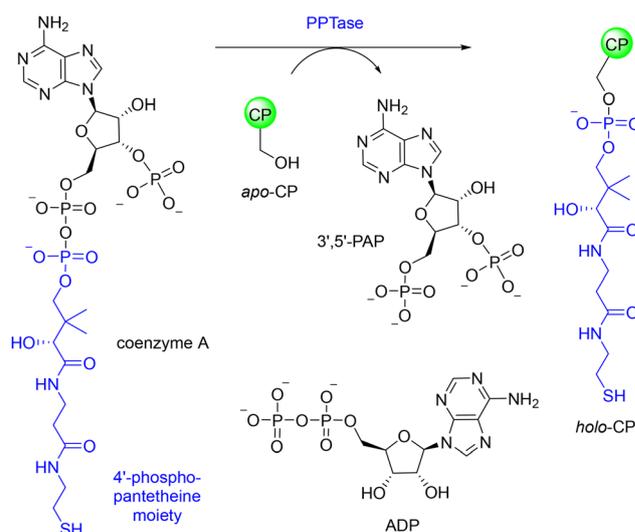
**Phosphopantetheinyl transferases activate carrier proteins through attachment of a coenzyme A derived phosphopantetheinyl linker. This study describes a method to monitor this process through a modified HSCoA with an alkyne group, allowing for the Cu-catalysed alkyne–azide cycloaddition of a fluorescent tag. Application of the method in an enzyme screening resulted in the identification of new promiscuous PPTases.**

Natural products synthesised by polyketide synthases (PKSs) and nonribosomal peptide synthetases (NRPSs) are immensely important because of their diverse biological activities and therapeutic potential. Notable examples include the antibiotic erythromycin,<sup>1</sup> the anticancer agent bleomycin,<sup>2</sup> and the immunosuppressant rapamycin.<sup>3</sup> Like fatty acids, these complex molecules are assembled through modular and iterative processes that rely on carrier proteins (CPs) for the efficient transfer of biosynthetic intermediates. A critical step in the biosynthesis of these natural products is the activation of CPs by phosphopantetheinyl transferases (PPTases).<sup>4</sup>

PPTases catalyse the transfer of a 4'-phosphopantetheinyl (PPant) arm from coenzyme A (HSCoA) to a conserved serine residue on the CPs (Scheme 1). This modification is pivotal in converting inactive *apo*-CPs into their active *holo* forms. The active *holo*-CPs are able to carry biosynthetic intermediates and to shuttle them through the various stages of the PKS or NRPS assembly line. The PPant arm is a 20 Å long prosthetic group that functions as a flexible linker and ends with a thiol group. This thiol group holds the biosynthetic intermediates *via* a thioester bond, which is crucial for maintaining the dynamic interactions required for the complex enzymatic transformations in the biosynthesis of natural products. The phosphopantetheinylation of CPs by PPTases is not only a

necessary step for the functionality of fatty acid synthases (FASs), PKSs and NRPSs, but also a key regulatory mechanism in fatty acid, polyketide and non-ribosomal peptide biosynthesis.<sup>5</sup>

The PPTase Sfp from *Bacillus subtilis*, the enzyme that modifies surfactin synthase,<sup>6</sup> stands out for its remarkable substrate promiscuity and broad applicability in metabolic engineering. Sfp is notable for its ability to accommodate a diverse array of functionalised thioesters derived from HSCoA such as acetyl-S-CoA and malonyl-S-CoA and can also act on a large variety of CPs. This flexibility has been exploited to attach non-natural substrates and intermediates to CPs in mechanistic studies of diverse PKSs and NRPSs.<sup>7–9</sup> The R4-4 variant of Sfp displaying three amino acid exchanges (K28E, T44E and C77Y) has an increased catalytic efficiency and a substrate preference for 3'-dephospho-CoA.<sup>10</sup> This enzyme variant can even upload adenosine diphosphate (ADP), lacking the 3'-phosphate and the complete pantetheine portion, and its phosphate ester



Scheme 1 Phosphopantetheinyl transferases activate carrier proteins.

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

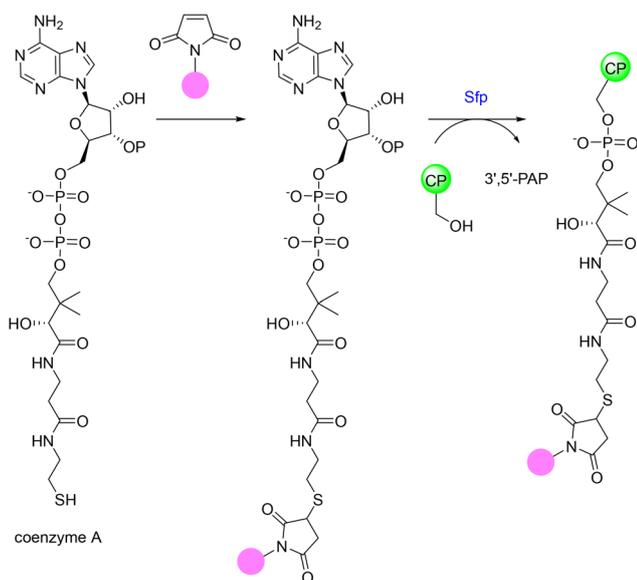
† Electronic supplementary information (ESI) available: Experimental procedures, information about the enzymes used in this study, results of PPTase fluorescence assays, and NMR spectra of synthetic compounds. See DOI: <https://doi.org/10.1039/d4ob01485e>

derivatives.<sup>11</sup> Also *Escherichia coli* holo-ACP synthase (ACPS) involved in fatty acid biosynthesis is known to have a broad substrate tolerance.<sup>12</sup>

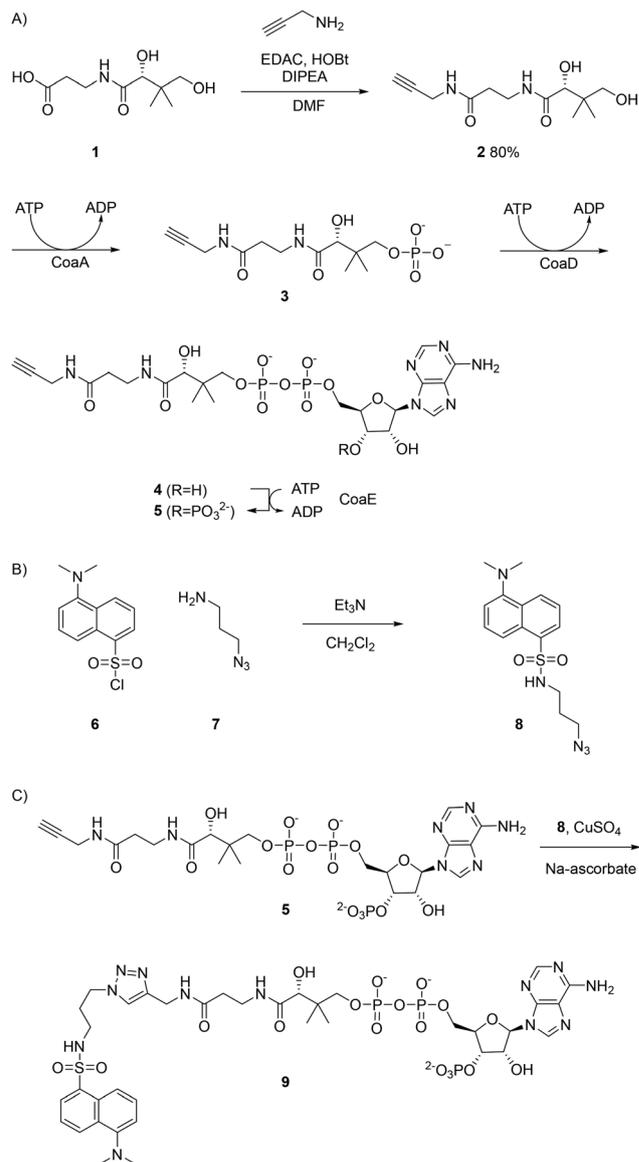
Monitoring the PPTase-catalysed phosphopantetheinylation of CPs is crucial for understanding the biosynthesis of polyketides and non-ribosomal peptides. Conventional methods for assaying this modification include the use of radiolabelled HSCoA<sup>4,6,13–15</sup> and HPLC-MS analysis of purified proteins.<sup>9,12,14</sup> These techniques, although effective, present challenges, either because they are associated with safety issues, or because they require costly equipment. An alternative interesting approach developed by Burkart and coworkers makes use of fluorescence labelling tags.<sup>8</sup> The strategy applies the Michael addition of HSCoA to a maleimide-fluorescence label conjugate, followed by upload to a CP using the PPTase Sfp (Scheme 2). Proteins carrying the fluorescence tag can then easily be visualised *e.g.* in protein gel electrophoresis (SDS-PAGE).

As a drawback, this method relies on the highly promiscuous PPTase Sfp, while other PPTases with a narrow substrate tolerance may not accept the substantial structural modifications associated with the HSCoA-Michael adducts. Here we present a new method making use of fluorescence labels for the visualisation of CP modifications by PPTases. Our approach uses an HSCoA derivative with an alkyne substituting for the thiol group. This minimal structural modification is tolerated by various PPTases, and after attachment of the modified PPant linker to CPs the fluorescence tag can be added through Cu-catalysed alkyne-azide cycloaddition (CuAAC, “click chemistry”).

The clickable HSCoA derivative was prepared in a combined strategy using chemical synthesis and a one-pot enzymatic transformation (Scheme 3A). Starting from pantothenate (**1**), a coupling with propargylamine using the hydrochloride salt of



Scheme 2 Burkart's fluorescence labelling strategy of CPs using Sfp.



Scheme 3 Synthesis of (A) the HSCoA derivative **5**, (B) the fluorescent probe **8**, and (C) the HSCoA derivative **9**.

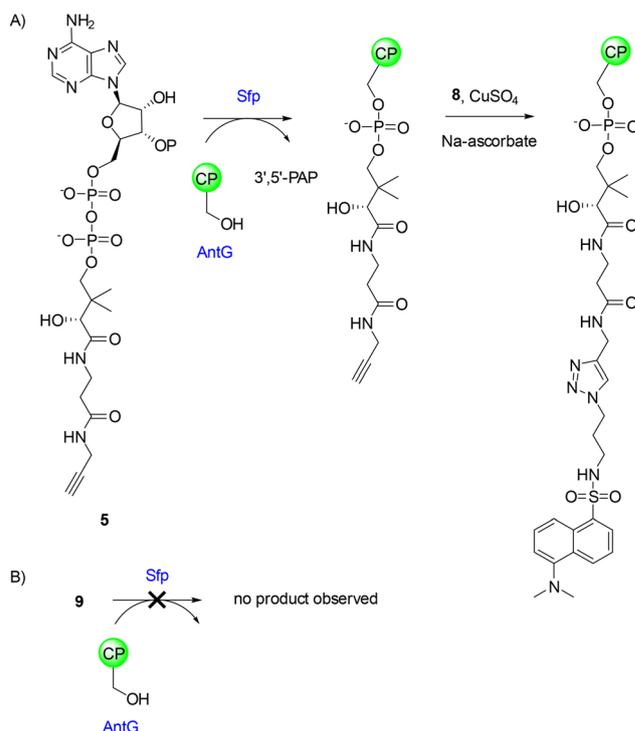
*N*-(3-dimethylaminopropyl)-*N*-ethylcarbodiimide (EDAC), 1-hydroxybenzotriazole (HOBt) and Hünig's base (DIPEA) yielded **2**. This material was further transformed into the HSCoA derivative **5** using the purified recombinant enzymes CoaA, CoaD and CoaE (Fig. S1†) from the *E. coli* coenzyme A biosynthesis pathway<sup>16–19</sup> in conjunction with ATP. The resulting product **5** was purified by HPLC and subsequently used as a substrate with a variety of CPs and PPTases. This approach leverages the promiscuity of *Escherichia coli* CoaA, CoaD and CoaE that has previously been used by other researchers in the enzymatic synthesis of HSCoA derivatives.<sup>20–27</sup> One example that is particularly notable in the context of this study is the development of modified pantetheine analogs, including analogs carrying fluorescent dyes, that can be converted with CoaADE into the corresponding coenzyme A analogs,<sup>28</sup>

followed by Sfp coupling to *E. coli* AcpP<sup>29</sup> from fatty acid biosynthesis. This system can also be used *in vivo*, but also in this case the intermediate coenzyme A analogs show substantial structural modifications.<sup>28</sup>

The synthesis of the fluorescent probe (Scheme 3B) involved the Einhorn amidation of 5-(dimethylamino)naphthalene-1-sulfonyl chloride (**6**) with 3-azidopropan-1-amine (**7**) to yield **8**. For additional enzyme testings also the adduct **9** was obtained by CuAAC between **5** and **8** (Scheme 3C).

First experiments to investigate the activation of CPs were performed with the highly promiscuous PPTase Sfp. Recombinant AntG, a discrete CP from the antimycin biosynthetic gene cluster in *Streptomyces ambofaciens*,<sup>30</sup> was expressed in *E. coli* and purified (Fig. 1, lane 1; Fig. S2†). The incubation of AntG with Sfp and the HSCoA analog **5** using a reported protocol for PPTase activity assays<sup>31</sup> did not lead to any obvious changes (Fig. 1, lane 2), but supposedly resulted in the attachment of a PPant arm analog to the highly conserved Ser of AntG. Successful attachment of the PPant arm analog was demonstrated through addition of the azide **8**, CuSO<sub>4</sub> and sodium ascorbate to facilitate the CuAAC reaction (Scheme 4A), leading to a fluorescent band on the expected height of AntG on the protein gel (Fig. 1, lane 3). Such a fluorescent band was not observed when Sfp and AntG were incubated with the cycloadduct **9**, demonstrating that the substantial structural modification in this HSCoA analog was not tolerated by the enzyme system (Scheme 4B; Fig. 1, lane 4).

To exclude a non-covalent interaction between **5** and AntG and to demonstrate substrate upload to its highly conserved serine, the AntG-S34A variant was generated *via* site-directed mutagenesis. The incubation of wild-type AntG and the AntG-S34A variant with Sfp and **5** followed by cycloaddition of

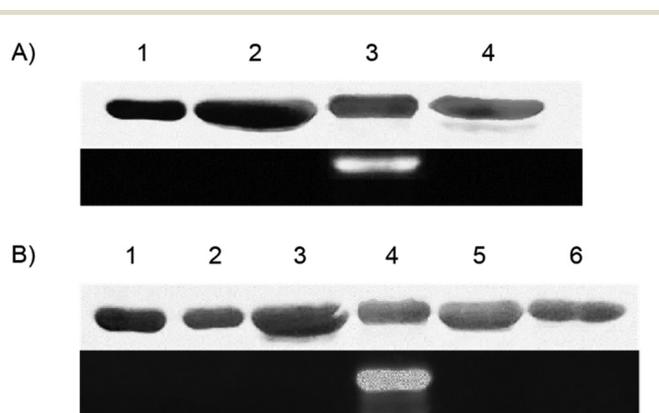


**Scheme 4** Fluorescence assay for the activation of AntG by Sfp. (A) Upload of the HSCoA analog **5** to AntG and subsequent CuAAC with **8**. (B) Upload of the cycloadduct **9** to AntG by Sfp is unsuccessful.

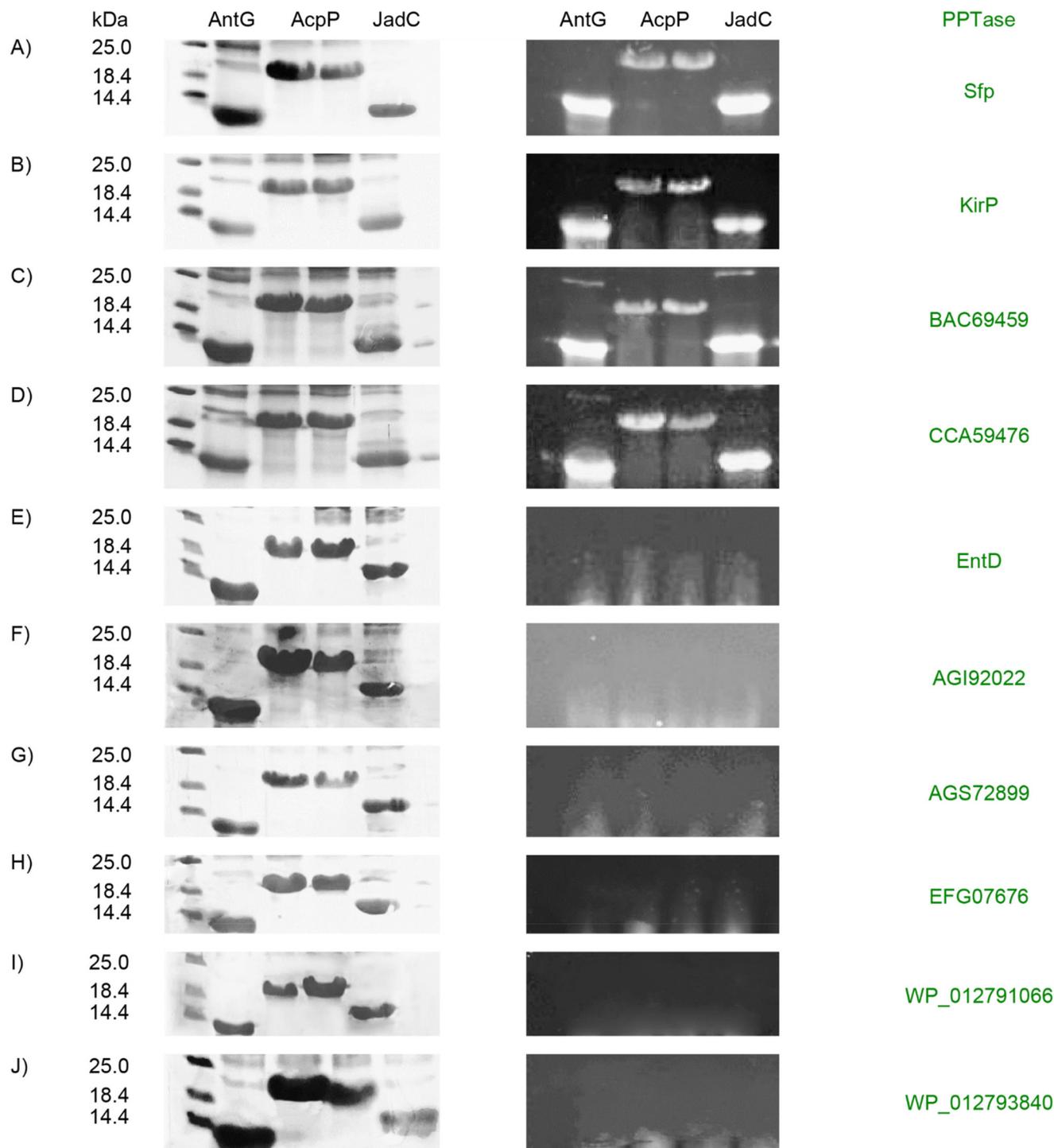
**8** only resulted in a fluorescent band for wildtype AntG, but not for its S34A variant (Fig. 1, lanes 4 and 6; Fig. S3†). Taken together, these findings underscore the applicability of the HSCoA analog **5** in monitoring the activation of AntG by Sfp and demonstrate its upload to the highly conserved Ser34 of AntG.

After establishment of the fluorescence assay, additional experiments were performed to investigate the activation of other CPs by Sfp (Table S4†). This included AcpP from fatty acid biosynthesis in *E. coli* K-12<sup>29</sup> and JadC from jadomycin biosynthesis in *Streptomyces venezuelae* that is naturally activated by the PPTase JadM.<sup>32,33</sup> Also additional PPTases were administered in an enzyme screening, including KirP from kirromycin biosynthesis in *Streptomyces collinus*,<sup>12,34</sup> EntD from enterobactin biosynthesis in *E. coli*,<sup>35,36</sup> and several other PPTases from different branches of a phylogenetic tree constructed from the amino acid sequences of 219 PPTase homologs (Fig. S4†).

Both AcpP and JadC were also activated with **5** by Sfp, leading to fluorescent bands in the protein gels after cycloaddition of **8** (Fig. 2A and Fig. S5†). Activations of AntG, AcpP and JadC with **5** were also observed with KirP and two PPTases from *Streptomyces avermitilis* (accession number BAC69459) and *S. venezuelae* (CCA59476) (Fig. 2B–D and Fig. S6–S8†). All other tested PPTases (EntD, AGI92022 from *Streptomyces albus*, AGS72899 from *S. collinus*, EFG07676 from *Streptomyces clavuligerus*, and WP\_012791066 and WP\_012793840 from



**Fig. 1** SDS-PAGE fluorescence assay for the activation of AntG by Sfp. (A) Lane 1: purified AntG; lane 2: purified Sfp + AntG + **5**; lane 3: purified Sfp + AntG + **5** + **8** after CuAAC; lane 4: purified Sfp + AntG + **9** after CuAAC. (B) Lane 1: purified AntG; lane 2: purified AntG-S34A; lane 3: Sfp + AntG + **5**; lane 4: Sfp + AntG + **5** + **8** after CuAAC; lane 5: Sfp + AntG-S34A + **5**; lane 6: Sfp + AntG-S34A + **5** + **8** after CuAAC. The top rows show the protein gel after Coomassie Brilliant Blue staining, the bottom rows show the same gel before staining under UV irradiation ( $\lambda = 360$  nm). Full size gels including a protein marker and additional control experiments are shown in Fig. S2 and S3.†



**Fig. 2** SDS-PAGE fluorescence assay for the activation of three different carrier proteins by ten different Sfps. On each gel the lanes contain from left to right: protein marker, assay with AntG, assay with AcpP (two lanes), and the assay with JadC (as mentioned on top). The assays were performed with the PPTases as mentioned in green to the right of the gels. The left column shows the protein gels after Coomassie Brilliant Blue staining, the right column shows the same gel before staining under UV irradiation ( $\lambda = 360$  nm). Full size gels including a protein marker and additional control experiments are shown in Fig. S5–S14.†

*Chitinophaga pinensis*) failed to activate any of the three tested CPs (Fig. 2E–J and Fig. S9–S14†). All CP activations that were successful according to the fluorescence assay developed in this study were confirmed by MALDI-ToF MS (Fig. S15–S17†).

## Conclusions

Classical methods to investigate the PPTase mediated activation of CPs include radiolabelling experiments or the appli-

cation of MS-based methods. Disadvantages of these approaches reside in safety concerns and high costs of the analytical instrumentation. An alternative method presented by Burkart made use of fluorescence tags. In this case the label was first attached to HSCoA through Michael addition to a fluorescent maleimide derivative, followed by the PPTase catalysed upload of the HSCoA derivative to a CP.<sup>8</sup> This method is successful, if the investigated enzyme system shows a broad substrate tolerance, requiring flexibility of both the PPTase and the CP. As we have demonstrated here, the combination of the highly promiscuous PPTase Sfp and the CP AntG failed to show this flexibility. For this reason, the HSCoA derivative 5 with a minor structural modification by exchange of a thiol to an alkyne function was developed. After its PPTase mediated upload to a CP the CuAAC reaction can be used to click the alkyne to a fluorescent azide, allowing for an efficient monitoring of CP activations by PPTases. A subsequent screening of various PPTases and CPs revealed that not only Sfp and KirP, but also two previously uncharacterised PPTases from *S. avermitilis* (BAC69459) and *S. venezuelae* (CCA59476) can activate a wide range of CPs. These findings highlight the potential of the clickable HSCoA derivative 5 in studying PPTase-CP interactions and expand the options of promiscuous PPTases beyond Sfp.

## Data availability

The data supporting this article have been included as part of the ESI.†

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

This work was funded by the DFG (SFB TRR 261, project ID 398967434). We thank Prof. Jörn Piel (ETH Zürich) for *Bacillus amyloliquefaciens* FZB42 and Andreas Schneider for HPLC purifications.

## References

- J. M. McGuire, R. L. Bunch, R. C. Anderson, H. E. Boaz, E. H. Flynn, H. M. Powell and J. W. Smith, *Antibiot. Chemother.*, 1952, **2**, 281.
- H. Suzuki, K. Nagai, H. Yamaki, N. Tanaka and H. Umezawa, *J. Antibiot.*, 1968, **21**, 379.
- C. Vézina, A. Kudelski and S. N. Sehgal, *J. Antibiot.*, 1975, **28**, 721.
- R. H. Lambalot, A. M. Gehring, R. S. Flugel, P. Zuber, M. LaCelle, M. A. Marahiel, R. Reid, C. Khosla and C. T. Walsh, *Chem. Biol.*, 1996, **3**, 923.
- J. Beld, E. C. Sonnenschein, C. R. Vickery, J. P. Noel and M. D. Burkart, *Nat. Prod. Rep.*, 2014, **31**, 61.
- L. E. Quadri, P. H. Weinreb, M. Lei, M. M. Nakano, P. Zuber and C. T. Walsh, *Biochemistry*, 1998, **37**, 1585.
- P. J. Belshaw, C. T. Walsh and T. Stachelhaus, *Science*, 1999, **284**, 486–489.
- J. J. La Clair, T. L. Foley, T. R. Schegg, C. M. Regan and M. D. Burkart, *Chem. Biol.*, 2004, **11**, 195.
- S. A. Sieber, C. T. Walsh and M. A. Marahiel, *J. Am. Chem. Soc.*, 2003, **125**, 10862.
- M. Sunbul, N. J. Marshall, Y. Zou, K. Zhang and J. Yin, *J. Mol. Biol.*, 2009, **387**, 883.
- Y. Zou and J. Yin, *J. Am. Chem. Soc.*, 2009, **131**, 7548.
- A. M. Gehring, R. H. Lambalot, K. W. Vogel, D. G. Drueckhammer and C. T. Walsh, *Chem. Biol.*, 1997, **4**, 17.
- R. Finking, J. Solsbacher, D. Konz, M. Schobert, A. Schäfer, D. Jahn and M. A. Marahiel, *J. Biol. Chem.*, 2002, **277**, 50293.
- M. Pavlidou, E. K. Pross, E. M. Musiol, A. Kulik, W. Wohlleben and T. Weber, *FEMS Microbiol. Lett.*, 2011, **319**, 26.
- M. R. Mofid, R. Finking and M. A. Marahiel, *J. Biol. Chem.*, 2002, **277**, 17023.
- A. Geerlof, A. Lewendon and W. V. Shaw, *J. Biol. Chem.*, 1999, **274**, 27105.
- T. Kupke, M. Uebele, D. Schmid, G. Jung, M. Blaesse and S. Steinbacher, *J. Biol. Chem.*, 2000, **275**, 31838.
- E. Strauss, C. Kinsland, Y. Ge, F. W. McLafferty and T. P. Begley, *J. Biol. Chem.*, 2001, **276**, 13513.
- T. P. Begley, C. Kinsland and E. Strauss, *Vitam. Horm.*, 2001, **61**, 157.
- E. Strauss, M. de Villiers and I. Rootman, *ChemCatChem*, 2010, **2**, 929.
- D. P. Martin, R. T. Bibart and D. G. Drueckhammer, *J. Am. Chem. Soc.*, 1994, **116**, 4660.
- E. Strauss and T. P. Begley, *J. Biol. Chem.*, 2002, **277**, 48205.
- I. Nazi, K. P. Koteva and G. D. Wright, *Anal. Biochem.*, 2004, **324**, 100.
- M. van Wyk and E. Strauss, *Chem. Commun.*, 2007, 398.
- M. Tosin, D. Spitteller and J. B. Spencer, *ChemBioChem*, 2009, **10**, 1714.
- I. Rootman, M. de Villiers, L. A. Brand and E. Strauss, *ChemCatChem*, 2010, **2**, 1239.
- R. B. Hamed, L. Henry, J. R. Gomez-Castellanos, A. Asghar, J. Brem, T. D. Claridge and C. J. Schofield, *Org. Biomol. Chem.*, 2013, **11**, 8191.
- J. L. Meier, A. C. Mercer, H. Rivera and M. D. Burkart, *J. Am. Chem. Soc.*, 2006, **128**, 12174.
- M. Rawlings and J. E. Cronan, *J. Biol. Chem.*, 1992, **267**, 5751.
- R. F. Seipke, J. Barke, C. Brearley, L. Hill, D. W. Yu, R. J. M. Goss and M. I. Hutchings, *PLoS One*, 2011, **6**, e22028.

- 31 V. Agarwal, S. Diethelm, L. Ray, N. Garg, T. Awakawa, P. C. Dorrestein and B. S. Moore, *Org. Lett.*, 2015, **17**, 4452–4455.
- 32 L. Han, K. Yang, E. Ramalingam, R. H. Mosher and L. C. Vining, *Microbiology*, 1994, **140**, 3379.
- 33 L. Wang, J. McVey and L. C. Vining, *Microbiology*, 2001, **147**, 1535.
- 34 T. Weber, K. J. Laiple, E. K. Pross, A. Textor, S. Grond, K. Welzel, S. Pelzer, A. Vente and W. Wohlleben, *Chem. Biol.*, 2008, **15**, 175.
- 35 P. E. Coderre and C. F. Earhart, *J. Gen. Microbiol.*, 1989, **135**, 3043.
- 36 A. M. Gehring, K. A. Bradley and C. T. Walsh, *Biochemistry*, 1997, **36**, 8495.

## Appendix B

### **Substrate Specificity of a Ketosynthase Domain Involved in Bacillaene Biosynthesis**

*Beilstein J. Org. Chem.* **2024**, *5*, 734–740.

[doi.org/10.3762/bjoc.20.67](https://doi.org/10.3762/bjoc.20.67)



# Substrate specificity of a ketosynthase domain involved in bacillaene biosynthesis

Zhiyong Yin and Jeroen S. Dickschat\*

## Letter

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:  
bacillaene; biosynthesis; enzyme mechanisms; isotopes; *trans*-AT  
polyketide synthases

*Beilstein J. Org. Chem.* **2024**, *20*, 734–740.  
<https://doi.org/10.3762/bjoc.20.67>

Received: 15 December 2023  
Accepted: 27 March 2024  
Published: 05 April 2024

Associate Editor: S. Bräse



© 2024 Yin and Dickschat; licensee Beilstein-Institut.  
License and terms: see end of document.

## Abstract

An isotopic labelling method was developed to investigate substrate binding by ketosynthases, exemplified by the second ketosynthase of the polyketide synthase BaeJ involved in bacillaene biosynthesis (BaeJ-KS2). For this purpose, both enantiomers of a  $^{13}\text{C}$ -labelled *N*-acetylcysteamine thioester (SNAC ester) surrogate of the proposed natural intermediate of BaeJ-KS2 were synthesised, including an enzymatic step with glutamate decarboxylase, and incubated with BaeJ-KS2. Substrate binding was demonstrated through  $^{13}\text{C}$  NMR analysis of the products against the background of various control experiments.

## Introduction

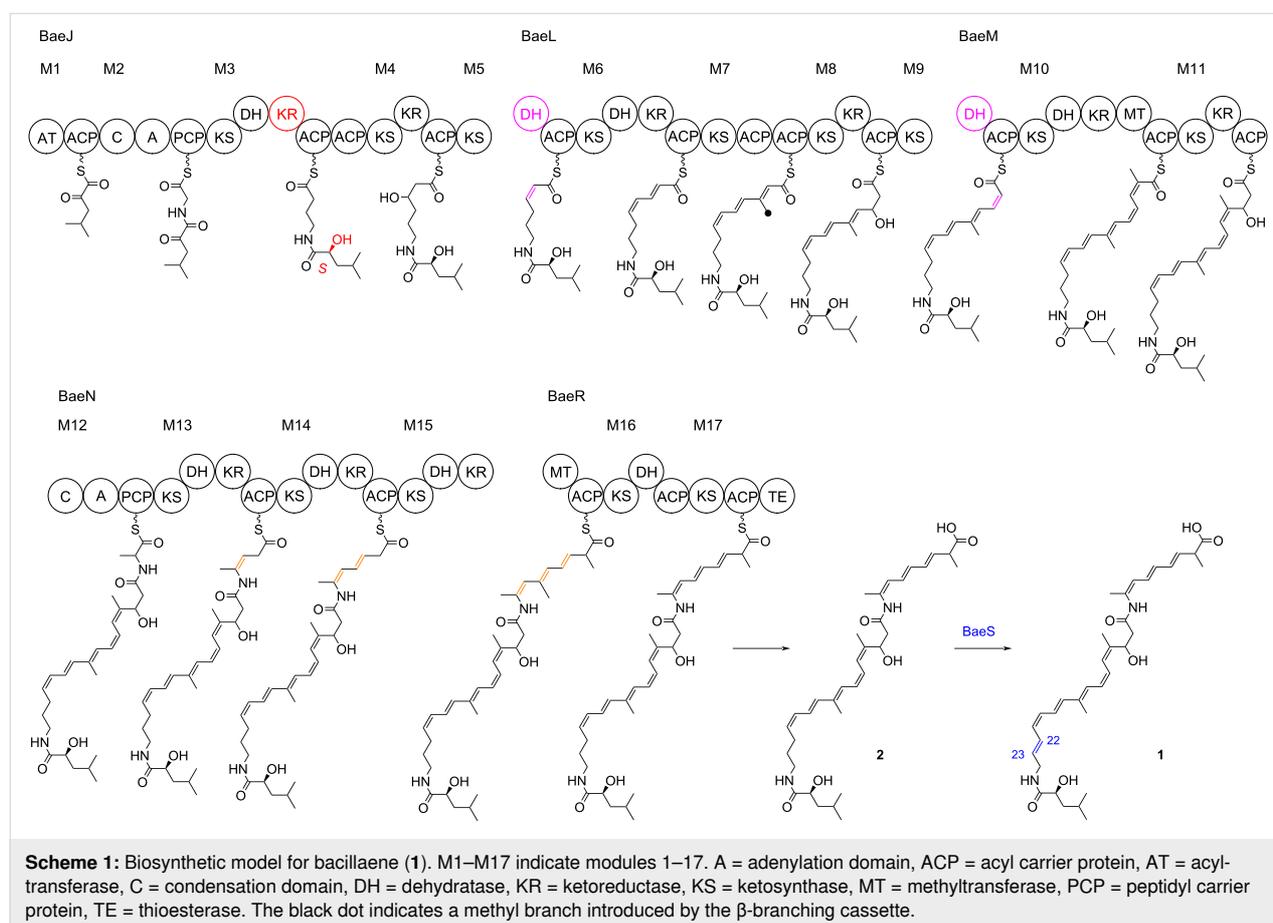
Polyketides are a large class of natural products which often exhibit potent biological activities for their application in medicine, e.g., as antibiotics or immunosuppressants [1]. Despite their high structural variability all compounds from this class are commonly made through the action of polyketide synthases (PKS). The type I of these enzymes are megasynthases composed of several catalytically active domains that can either act iteratively with the same set of domains catalysing the incorporation of several extender units into a growing polyketide chain, or non-iteratively with one set of domains acting only for the incorporation of one extender unit [2,3]. Although enzyme domains with various specialised catalytic functions can be

found as integral part of polyketide synthases, three domain types are fundamental to their biosynthesis, resembling the same logic as observed for fatty acid biosynthesis: the acyl transferases (AT) for loading of the starter or extender units, the acyl carrier proteins (ACP) for anchoring the growing polyketide chain, and the ketosynthases (KS) for merging of the next extender unit with the existing chain by a decarboxylative Claisen condensation [2,4]. Today a high understanding of polyketide biosynthesis has been reached, including a detailed knowledge of the extender unit selection and the stereochemical implications that are predictable from amino acid sequences [5,6].

The polyene antibiotic bacillaene was first isolated from *Bacillus subtilis* [7]. This soil-dwelling organism shares the same habitat with the predator *Myxococcus xanthus* that feeds on other bacteria including *B. subtilis*, and bacillaene is the primary factor conferring *B. subtilis* cells resistance to predation by *M. xanthus* [8]. The identification of its biosynthetic gene cluster (*bae*) revealed that the compound is made through a *trans*-AT polyketide synthase–non-ribosomal peptide synthase (PKS-NRPS) hybrid [9,10]. Instead of using the classical domain organisation KS-AT-ACP with AT domains integrated into the PKS, *trans*-AT PKSs utilize discrete ATs that are not an integral part of the PKS, but rather cooperate with the PKS “*in trans*” [11,12]. Notably, in *B. subtilis* the giant bacillaene biosynthesis machinery forms an organelle-like complex that can be observed through cryoelectron microscopy [13]. The structure elucidation of “bacillaene” through extensive NMR spectroscopic methods revealed the presence of two major compounds, bacillaene (**1**) and dihydrobacillaene (**2**) (Scheme 1), and allowed for a detailed biosynthetic model [14]. As is typical for *trans*-AT systems, the bacillaene PKS-NRPS contains several irregular features such as split modules and duplicate ACP domains. Because of the absence of a ketoreductase (KR) domain in module 1, the starter unit was initially suggested to

be  $\alpha$ -hydroxyisocaproate [14], but later it was shown that the KR of module 3 acts twice, in the reduction of the  $\beta$ -ketoacyl intermediate of the elongation step of module 3 and in the reduction of the  $\alpha$ -ketoisocaproate starter unit with introduction of an *S* configured stereocentre (highlighted in red in Scheme 1) [15]. The domain organisation of module 3 containing no enoylreductase (ER) domain furthermore suggests the formation of an  $\alpha,\beta$ -unsaturated intermediate, and not a full reduction at this stage, in agreement with the presence of a double bond between C22 and C23 in **1**. A contrasting picture was obtained through deletion of the TE domain that resulted in off-loading of all premature intermediates from the PKS [16], possibly catalysed by a proofreading AT-like enzyme encoded in the *bae* cluster [17]. These intermediates consistently showed masses two Da higher than expected, in agreement with a fully saturated intermediate at module 3 [16]. Subsequent investigations demonstrated that the C22=C23 double bond in **1** is incorporated in a post-PKS step through the action of BaeS (highlighted in blue) [18].

Further interesting features are the presence of a  $\beta$ -branching cassette for the installation of a  $\beta$ -methyl group at the stage of the module 7 intermediate (highlighted by the black dot)



[14,16], and split modules between BaeJ and BaeL (module 5), as well as between BaeL and BaeM (module 9). These modules are suggested not to catalyse elongations, but the KS domains of modules 5 and 9 may only act in the translocation of the intermediate from the ACP of the previous module to the ACP of modules 5 and 9, respectively. Modules 5 and 9 then only catalyse the dehydration of the alcohol functions installed by the preceding modules (highlighted in purple) [14,16]. Furthermore, the structures of **1** and **2** show a shifted triene portion that is not in conjugation with the carboxylic acid function. NMR studies of off-loaded intermediates with the TE deletion mutant revealed that these double bond shifts are introduced during the elongation steps of modules 13–15, and not after complete assembly of the PKS backbone [19].

A phylogenetic analysis of KS domains from *trans*-AT PKSs revealed that these domains group together, if the structures of their processed substrates are similar, allowing for a prediction of the substrate specificity of such KS domains [20]. Especially the KS domains of a PKS module downstream of an NRPS module need to be able to process unconventional NRPS-derived intermediates [21]. Despite the above mentioned predictability of KS domain substrate specificities, functional testing of KS domains is still important. Previous approaches to determine KS domain specificities have involved mass spectrometry (MS)-based methods [22,23], MS analysis of trypsin-digested proteins [24], and radiochemical assays [25]. Here, we report on a new method using <sup>13</sup>C-labelled substrate surrogates in conjunction with <sup>13</sup>C NMR to investigate substrate acceptance by BaeJ-KS2 from module 4 of the bacillaene PKS.

## Results and Discussion

### Synthesis of *N*-acetylcysteamine thioesters

Previous studies have shown that *N*-acetylcysteamine thioesters (SNAC esters) can be uploaded to KS domains [22–25]. Therefore, to investigate the function of the KS domain BaeJ-KS2 the synthesis of <sup>13</sup>C-labelled (*S*)-**11** as a mimic of the intermediate bound to the ACP of module 3 was performed. It was planned to introduce the <sup>13</sup>C-labelling from (5-<sup>13</sup>C)glutamate into the  $\gamma$ -aminobutyrate portion of (*S*)-**11**. For this purpose, the gene coding for the glutamate decarboxylase from *Escherichia coli* K12 (accession no. AAA23833) was cloned through homologous recombination in yeast into the expression vector pYE-Express [26]. Heterologous expression and purification of the recombinant enzyme (Figure S1, Supporting Information File 1) allowed for test incubations with unlabelled glutamic acid in the presence of the cofactor pyridoxal phosphate (PLP), showing the complete conversion of glutamic acid (**3**) into  $\gamma$ -aminobutyric acid (**4**) (Scheme 2). The crude product was used without purification for a Schotten–Baumann esterification with benzyl alcohol to obtain the unlabelled ester **5** with a quantita-

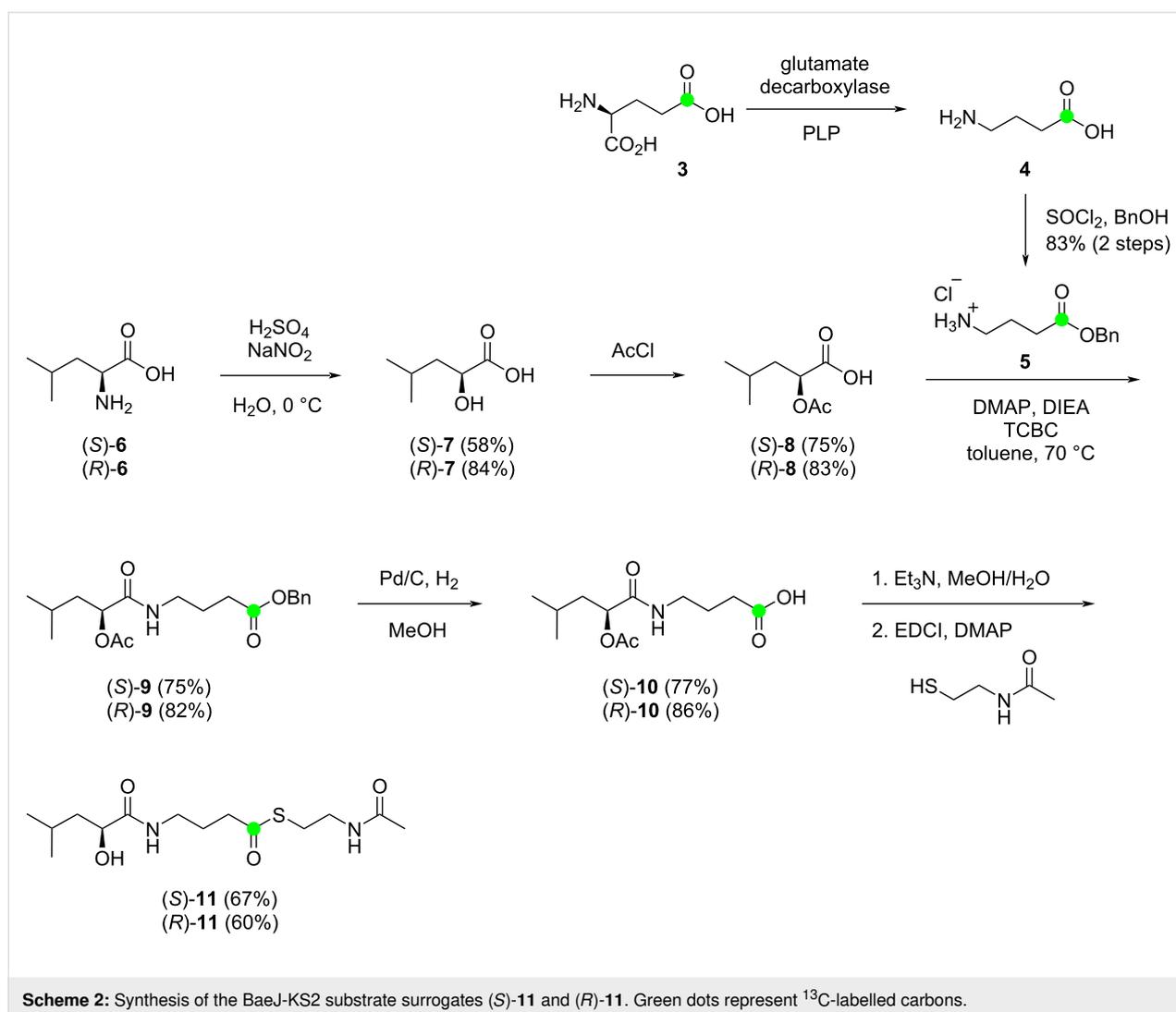
tive yield over two steps. After having established this method, (5-<sup>13</sup>C)glutamic acid (**3**) was converted analogously into (1-<sup>13</sup>C)-**5**, unfortunately with a little lower, but still very good yield of 83%.

By employing Yamaguchi conditions, the ester **5** was coupled with the carboxylic acid (*S*)-**8**, derived from *L*-leucine ((*S*)-**6**) via hydroxyacid (*S*)-**7**, to yield the amide (*S*)-**9**. Deprotection through catalytic hydrogenation to (*S*)-**10**, saponification of the acetate ester and Steglich esterification with *N*-acetylcysteamine gave access to the desired SNAC ester (*S*)-**11**, with a yield of 32% over 5 steps starting from <sup>13</sup>C-labelled **3**. Analogous reactions were performed to obtain (*R*)-**11** from *D*-leucine ((*R*)-**6**) with a comparable overall yield of 35% from labelled **3**.

### BaeJ-KS2 activity assay

The substrate specificities of ketosynthases in *trans*-AT PKSs hold significant importance. Many KSs within *trans*-AT PKSs function as gatekeepers, facilitating the transfer of intermediates along the assembly line without participating in chain elongation. For instance, the final KS of the bacillaene PKS lacks condensation capability, but may serve to ensure proper double bond isomerisation in the late polyketide intermediates before they are passed on for their TE-mediated release from the PKS. Such gatekeeping roles have also been discussed for condensation-competent KSs [27] that may process the substrate for the next round of elongation only after installation of the correct functional groups through optional reductive loop or other modifying domains. Located in module 4, BaeJ-KS2 was in the focus of our investigation to assess its response to both enantiomers of the full-length substrate surrogate **11**.

To investigate the function of BaeJ-KS2 the recombinant His-tagged enzyme was obtained through heterologous expression and purification by Ni<sup>2+</sup>-NTA affinity chromatography (Figure S1, Supporting Information File 1). After incubation of the purified enzyme with the <sup>13</sup>C-labelled SNAC derivatives (*S*)-**11** or (*R*)-**11** for 30 minutes at 25 °C, the incubation buffer of the reaction mixture potentially containing the free substrate surrogates was exchanged through repeated centrifugation using an ultrafiltration centrifugal tube (3 kDa cut-off), followed by the addition of incubation buffer. In total, through this method five successive  $\approx$ 10:1 dilutions were performed to eliminate any unreacted free **11** in the reaction mixture. The resulting protein preparations were subsequently analysed by <sup>13</sup>C NMR spectroscopy. While the signal for the thioester carbonyl group of free **11** dissolved in incubation buffer was observed at  $\delta$  = 203.33 ppm (Figure 1A), for both samples obtained from the incubations with (*S*)-**11** and (*R*)-**11** signals were detected with a small, but reproducible difference of the chemical shift at



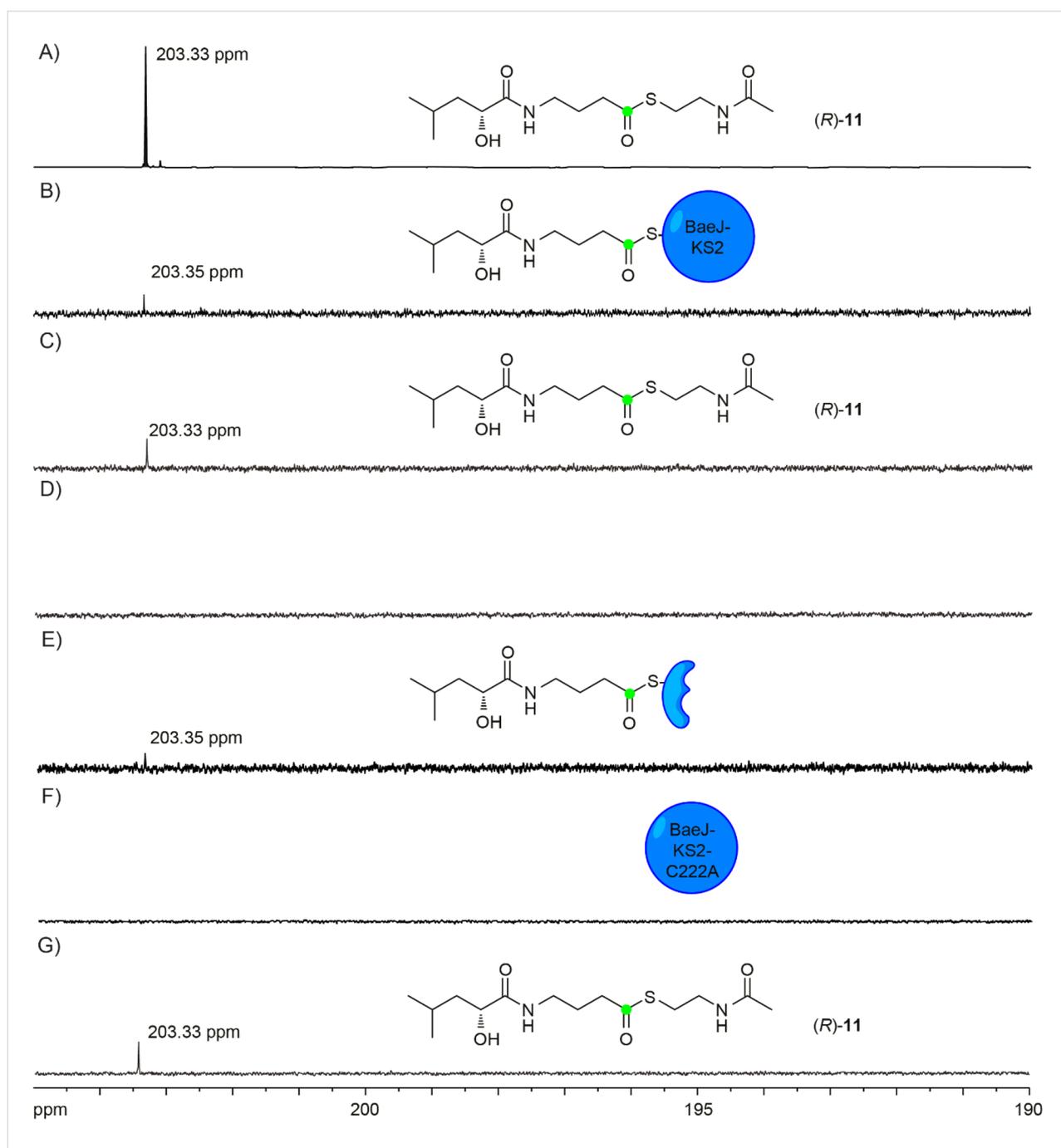
$\delta = 203.35$  ppm (Figure 1B), likely representing enzyme bound **11**. Furthermore, the filtrates obtained from the first and the fifth centrifugation were analysed by <sup>13</sup>C NMR, showing the presence of free **11** after the first centrifugation step (Figure 1C), but not after the last round of centrifugation (Figure 1D).

Protein binding of the substrate surrogates **11** was confirmed through digestion of BaeJ-KS2 using protease K after buffer exchange. The digested sample was subsequently subjected to another round of centrifugation using an ultrafiltration centrifugal tube, resulting in the detection of a signal for a thioester carbonyl group in the filtrate for both samples derived from (*S*)- and (*R*)-**11** at  $\delta = 203.35$  ppm (Figure 1E). This observation supports binding of both substrate surrogates to BaeJ-KS2, but it is unclear from these experiments, if **11** is bound covalently to the protein or through non-covalent interactions.

To gain further evidence for the covalent binding of **11** to BaeJ-KS2, the highly conserved Cys residue involved in substrate attachment [28,29] was exchanged through site-directed mutagenesis, resulting in the BaeJ-KS2-C222A enzyme variant. After heterologous expression and protein purification (Figure S1, Supporting Information File 1), the same protocol for the incubation with (*S*)-**11** and (*R*)-**11** followed by exchange of the incubation buffer as described above was applied. No signal corresponding to a thioester carbonyl group was detected in the protein preparations (Figure 1F), while signals at  $\delta = 203.33$  ppm for free **11** were observed in the filtrates obtained after the first step (Figure 1G).

## Conclusion

Taken together, we have established a new method based on stable isotope (<sup>13</sup>C) labelling to investigate the KS domain substrate specificity that makes use of simple <sup>13</sup>C NMR analysis of protein preparations obtained by a buffer exchange after en-



**Figure 1:**  $^{13}\text{C}$  NMR spectra of  $(R)$ -11 incubated with BaeJ-KS2 and BaeJ-KS2-C222A. A) Free 11 dissolved in incubation buffer; B)  $(R)$ -11 bound to BaeJ-KS2 after incubation and buffer exchange (5 centrifugations); C) the filtrate obtained from the incubation of  $(R)$ -11 with BaeJ-KS2 containing free  $(R)$ -11 (first centrifugation); D) the filtrate from the same experiment containing no  $(R)$ -11 (fifth centrifugation); E) the filtrate obtained from the incubation of  $(R)$ -11 with BaeJ-KS2 followed by buffer exchange and then digestion with proteinase K; F)  $(R)$ -11 is not bound to BaeJ-KS2-C222A after incubation and buffer exchange (5 centrifugations); G) the filtrate obtained from the incubation of  $(R)$ -11 with BaeJ-KS2-C222A containing free  $(R)$ -11 (first centrifugation). Green dots represent  $^{13}\text{C}$ -labelled carbons.

zyme incubations with substrate surrogates (SNAC esters), and its application to investigate BaeJ-KS2 from the bacillaene PKS. While this enzyme has been investigated for its substrate scope before [27], the present study is the first that investigates the acceptance of a substrate surrogate representing the natural

module 3 intermediate. The obtained results confirm the structure of this intermediate as  $N$ -(( $S$ )- $\alpha$ -hydroxyisocaproyl)- $\gamma$ -aminobutanoyl-S-ACP and also demonstrate that BaeJ-KS2 is not sensitive towards an inversion of the configuration in the  $\alpha$ -hydroxyisocaproate moiety. This finding is in agreement with

the phylogenetic analysis of KS domains from *trans*-AT PKSs, showing that the amino acid sequences of these KSs correlate with the structures of the processed PKS intermediates up to the  $\beta$ -carbon [20].

## Supporting Information

### Supporting Information File 1

Experimental part and NMR spectra.

[<https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-20-67-S1.pdf>]

## Acknowledgements

We thank Jörn Piel (Zürich) for generously providing the strain *Bacillus amyloliquefaciens* FZB42 to us.

## Funding

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, TRR 261, project ID 398967434).

## ORCID® iDs

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

## Data Availability Statement

All data that supports the findings of this study is available in the published article and/or the supporting information to this article.

## References

- Koehn, F. E. *MedChemComm* **2012**, *3*, 854–865. doi:10.1039/c2md00316c
- Staunton, J.; Weissman, K. J. *Nat. Prod. Rep.* **2001**, *18*, 380–416. doi:10.1039/a909079g
- Hertweck, C. *Angew. Chem., Int. Ed.* **2009**, *48*, 4688–4716. doi:10.1002/anie.200806121
- Cortes, J.; Haydock, S. F.; Roberts, G. A.; Bevitt, D. J.; Leadlay, P. F. *Nature* **1990**, *348*, 176–178. doi:10.1038/348176a0
- Weissman, K. J. *Beilstein J. Org. Chem.* **2017**, *13*, 348–371. doi:10.3762/bjoc.13.39
- Yin, Z.; Dickschat, J. S. *Nat. Prod. Rep.* **2021**, *38*, 1445–1468. doi:10.1039/d0np00091d
- Patel, P. S.; Huang, S.; Fisher, S.; Pirnik, D.; Aklonis, C.; Dean, L.; Meyers, E.; Fernandes, P.; Mayerl, F. *J. Antibiot.* **1995**, *48*, 997–1003. doi:10.7164/antibiotics.48.997
- Müller, S.; Strack, S. N.; Hoefler, B. C.; Straight, P. D.; Kearns, D. B.; Kirby, J. R. *Appl. Environ. Microbiol.* **2014**, *80*, 5603–5610. doi:10.1128/aem.01621-14
- Scotti, C.; Piatti, M.; Cuzzoni, A.; Perani, P.; Tognoni, A.; Grandi, G.; Galizzi, A.; Albertini, A. M. *Gene* **1993**, *130*, 65–71. doi:10.1016/0378-1119(93)90347-6
- Chen, X.-H.; Vater, J.; Piel, J.; Franke, P.; Scholz, R.; Schneider, K.; Koumoutsis, A.; Hitzeroth, G.; Grammel, N.; Strittmatter, A. W.; Gottschalk, G.; Süßmuth, R. D.; Borriss, R. *J. Bacteriol.* **2006**, *188*, 4024–4036. doi:10.1128/jb.00052-06
- Piel, J. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 14002–14007. doi:10.1073/pnas.222481399
- Helfrich, E. J. N.; Piel, J. *Nat. Prod. Rep.* **2016**, *33*, 231–316. doi:10.1039/c5np00125k
- Straight, P. D.; Fischbach, M. A.; Walsh, C. T.; Rudner, D. Z.; Kolter, R. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 305–310. doi:10.1073/pnas.0609073103
- Butcher, R. A.; Schroeder, F. C.; Fischbach, M. A.; Straight, P. D.; Kolter, R.; Walsh, C. T.; Clardy, J. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 1506–1509. doi:10.1073/pnas.0610503104
- Calderone, C. T.; Bumpus, S. B.; Kelleher, N. L.; Walsh, C. T.; Magarvey, N. A. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 12809–12814. doi:10.1073/pnas.0806305105
- Moldenhauer, J.; Chen, X.-H.; Borriss, R.; Piel, J. *Angew. Chem., Int. Ed.* **2007**, *46*, 8195–8197. doi:10.1002/anie.200703386
- Jensen, K.; Niederkrüger, H.; Zimmermann, K.; Vagstad, A. L.; Moldenhauer, J.; Brendel, N.; Frank, S.; Pöplau, P.; Kohlhaas, C.; Townsend, C. A.; Oldiges, M.; Hertweck, C.; Piel, J. *Chem. Biol.* **2012**, *19*, 329–339. doi:10.1016/j.chembiol.2012.01.005
- Reddick, J. J.; Antolak, S. A.; Raner, G. M. *Biochem. Biophys. Res. Commun.* **2007**, *358*, 363–367. doi:10.1016/j.bbrc.2007.04.151
- Moldenhauer, J.; Götz, D. C. G.; Albert, C. R.; Bischof, S. K.; Schneider, K.; Süßmuth, R. D.; Engeser, M.; Gross, H.; Bringmann, G.; Piel, J. *Angew. Chem., Int. Ed.* **2010**, *49*, 1465–1467. doi:10.1002/anie.200905468
- Nguyen, T.; Ishida, K.; Jenke-Kodama, H.; Dittmann, E.; Gurgui, C.; Hochmuth, T.; Taudien, S.; Platzer, M.; Hertweck, C.; Piel, J. *Nat. Biotechnol.* **2008**, *26*, 225–233. doi:10.1038/nbt1379
- Du, L.; Sánchez, C.; Shen, B. *Metab. Eng.* **2001**, *3*, 78–95. doi:10.1006/mben.2000.0171
- Jenner, M.; Frank, S.; Kampa, A.; Kohlhaas, C.; Pöplau, P.; Briggs, G. S.; Piel, J.; Oldham, N. J. *Angew. Chem., Int. Ed.* **2013**, *52*, 1143–1147. doi:10.1002/anie.201207690
- Kohlhaas, C.; Jenner, M.; Kampa, A.; Briggs, G. S.; Afonso, J. P.; Piel, J.; Oldham, N. J. *Chem. Sci.* **2013**, *4*, 3212–3217. doi:10.1039/c3sc50540e
- Schnarr, N. A.; Chen, A. Y.; Cane, D. E.; Khosla, C. *Biochemistry* **2005**, *44*, 11836–11842. doi:10.1021/bi0510781
- Wu, J.; Kinoshita, K.; Khosla, C.; Cane, D. E. *Biochemistry* **2004**, *43*, 16301–16310. doi:10.1021/bi048147g
- Dickschat, J. S.; Pahirulzaman, K. A. K.; Rabe, P.; Klapschinski, T. A. *ChemBioChem* **2014**, *15*, 810–814. doi:10.1002/cbic.201300763
- Gay, D. C.; Gay, G.; Axelrod, A. J.; Jenner, M.; Kohlhaas, C.; Kampa, A.; Oldham, N. J.; Piel, J.; Keatinge-Clay, A. T. *Structure* **2014**, *22*, 444–451. doi:10.1016/j.str.2013.12.016
- Olsen, J. G.; Kadziola, A.; von Wettstein-Knowles, P.; Siggaard-Andersen, M.; Larsen, S. *Structure* **2001**, *9*, 233–243. doi:10.1016/s0969-2126(01)00583-4
- von Wettstein-Knowles, P.; Olsen, J. G.; McGuire, K. A.; Henriksen, A. *FEBS J.* **2006**, *273*, 695–710. doi:10.1111/j.1742-4658.2005.05101.x

## 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.20.67>

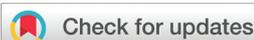
## Appendix C

### **An Isotopic Probe to Follow the Stereochemical Course of Dehydratase Reactions in Polyketide and Fatty Acid Biosynthesis**

*Org. Chem. Front.* **2022**, *9*, 2714–2720.

[doi.org/10.1039/D2QO00272H](https://doi.org/10.1039/D2QO00272H)

## RESEARCH ARTICLE

View Article Online  
View Journal | View IssueCite this: *Org. Chem. Front.*, 2022, 9, 2714

# An isotopic probe to follow the stereochemical course of dehydratase reactions in polyketide and fatty acid biosynthesis†

Zhiyong Yin,<sup>a</sup> Elisa Liebhart,<sup>b</sup> Evi Stegmann,<sup>b</sup> Heike Brötz-Oesterhelt <sup>b</sup> and Jeroen S. Dickschat \*<sup>a</sup>Received 18th February 2022,  
Accepted 7th April 2022

DOI: 10.1039/d2qo00272h

rsc.li/frontiers-organic

Four stereoisomeric and isotopically labelled probes that are suitable to easily follow the stereochemical course of dehydratases have been synthesised. The synthesis made use of an epoxide opening through catalytic hydrogenation that was shown to proceed with inversion of configuration. The obtained labelled probes were applied to study the stereochemical courses of several dehydratases involved in the biosynthesis of selected antibiotics and of fatty acids.

## Introduction

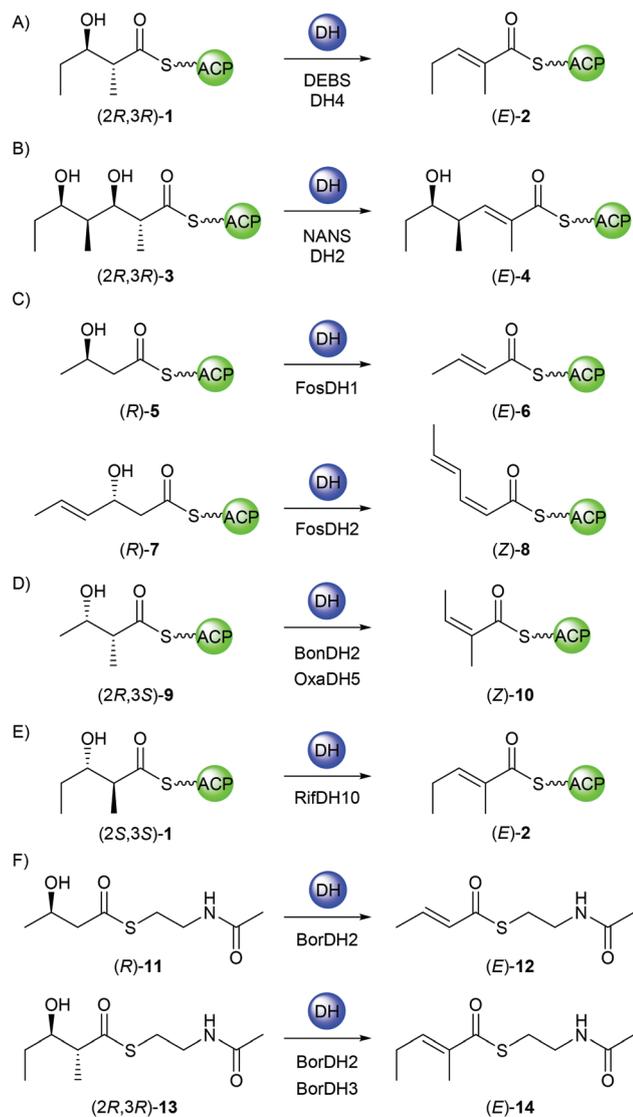
Polyketide biosynthesis<sup>1</sup> by a modular type I polyketide synthase (PKS) is a process in which small activated carboxylic acids (usually acetyl-CoA or propionyl-CoA) are fused in a biological Claisen condensation. Herein, one module (set of domains) is responsible for the incorporation of usually one biosynthetic building block. The elongation units are first carboxylated to malonyl-CoA or methylmalonyl-CoA to further increase their reactivity and to make the Claisen condensation irreversible by release of the thermodynamically stable byproduct carbon dioxide, which drives the multistep assembly of a polyketide forward. This all happens in an enzyme-bound manner with attachment of the growing polyketide chain to a coenzyme A-derived phosphopantetheinyl (PPant) linker that is itself bound to an acyl carrier protein (ACP). Upload of the building blocks and their condensation is catalysed by acyl transferase (AT) and ketosynthase (KS) domains, respectively. Optional processing steps of the resulting  $\beta$ -ketoacyl intermediates include reduction by a ketoreductase (KR), elimination of water by a dehydratase (DH) and again reduction by an enoyl reductase (ER), before the polyketide chain is transferred to the next module. The stereochemical course of the KR domain can be predicted from signature sequence motifs,

grouping KR domains into A type (yielding 3L configured alcohols) or B type (3D alcohols).<sup>2</sup> In analogy to the reactions of fatty acid synthases that exclusively yield *E*-configured double bonds through *syn* dehydration of 3D alcohols,<sup>3,4</sup> the subsequently acting DH domain is then believed to generally catalyse a *syn*-elimination of water<sup>5–7</sup> with stereospecific removal of the 2-*pro-S* proton.<sup>8</sup> As suggested by Reid *et al.*, the KR domain stereospecificity has thus consequences for the dehydration step:<sup>9</sup> B type (3D) alcohols will be converted into products with *E* configured double bond as found in most polyketides, while the elimination of water from A type (3L) alcohols leads to products with a *Z* double bond, as is also occasionally observed.<sup>10</sup> This model is supported by a correlation of the structures of many polyketides with C=C double bonds to the observed KR signature motifs. Dehydratases have been extensively studied for various aspects,<sup>11–15</sup> but only little experimental work has been conducted to follow the stereochemical course of PKS DH domains experimentally. For some systems direct insights were obtained by *in vitro* incubations: DEBS DH4 from erythromycin biosynthesis was shown to convert the 3-hydroxyacyl-SACP (2*R*,3*R*)-**1** by *syn* dehydration into (*E*)-**2**, but does not accept other stereoisomers of **1** (Scheme 1A).<sup>6</sup> Similarly, NANS DH2 from the nanchangmycin PKS catalyses *syn* dehydration of the more complex substrate (2*R*,3*R*)-**3** to yield (*E*)-**4** (Scheme 1B).<sup>7</sup> Also two dehydratases from the fostriecin PKS were investigated. FosDH1 catalyses the dehydration of (*R*)-**5** to (*E*)-**6**, while the *S* enantiomer is not accepted. In contrast, FosDH2 converts (*R*)-**7** into (*Z*)-**8**, explaining *Z* double bond formation in fostriecin biosynthesis (Scheme 1C).<sup>16</sup> For both enzymes, the stereoselectivity at *C*<sub>α</sub> has not been investigated and it remains unknown, whether the reactions represent *syn* eliminations. After characterisation of this first *Z* double bond forming DH, also BonDH2 from bongkrecic acid biosynthesis

<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>Interfaculty Institute of Microbiology and Infection Medicine, Department of Microbial Bioactive Compounds, Auf der Morgenstelle 28, 72076 Tübingen, Germany. E-mail: heike.broetz-oesterhelt@uni-tuebingen.de

†Electronic supplementary information (ESI) available: Experimental procedures, spectroscopic data, amino acid sequences of proteins and results from labelling experiments. See DOI: <https://doi.org/10.1039/d2qo00272h>



**Scheme 1** Previous *in vitro* experiments with dehydratases from different PKSs.

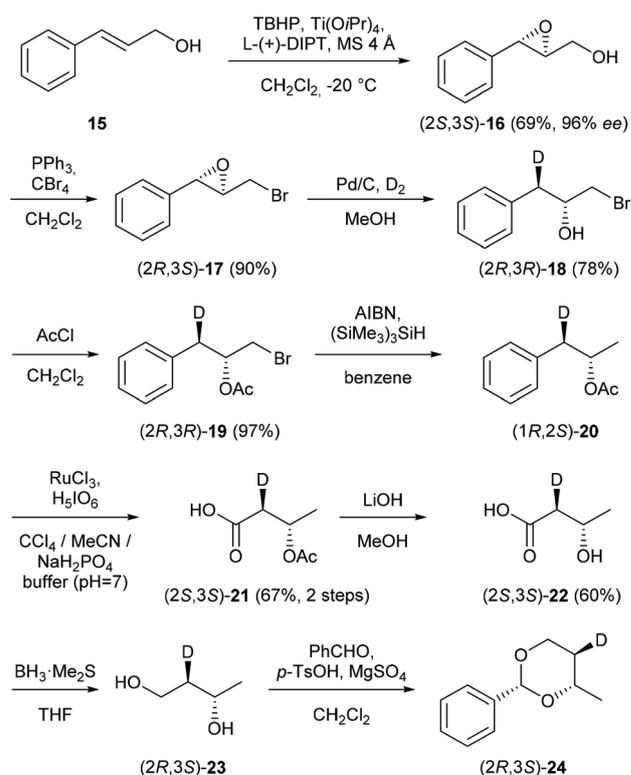
and OxaDH5 from the oxazolomycin PKS were shown to catalyse *Z* double bond installation by *syn* dehydration of (2*R*,3*S*)-9 to yield (*Z*)-10 (Scheme 1D).<sup>17</sup> RifDH10 from the rifamycin PKS also catalyses *syn* elimination of water, but is unusual, because it acts on a 3L alcohol and removes the 2-*pro-R* proton in the conversion of (2*S*,3*S*)-1 into (*E*)-2 (Scheme 1E).<sup>5</sup> While all these experiments have been performed on ACP-bound substrates, Leadlay and coworkers have shown that the more easily accessible *N*-acetylcysteamine thioesters (SNAC esters) can also be converted by DHs. BorDH2, but not BorDH3 from the borrelidin PKS catalyses the dehydration of (*R*)-11 to (*E*)-12, and both enzymes can convert (2*R*,3*R*)-13 into (*E*)-14, while in all cases no distinct activity was observed with other stereoisomers of 11 and 13 (Scheme 1F).<sup>18</sup> Also in these cases no detailed picture of the stereochemical course of the reaction is available. Here we report on the synthesis of isotopically labelled

probes that allow to efficiently follow the stereochemical course of dehydrations by <sup>13</sup>C-NMR and their application on various DHs from polyketide and fatty acid biosynthesis.

## Results and discussion

### Synthesis of enantioselectively deuterated probes

The stereochemical course of dehydratases involves two aspects: (1) the stereoselectivity for (*R*)- or (*S*)-3-hydroxyacids, and 2. the selectivity of the deprotonation at *C<sub>α</sub>* regarding loss of the 2-*pro-R* or 2-*pro-S* proton. Both aspects can be investigated by using a set of all four stereoisomers of (2-<sup>2</sup>H)-3-hydroxybutyric acid, activated as *N*-acetylcysteamine thioesters (SNAC esters) that represent surrogates of ACP thioesters. For this purpose, first a route was established for the stereoselective synthesis of (2-<sup>2</sup>H)-3-hydroxybutyric acid (Scheme 2). Cinnamic alcohol (15) was converted into (2*S*,3*S*)-16 by Sharpless epoxidation with *L*-(+)-DIPT,<sup>19</sup> yielding the epoxide with high enantiomeric purity as determined by conversion into the Mosher ester with (*S*)-Mosher acid chloride and NMR spectroscopic analysis (Fig. S7†).<sup>20</sup> After conversion of the alcohol into the bromide (17) by Appel reaction, the epoxides were reductively opened with D<sub>2</sub> and Pd/C to yield 18 regioselectively.<sup>21</sup> The NMR spectra showed that deuterium incorporation proceeded selectively into one of the diastereotopic posi-



**Scheme 2** Synthesis of enantioselectively deuterated 3-hydroxybutyric acid and further conversion into dioxane 24 to study the stereochemical course of the epoxide opening step.

tions at C3, but at this stage it was unclear, which stereoisomer was obtained.

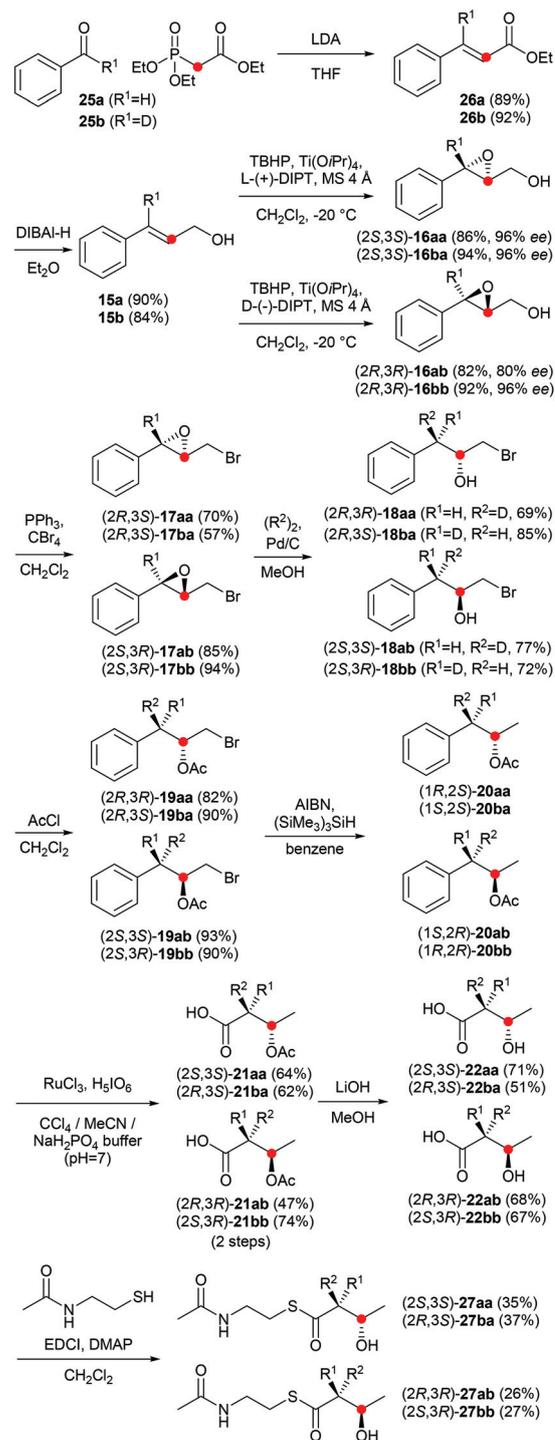
Compound **18** was then further transformed by acetylation into **19**, followed by a radical dehalogenation using  $(\text{SiMe}_3)_3\text{SiH}$  and AIBN in benzene to give **20**. Oxidative degradation of the phenyl ring with catalytic  $\text{RuCl}_3$  and stoichiometric amounts of  $\text{H}_5\text{IO}_6$  released the carboxylic acid **21** that was saponified to yield the target compound **22**. To establish the stereochemical course of the epoxide opening, hydroxy acid **22** was reduced to the diol **23** and then converted into the benzaldehyde acetal **24**. For this compound with its rigidified structure, the site of deuterium incorporation could be determined from its NOESY spectrum (Fig. S28<sup>†</sup>), revealing *2R* configuration for **24** and thus epoxide opening of **17** to **18** with inversion of configuration at C3.

After establishing the route, all four stereoisomers of the hydroxy acids **22** with an additional  $^{13}\text{C}$ -labelling at C3 were synthesised and converted into the corresponding SNAC esters **27** (Scheme 3). The synthesis was started from non-labelled (**25a**) or deuterated benzaldehyde (**25b**), prepared itself by DIBAL-D reduction of methyl benzoate and PCC oxidation. A Horner–Wadsworth–Emmons reaction with triethyl ( $2\text{-}^{13}\text{C}$ ) phosphonoacetate yielded **26a** and **26b** that were reduced to cinnamyl alcohols **15a** and **15b**. The Sharpless epoxidation was performed either with *L*-(+)-DIPT to obtain non-deuterated **16aa** and deuterated **16ba**, or with *D*-(-)-DIPT to yield non-deuterated **16ab** and deuterated **16bb**. After conversion into the bromides **17**, for the non-deuterated compounds the epoxides were opened with  $\text{D}_2$  and Pd/C, while for the deuterated compounds  $\text{H}_2$  was used, resulting in all four stereoisomers of **18**. These materials were further transformed into the hydroxy acids **22** according to the established route, followed by conversion into the SNAC esters **27**. All compounds were obtained with high stereoisomeric purity as judged from the NMR analysis of the Mosher esters of **16** (Fig. S8<sup>†</sup>).

The  $^{13}\text{C}$  labels in **27** serve in the sensitive detection of the outcome of DH-catalysed dehydrations with these substrates (Scheme 4). The chemical shift in the  $^{13}\text{C}$ -NMR spectrum for the labelled carbon is  $\delta = 141.13$  ppm for the SNAC ester of (*E*)-but-2-enoic acid (**28a**) and  $\delta = 141.87$  ppm for the SNAC ester of (*Z*)-but-2-enoic acid (**29a**), as confirmed by authentic standards synthesised in this study (Fig. S33 and S35<sup>†</sup>), if the elimination yields non-deuterated products. A deuterium atom in the neighbouring position will lead to a small upfield shift of *ca.*  $\Delta\delta = -0.1$  ppm for the products **28b** and **29b**. Thus, the set of the four SNAC esters of **27** can indicate, through the detection of a single diagnostic  $^{13}\text{C}$  NMR signal, which enantiomer regarding the configuration at C3 is accepted by a DH, which hydrogen is removed from  $\text{C}_\alpha$ , and what is the configuration of the newly installed double bond in the product. Taken together, these data can distinguish between *syn* and *anti* elimination.

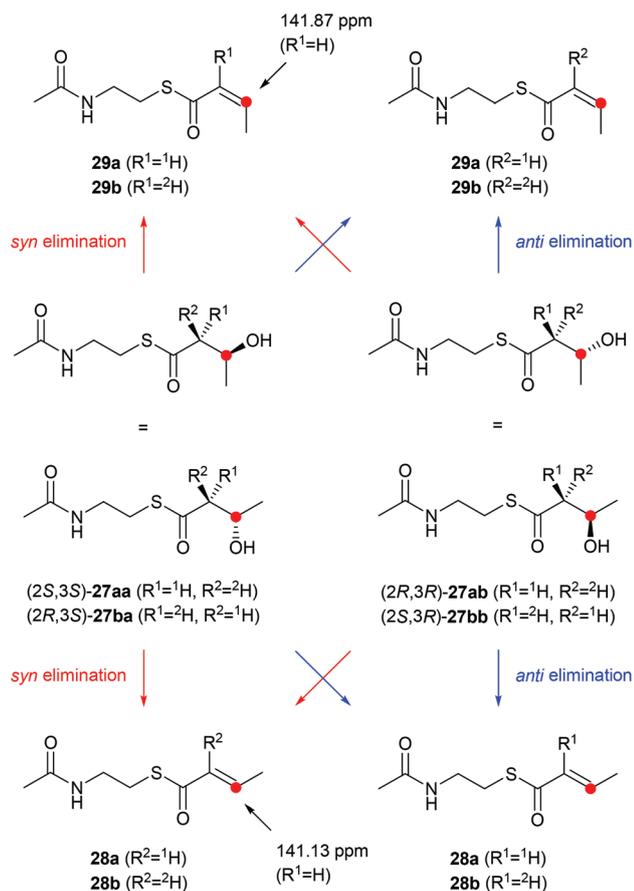
### The stereochemical course of BorDH2, BorDH3 and BorDH5

In the early steps of borrelidin (**30**) biosynthesis,<sup>22</sup> two double bonds are installed by BorDH2 and BorDH3 (Fig. 1A). The first



Scheme 3 Synthesis of  $^{13}\text{C}$ - and  $^2\text{H}$ -labelled probes **27**.

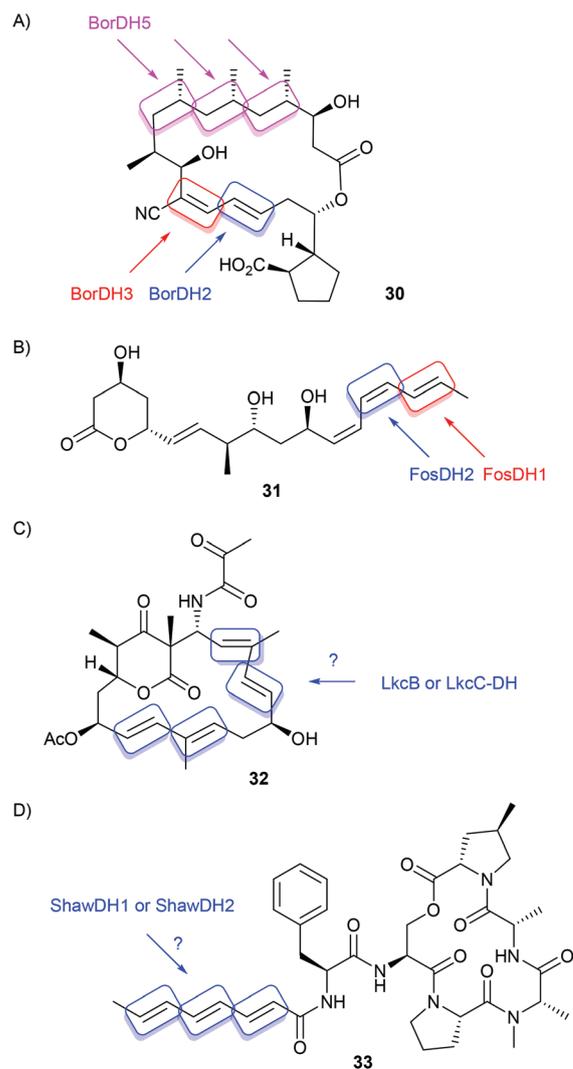
ends up with *E* configuration in **30** and the second is *Z* configured, but both modules contain a B type KR, which raised the question if the borrelidin PKS may not follow the usual stereochemical course with respect to dehydration by BorDH3. As outlined above, incubation experiments with both enzymes (for SDS PAGE analysis of purified enzymes *cf.* Fig. S36<sup>†</sup>) revealed a preference for (*R*)-**11**, while the experiments with



**Scheme 4** Possible results for enzymatic conversions of the  $^{13}\text{C}$  and  $^2\text{H}$  labelled probes **27** with dehydratases.

(*2R,3R*)-**13** pointed to *syn* elimination (Scheme 1F).<sup>18</sup> Today, it is also known that BorDH3 first installs an *E* configured double bond that is isomerised during post-PKS installation of the adjacent nitrile function.<sup>23</sup> Small scale incubations of all four  $^{13}\text{C}$  and  $^2\text{H}$  labelled substrates **27aa–27bb** (0.1 mg each in 100  $\mu\text{L}$  buffer) with BorDH2 showed a conversion of the *3R* configured stereoisomers **27ab** and **27bb**, as indicated by the signals in the olefinic region in the  $^{13}\text{C}$  NMR spectra of the products, but not of the *3S* configured compounds **27aa** and **27ba** (Fig. S37<sup>†</sup>). A detailed inspection showed a slight upfield shift for the  $^{13}\text{C}$  signal of **28b** ( $\delta = 141.03$  ppm) observed with substrate **27ab** in comparison to the signal detected with **27bb** ( $\delta = 141.13$  ppm for **28a**), revealing *syn* elimination for BorDH2 as reported before.<sup>18</sup> Thus, these findings established the proof of concept for the application of the labelled probes to follow the stereochemical course of DHs.

One of the advantages of the labelled probes is that with them small scale reactions can be performed without the need of purification of the products. The obtained products can simply be extracted with  $\text{C}_6\text{D}_6$  and then directly be analysed by  $^{13}\text{C}$  NMR for detection of the  $^{13}\text{C}$  label with high sensitivity. To investigate the limits of detection, different concentrations of the substrate **27bb** (10-fold dilution series from 1 mg to 0.1  $\mu\text{g}$



**Fig. 1** Double bond installations in polyketide biosynthesis. Dehydrations by (A) BorDH2, BorDH3 and BorDH5 in the biosynthesis of borrelidin (**30**), (B) FosDH1 and FosDH2 in the biosynthesis of fostriecin (**31**), (C) LkcB and LkcC-DH in the biosynthesis of lankacidins, represented by lankacidin A (**32**), and (D) candidate enzymes ShawDH1 and ShawDH2 in the biosynthesis of ADEP1 (**33**).

in 100  $\mu\text{L}$  buffer) were incubated with BorDH2. While the experiments with 1 mg and 0.1 mg substrate gave clearly visible peaks and the experiment with 10  $\mu\text{g}$  substrate still produced a detectable signal, lower substrate concentrations did not allow to monitor the reaction (Fig. S38<sup>†</sup>). Therefore, for standard applications, small scale reactions (0.1 mg substrate in 100  $\mu\text{L}$  buffer) were carried out in all subsequent experiments.

In agreement with a previous report, BorDH3 did not accept any of the stereoisomers of **27** (Fig. S39<sup>†</sup>).<sup>18</sup> A third dehydratase in the biosynthetic gene cluster of **31** (BorDH5) has a cryptic stereochemical course, because the double bonds generated in the water eliminations by this iteratively acting DH become subsequently reduced and thus their configur-

ations are not visible in the final product. Incubation of all stereoisomers of **27** with BorDH5 revealed a specificity for the 3*R* configured stereoisomers and *syn* elimination of water with installation of an *E* configured double bond (Fig. S40†).

### The stereochemical course of FosDH1 and FosDH2

Four dehydratase domains are found in the PKS for fostriecin (**31**) biosynthesis from *Streptomyces pulveraceus*.<sup>24</sup> FosDH1 and FosDH2 of the first two elongation modules catalyse the elimination of water to an *E* and a *Z* configured olefin, respectively (Fig. 1B), but with unknown stereochemical course regarding *syn* or *anti* elimination. Incubation of the four **27** stereoisomers with FosDH1 resulted in conversion of only the 3*R* configured substrates with *syn* elimination of water to yield **28b** from **27ab** and **28a** from **27bb**, respectively (Fig. S41†). Notably, the same observations were made with FosDH2 (Fig. S42†), but no *Z* configured products were obtained with this enzyme from the 3*S* configured substrates. This finding reflects the earlier observation that FosDH2 catalyses the elimination of water from the ACP-bound substrate (*R*)-3-hydroxybutanoyl-SACP to yield (*E*)-crotonyl-SACP, while (*S*)-3-hydroxybutanoyl-SACP did not result in any elimination product.<sup>16</sup> The latter finding was explained by an unfavourable equilibrium for the elimination of water on the side of the starting material; indirect proof for the elimination and back reaction was obtained through incorporation of labelling from H<sub>2</sub><sup>18</sup>O.<sup>16</sup> Taken together, these data suggest that selection of the 3*S* configured substrate and installation of a *Z* configured double bond require a longer acyl chain length and eventually ACP bound substrates (Scheme 1C), pointing to a possible involvement of protein–protein interaction between FosDH2 and the corresponding ACP.

### The stereochemical course of LkcB

Lankacidin (**32**) biosynthesis in *Streptomyces rochei* is performed by a *trans*-AT biosynthetic gene cluster<sup>25–27</sup> that encodes the discrete dehydratase LkcB and a later recognised C-terminal dehydratase domain as part of the multidomain enzyme LkcC (LkcC-DH, Fig. 1C).<sup>28</sup> Site-directed mutagenesis of the active site His required for enzyme activity and gene knockout experiments revealed that both dehydratases are required for the biosynthesis of **32** and its cometabolites, but it is currently unknown which of the four elimination steps is performed by which DH. The incubation of LkcB with the four stereoisomers of **27** revealed acceptance of the 3*R* configured substrates and conversion of **27ab** into **28b** and of **27bb** into **28a** through *syn* elimination (Fig. S43†), albeit with lower conversion as compared to the other investigated DHs.

### Activity and stereochemical course of ShawDH1 and ShawDH2

The biosynthesis of the acyldepsipeptide ADEP1 (**33**) in *Streptomyces hawaiiensis* (Fig. 1D) is performed with involvement of a non-ribosomal peptide synthase (NRPS).<sup>29</sup> The short polyunsaturated fatty acyl side chain attached to the phenylalanine residue in **33** is likely made by a type II PKS encoded in the ADEP1 biosynthetic gene cluster that is composed of a

set of discrete enzymes including a putative ACP, KS<sub>α</sub> and KS<sub>β</sub> subunits, and a KR, while a DH for introduction of the double bonds into the unsaturated fatty acyl group is missing in the ADEP1 cluster.<sup>29</sup> A bioinformatic analysis revealed two possible candidates outside the cluster, ShawDH1 and ShawDH2, that may be involved in ADEP1 biosynthesis.<sup>29</sup> Incubation of the four **27** stereoisomers with both enzymes gave no activity with ShawDH1, but selective conversion of the 3*R* configured substrates by ShawDH2, with formation of **28b** from **27ab** and **28a** from **27bb** through *syn* elimination (Fig. S45†). While these data cannot ultimately prove the involvement of ShawDH2 in ADEP1 biosynthesis, the observed inactivity of ShawDH1 *in vitro* may disfavour its role in ADEP1 biosynthesis. Further research is required to clarify this point.

### The stereochemical course of *Escherichia coli* FabZ

While fatty acid biosynthesis is generally believed to proceed with *syn* elimination of water from *R* configured 3-hydroxyacyl intermediates, this has only been established experimentally in a few cases including the dehydratase of yeast fatty acid synthase and of mammalian fatty acid synthase from chicken liver.<sup>3,4</sup> While mammalian and fungal fatty acid biosynthesis is performed by multidomain megasynthases, bacterial fatty acid biosynthesis makes use of a set of discrete enzymes including the dehydratase FabZ,<sup>30–32</sup> but notably, the detailed stereochemical course for this enzyme has never been investigated experimentally from any bacterium. Incubation of the four stereoisomers of **27** with FabZ from *Escherichia coli* gave efficient conversion of the 3*R* configured substrates with *syn* elimination of water (Fig. S46†), while the 3*S* configured substrates were not accepted. Thus, the DH of bacterial fatty acid biosynthesis follows the same stereochemical course as known from yeast and mammalian fatty acid synthase.<sup>3,4</sup>

## Conclusions

The four synthetic isotopically labelled stereoisomers of **27** can be used to rapidly and sensitively investigate the stereochemical courses of dehydratases involved in polyketide and fatty acid biosynthesis. In the current work dehydratases from fatty acid (*E. coli* FabZ), borrelidin (BorDH2, BorDH3 and BorDH5), fostriecin (FosDH1, FosDH2), lankacidin (LkcB) and potentially ADEP1 biosynthesis (ShawDH1, ShawDH2) were investigated, showing activity with the 3*R* configured substrates **27ab** and **27bb** in all cases but BorDH3 and ShawDH1 that were inactive. Furthermore, all experiments uniformly revealed *syn* elimination with formation of an *E* configured double bond in the product. These results can be expected in most of the investigated cases, because *E* configured double bonds are also observed in the corresponding polyketides, but FosDH2 is involved in the formation of a *Z* configured double bond in fostriecin biosynthesis. Thus, the substrate's chain length or protein–protein interaction as provided by the cellular context of the fostriecin producer *S. pulveraceus* may be relevant in the natural system to install the observed *Z* double bond geometry.

In contrast to the stereospecificity of KR domains, it is impossible to predict from the amino acid sequences of dehydratases if their elimination product is *E* or *Z* configured, which suggests that protein–protein interactions may be more generally relevant to achieve elimination with *Z* double bond formation. While the labelled probes introduced here turned out to be very useful in following the stereochemical courses of dehydratase reactions, their limitations lie in questions about the timing of events, *i.e.* they cannot distinguish between concerted or highly asynchronous (stepwise) reactions that may proceed through partially developed enol(ate) intermediates.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft DFG (TRR 261, project ID 398967434). We thank Andreas Schneider for HPLC purifications.

## References

- 1 K. Weissman, Polyketide stereocontrol: a study in chemical biology, *Beilstein J. Org. Chem.*, 2017, **13**, 348–371.
- 2 P. Caffrey, Conserved Amino Acid Residues Correlating With Ketoreductase Stereospecificity in Modular Polyketide Synthases, *ChemBioChem*, 2003, **4**, 654–657.
- 3 B. Sedgwick, C. Morris and S. J. French, Stereochemical Course of Dehydration Catalysed by the Yeast Fatty Acid Synthetase, *J. Chem. Soc., Chem. Commun.*, 1978, 193–194.
- 4 V. E. Anderson and G. G. Hammes, Stereochemistry of the Reactions Catalyzed by Chicken Liver Fatty Acid Synthase, *Biochemistry*, 1984, **23**, 2088–2094.
- 5 D. Gay, Y.-O. You, A. Keatinge-Clay and D. E. Cane, Structure and Stereospecificity of the Dehydratase Domain from the Terminal Module of the Rifamycin Polyketide Synthase, *Biochemistry*, 2013, **52**, 8916–8928.
- 6 C. R. Valenzano, Y.-O. You, A. Garg, A. Keatinge-Clay, C. Koshla and D. E. Cane, Stereospecificity of the Dehydratase Domain of the Erythromycin Polyketide Synthase, *J. Am. Chem. Soc.*, 2010, **132**, 14697–14699.
- 7 X. Guo, T. Liu, C. R. Valenzano, Z. Deng and D. E. Cane, Mechanism and Stereospecificity of a Fully Saturating Polyketide Synthase Module: Nanchangmycin Synthase Module 2 and Its Dehydratase Domain, *J. Am. Chem. Soc.*, 2010, **132**, 14694–14696.
- 8 J. M. Schwab, A. Habib and J. B. Klassen, A Thorough Study of the Stereochemical Consequences of the Hydration/Dehydration Reaction Catalyzed by  $\beta$ -Hydroxydecanoyl Thioester Dehydrase, *J. Am. Chem. Soc.*, 1986, **108**, 5304–5308.
- 9 R. Reid, M. Piagentini, E. Rodriguez, G. Ashley, N. Viswanathan, J. Carney, D. V. Santi, C. R. Hutchinson and R. McDaniel, A Model of Structure and Catalysis for Ketoreductase Domains in Modular Polyketide Synthases, *Biochemistry*, 2003, **42**, 72–79.
- 10 Z. Yin and J. S. Dickschat, Cis double bond formation in polyketide biosynthesis, *Nat. Prod. Rep.*, 2021, **38**, 1445–1468.
- 11 Y. Li, G. J. Dodge, W. D. Fiers, R. A. Fecik, J. L. Smith and C. C. Aldrich, Functional Characterization of a Dehydratase Domain from the Pikromycin Polyketide Synthase, *J. Am. Chem. Soc.*, 2015, **137**, 7003–7006.
- 12 W. D. Fiers, G. J. Dodge, D. H. Sherman, J. L. Smith and C. C. Aldrich, Vinylogous Dehydration by a Polyketide Dehydratase Domain in Curacin Biosynthesis, *J. Am. Chem. Soc.*, 2016, **138**, 16024–16036.
- 13 G. J. Dodge, D. Ronnow, R. E. Taylor and J. L. Smith, Molecular Basis for Olefin Rearrangement in the Gephyronic Acid Polyketide Synthase, *ACS Chem. Biol.*, 2018, **13**, 2699–2707.
- 14 G. Berkhan, C. Merten, C. Holec and F. Hahn, The Interplay between a Multifunctional Dehydratase Domain and a C-Methyltransferase Effects Olefin Shift in Ambruticin Biosynthesis, *Angew. Chem., Int. Ed.*, 2016, **55**, 13589–13592.
- 15 G. Berkhan and F. Hahn, A Dehydratase Domain in Ambruticin Biosynthesis Displays Additional Activity as a Pyran-Forming Cyclase, *Angew. Chem., Int. Ed.*, 2014, **53**, 14240–14244.
- 16 D. D. Shah, Y.-O. You and D. E. Cane, Stereospecific Formation of *E*- and *Z*-Disubstituted Double Bonds by Dehydratase Domains from Modules 1 and 2 of the Fostriecin Polyketide Synthase, *J. Am. Chem. Soc.*, 2017, **139**, 14322–14330.
- 17 X. Xie and D. E. Cane, Stereospecific Formation of *Z*-Trisubstituted Double Bonds by the Successive Action of Ketoreductase and Dehydratase Domains from trans-AT Polyketide Synthases, *Biochemistry*, 2018, **57**, 3126–3129.
- 18 O. Vergnolle, F. Hahn, A. Baerga-Ortiz, P. F. Leadlay and J. N. Anderson, Stereoselectivity of Isolated Dehydratase Domains of the Borrelidin Polyketide Synthase: Implications for cis Double Bond Formation, *ChemBioChem*, 2011, **12**, 1011–1014.
- 19 T. Katsuki and K. B. Sharpless, The First Practical Method for Asymmetric Epoxidation, *J. Am. Chem. Soc.*, 1980, **102**, 5974–5976.
- 20 J. A. Dale, D. L. Dull and H. S. Mosher,  $\alpha$ -Methoxy- $\alpha$ -trifluoromethylphenylacetic acid, a versatile reagent for the determination of enantiomeric composition of alcohols and amines, *J. Org. Chem.*, 1969, **34**, 2543–2549.
- 21 N. Di Blasio, M. T. Lopardo and P. Lupattelli, 1,2-Diarylethanols by Alternative Regioselective Reductive Ring-Opening of 2,3-Diaryloxiranes, *Eur. J. Org. Chem.*, 2009, 938–944.
- 22 C. Olano, B. Wilkinson, C. Sanchez, S. J. Moss, R. Sheridan, V. Math, A. J. Weston, A. F. Brana, C. J. Martin, M. Oliynyk, C. Mendez, P. F. Leadlay and J. A. Salas, Biosynthesis of the

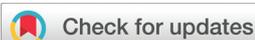
- Angiogenesis Inhibitor Borrelidin by *Streptomyces parvulus* Tü 4055: Cluster Analysis and Assignment of Functions, *Chem. Biol.*, 2004, **11**, 87–97.
- 23 N. Kandziora, J. N. Andexer, S. J. Moss, B. Wilkinson, P. F. Leadlay and F. Hahn, Uncovering the origin of Z-configured double bonds in polyketides: intermediate E-double bond formation during borrelidin biosynthesis, *Chem. Sci.*, 2014, **5**, 3563–3567.
- 24 R. Kong, X. Liu, C. Su, C. Ma, R. Qiu and L. Tang, Elucidation of the Biosynthetic Gene Cluster and the Post-PKS Modification Mechanism for Fostriecin in *Streptomyces pulveraceus*, *Chem. Biol.*, 2013, **20**, 45–54.
- 25 K. Arakawa, F. Sugino, K. Kodama, T. Ishii and H. Kinashi, Cyclization Mechanism for the Synthesis of Macrocyclic Antibiotic Lankacidin in *Streptomyces rochei*, *Chem. Biol.*, 2005, **12**, 249–256.
- 26 S. Tatsuno, K. Arakawa and H. Kinashi, Analysis of Modular-iterative Mixed Biosynthesis of Lankacidin by Heterologous Expression and Gene Fusion, *J. Antibiot.*, 2007, **60**, 700–708.
- 27 S. Tatsuno, K. Arakawa and H. Kinashi, Cyclization Mechanism for the Synthesis of Macrocyclic Antibiotic Lankacidin in *Streptomyces rochei*, *Biosci. Biotechnol. Biochem.*, 2009, **73**, 2712–2719.
- 28 J. S. Dickschat, O. Vergnolle, H. Hong, S. Garner, S. R. Bidgood, H. C. Dooley, Z. Deng, P. F. Leadlay and Y. Sun, An Additional Dehydratase-like Activity is Required for Lankacidin Antibiotic Biosynthesis, *ChemBioChem*, 2011, **12**, 2408–2412.
- 29 D. Thomy, E. Culp, M. Adamek, E. Y. Cheng, N. Ziemert, G. D. Wright, P. Sass and H. Brötz-Oesterhelt, The ADEP Biosynthetic Gene Cluster in *Streptomyces hawaiiensis* NRRL 15010 Reveals an Accessory *clpP* Gene as a Novel Antibiotic Resistance Factor, *Appl. Environ. Microbiol.*, 2019, **85**, e01292–e01219.
- 30 S. Mohan, T. M. Kelly, S. S. Eveland, C. R. H. Raetz and M. S. Anderson, An *Escherichia coli* Gene (*FabZ*) Encoding (3*R*)-Hydroxymyristoyl Acyl Carrier Protein Dehydratase, *J. Biol. Chem.*, 1994, **269**, 32896–32903.
- 31 R. J. Heath and C. O. Rock, Roles of the FabA and FabZ b-Hydroxyacyl-Acyl Carrier Protein Dehydratases in *Escherichia coli* Fatty Acid Biosynthesis, *J. Biol. Chem.*, 1996, **271**, 27795–27801.
- 32 M. S. Kimber, F. Martin, Y. Lu, S. Houston, M. Vedadi, A. Dharamsi, K. M. Fiebig, M. Schmid and C. O. Rock, The Structure of (3*R*)-Hydroxyacyl-Acyl Carrier Protein Dehydratase (FabZ) from *Pseudomonas aeruginosa*, *J. Biol. Chem.*, 2004, **279**, 52593–52602.

## Appendix D

### **The Substrate Scope of Dehydratases in Antibiotic Biosynthesis and their Application in Kinetic Resolutions**

*Org. Biomol. Chem.* **2022**, *20*, 9103–9107.

[doi.org/10.1039/D2OB01879A](https://doi.org/10.1039/D2OB01879A)

Cite this: *Org. Biomol. Chem.*, 2022, **20**, 9103Received 13th October 2022,  
Accepted 3rd November 2022

DOI: 10.1039/d2ob01879a

rsc.li/obc

## The substrate scope of dehydratases in antibiotic biosynthesis and their application in kinetic resolutions†

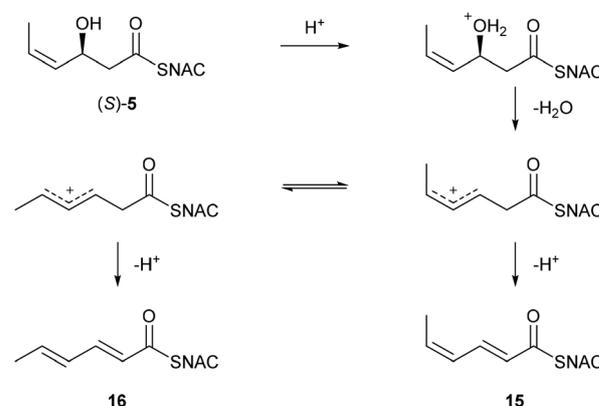
Zhiyong Yin,<sup>a</sup> Daniel Bär,<sup>b</sup> Bertolt Gust <sup>b</sup> and Jeroen S. Dickschat <sup>\*a</sup>

**Nine dehydratases involved in the biosynthesis of secondary metabolites in addition to FabZ from fatty acid biosynthesis were investigated for their substrate scope using a panel of *N*-acetylcysteamine (SNAC) thioesters. The best performing enzyme BorDH2 was applied in kinetic resolutions.**

Antibiotic biosynthesis in bacteria often proceeds through the action of type I polyketide synthases (PKSs). These remarkable megasynthases are composed of modules each of which is usually responsible for the incorporation of one PKS extender unit. The modules are composed of several domains including an acyl carrier protein (ACP) domain to which the growing polyketide chain is tethered, an acyltransferase (AT) domain for substrate upload, and a ketosynthase (KS) domain for the fusion of the next extender unit with the existing chain. Optional (reductive loop) domains for the modification of the incorporated extender unit include ketoreductases (KR), dehydratases (DH) and enoylreductases (ER). The domain organisation often logically follows the structure of the PKS product (colinearity),<sup>1</sup> but especially for *trans*-AT PKSs many exceptions are known.<sup>2</sup> The stereochemical implications of the reductive loop enzymes are well known,<sup>3</sup> e.g. KRs fall into two types, with type A catalysing the reduction of 3-ketoacyl intermediates to 3L alcohols, while B type KRs generate 3D alcohols.<sup>4</sup> Usually through *syn* elimination of water<sup>5–7</sup> with the specific abstraction of the 2-*pro-S* proton,<sup>8</sup> DHs convert 3D alcohols into *E* configured olefins, which represents the most frequently observed case, but some polyketides with *Z*-configured double bonds are known<sup>9</sup> to typically originate from 3L alcohols.<sup>10</sup> We have recently established a set of isotopically labelled probes

with stereoselective deuterations that allow the investigation of the stereochemical course of DHs *in vitro*.<sup>11</sup> However, the operability of the natural stereochemical course of the elimination reaction may depend on protein–protein interactions between an ACP and a DH. For instance, FosDH2 involved in fostriecin biosynthesis only catalyses the installation of a *Z* double bond when the substrate is ACP-bound,<sup>12</sup> while with a substrate surrogate (*N*-acetylcysteamine thioester, SNAC ester), an *E* configured double bond is obtained.<sup>11</sup> With respect to the substrate scope of PKS domains, much research has been performed on KS domains,<sup>13</sup> KR domains,<sup>14–16</sup> AT domains,<sup>17–21</sup> and TE domains,<sup>22</sup> and also a few investigations have addressed the substrate preference of DHs.<sup>23–26</sup> Here, we report on the substrate scope and stereoselectivity of various DHs involved in antibiotic and fatty acid biosynthesis.

To investigate the catalytic potential of DHs, the SNAC esters shown in Scheme 1 were synthesised. Compounds **1** and **2** were directly accessible from the commercially available corresponding carboxylic acids (Schemes S1 and S2†). SNAC esters **3–7** were prepared stereoselectively using aldol reactions with Evans' oxazolidinones<sup>27</sup> as a key step (Schemes S3–S7†). Despite several attempts, this chemistry failed for compound



Scheme 1 Stepwise elimination of water with (S)-5.

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

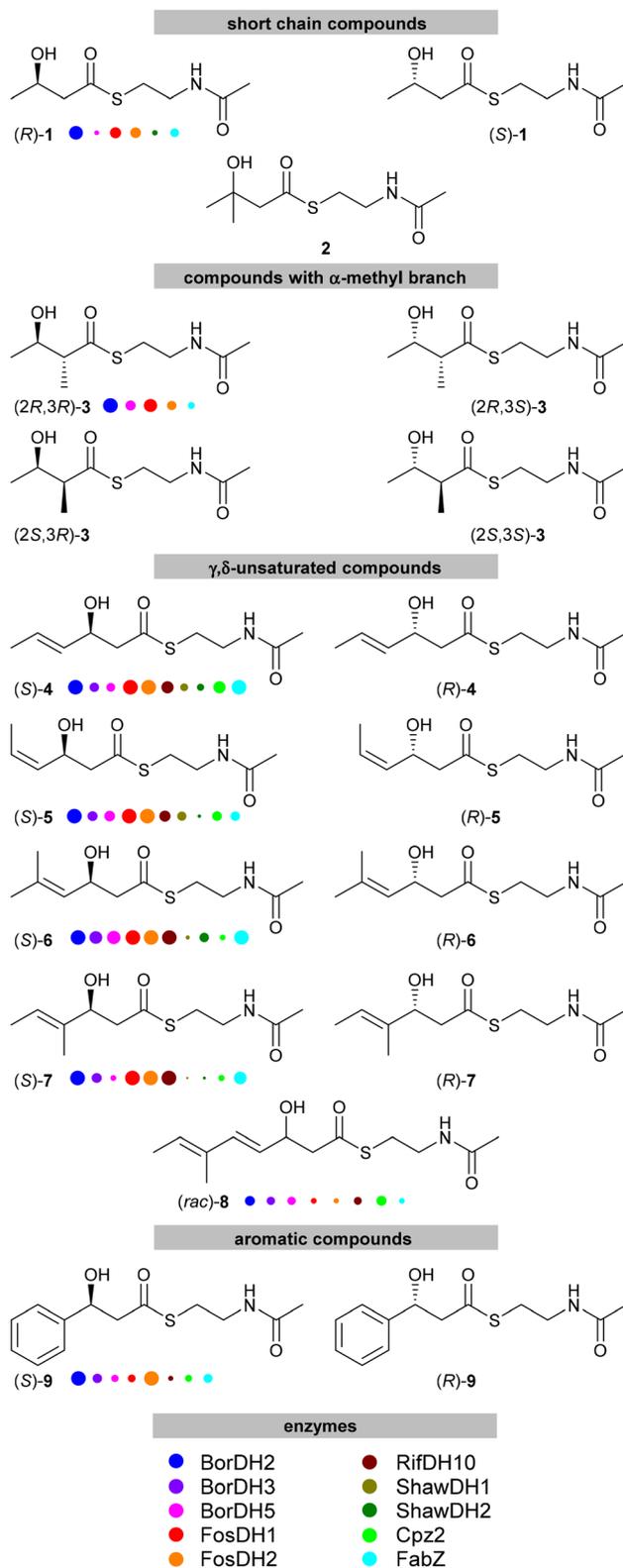
<sup>b</sup>Pharmaceutical Biology, University of Tübingen, Auf der Morgenstelle 8, 72076 Tübingen, Germany

†Electronic supplementary information (ESI) available: Synthetic procedures, gene cloning and expression, screening of dehydratases, and chromatograms of the products obtained from kinetic resolutions. See DOI: <https://doi.org/10.1039/d2ob01879a>

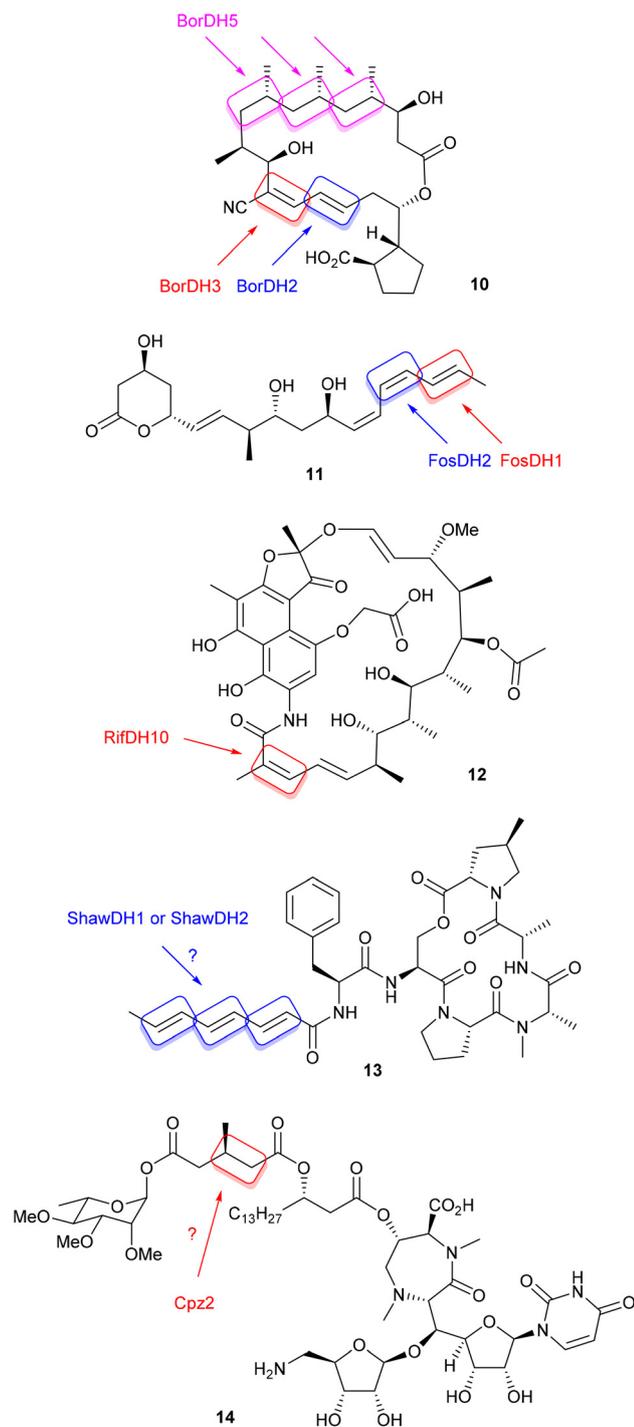
**8**, which was thus obtained as a racemate using a standard aldol reaction of 4-methylsorbic aldehyde with the ester enolate of ethyl acetate (Scheme S8<sup>†</sup>). The enantiomers of **9** were also prepared from commercially available carboxylic acids (Scheme S9<sup>†</sup>). The obtained 3-hydroxy SNAC esters could be grouped into short chain compounds, compounds with an  $\alpha$ -methyl branch, compounds with  $\gamma,\delta$ -unsaturation, and aromatic compounds (Fig. 1).

The synthetic SNAC esters were tested as substrates with a panel of DHs (Fig. S1<sup>†</sup>), including the DH domains BorDH2, BorDH3 and BorDH5 from the borrelidin (**10**) PKS,<sup>28–30</sup> FosDH1 and FosDH2 from fostriecin (**11**) biosynthesis,<sup>12,31</sup> and RifDH10 from the PKS for rifamycin (**12**),<sup>32,33</sup> the discrete and non-cluster associated ShawDH1 and ShawDH2 potentially involved in the biosynthesis of the acyldepsipeptide ADEP1 (**13**),<sup>11,34</sup> the discrete DH Cpz2 from the caprazamycin (**14**) biosynthetic gene cluster,<sup>35</sup> and FabZ from fatty acid biosynthesis in *Escherichia coli* (for the structures of the compounds, cf. Fig. 2).<sup>36,37</sup>

For an initial qualitative substrate screening, small scale reactions with 1 mg of the substrate were performed. The reactions were monitored by <sup>1</sup>H-NMR spectroscopy with comparison of the recorded spectra to those of the starting material and synthetic reference standards of DH products (cf. Scheme S10<sup>†</sup> for their synthesis). None of these enzymes accepted any of the substrates with 3L configuration (compounds on the right side of Scheme 1; because of a change in the priority orders of the substituents, the formal 3R and 3S configurations are changing), and also compound **2** with a tertiary alcohol function was not converted by any of the tested DHs (Table S1<sup>†</sup>). Among the 3D-configured substrates, the short chain compound (*R*)-**1** was efficiently converted into crotonyl-SNAC by BorDH2, FosDH1, FosDH2 and FabZ, but not converted or only poorly converted (BorDH5, ShawDH2) by the other DHs (Fig. S2<sup>†</sup>). Compounds with an  $\alpha$ -methyl branch as in 3-hydroxy-2-methylbutyryl-SNAC (**3**) were only accepted with the (2*R*,3*R*) configuration allowing for *syn* dehydration to an *E*-configured product, while (2*S*,3*R*)-**3** was not converted by any enzyme. Successful cases with (2*R*,3*R*)-**3** included BorDH2 and FosDH1, and lower conversions were found for BorDH5, FosDH2 and FabZ (Fig. S3<sup>†</sup>). Substrate (*S*)-**4** with  $\gamma,\delta$ -unsaturation was readily accepted by all enzymes, and complete conversions were observed with BorDH2, FosDH1, FosDH2 and FabZ (Fig. S4<sup>†</sup>). The stereoisomer (*S*)-**5** was also dehydrated by all enzymes, with full consumption of the starting material by BorDH2, FosDH1 and FosDH2, while ShawDH2 showed only poor conversion (Fig. S5<sup>†</sup>). Interestingly, depending on the enzyme, not only was the 2*E*,4*Z*-configured elimination product **15** observed, but also varying amounts of the 2*E*,4*E* product **16** were obtained (this was the major product of ShawDH1, and BorDH3 and Cpz2 also yielded this product in substantial amounts). The formation of **16** may point to a step-wise elimination reaction through protonation of the alcohol function, followed by elimination of water to an allyl cation intermediate that may be configurationally unstable and subsequent deprotonation (Scheme 1). In other cases, the formation of **16** is less pronounced (e.g. BorDH5), which may point



**Fig. 1** Synthetic 3-hydroxy SNAC esters used as DH substrates. The colour code indicates which enzyme can convert which substrate. The area of the coloured dots is proportional to the conversion rates in % as detailed in Table S1.<sup>†</sup>



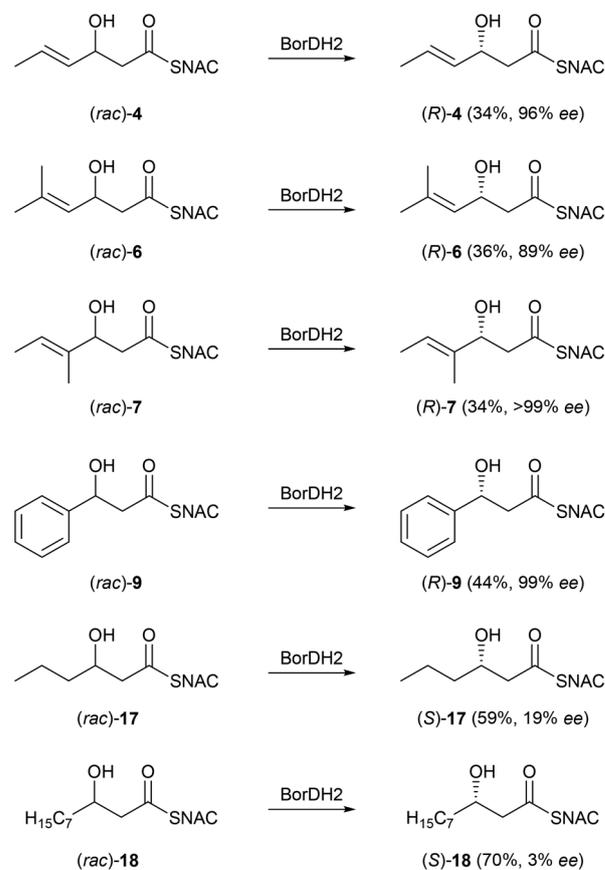
**Fig. 2** Structures of the compounds produced with the (potential) involvement of the DHs used in this study. The coloured boxes indicate the sites on which the DHs (potentially) act.

to a concrete elimination reaction, or steric constraints in the active site pocket do not allow for the configurational isomerisation of the allyl cation intermediate.

With substrate (*S*)-6, complete dehydration was obtained with BorDH2, FosDH1, FosDH2 and RifDH10, while poor substrate acceptance was found for ShawDH1, ShawDH2 and Cpz2

(Fig. S6†). Several enzymes including BorDH2, BorDH3, FosDH1, FosDH2, RifDH10 and FabZ catalysed the dehydration of (*S*)-7, with full conversion by BorDH2, FosDH1 and FosDH2, but BorDH5, ShawDH1, ShawDH2 and Cpz2 acted sluggishly on this substrate (Fig. S7†). For (*rac*)-8, the most efficient dehydration was observed with BorDH2 and Cpz2. In addition, a reduced conversion rate was observed for BorDH3, BorDH5, FosDH1, FosDH2, RifDH10 and FabZ, while ShawDH1 and ShawDH2 were inactive (Fig. S8†). With this substrate, the conversion never exceeded 50%, likely because only the *S* enantiomer of (*rac*)-8 was accepted. For (*S*)-9, a complete conversion was observed with BorDH2 and FosDH2, partial elimination of water was found for BorDH3, BorDH5, FosDH1, Cpz2 and FabZ, poor enzymatic conversion was observed with RifDH10, and ShawDH1 and ShawDH2 gave no elimination product (Fig. S9†). Taken together, the best results were obtained for BorDH2 that accepted most substrates and showed the highest efficiency in their dehydration. Therefore, this enzyme was selected for further study.

BorDH2 was first investigated for its efficiency with different substrates in triplicate experiments, showing the reproducibility of full conversion (100%) of the  $\gamma,\delta$ -unsaturated substrates (*S*)-4, (*S*)-5, (*S*)-6 and (*S*)-7 and the aromatic com-



**Scheme 2** Kinetic resolutions with BorDH2. Dehydration products are not shown.

pound (*S*)-**9** (Table S2†). Also for substrates (*R*)-**1** and (*2R,3R*)-**3**, a good performance was observed ( $92 \pm 4\%$  and  $99 \pm 1\%$  conversion, respectively), while for (*rac*)-**8**, the conversion reached  $33 \pm 5\%$ . Subsequently, the application of BorDH2 in preparative scale reactions was tested (up to 50 mg, 0.19 mmol of substrate; for details, cf. the section “Kinetic resolutions with BorDH2” in the ESI†). The reactions performed with BorDH2 using enantiomerically pure 3D alcohols are not very useful in synthesis, because the enantiomerically pure materials are difficult to be synthesized and the stereochemical information will be lost in the product, but for enzymatic kinetic resolutions, the application of BorDH2 can be useful. In order to test the practicability of this approach, a series of racemic 3-hydroxy-SNAC esters were synthesised (Scheme S11†) and treated with BorDH2 in kinetic resolutions (Scheme 2). The  $\gamma,\delta$ -unsaturated SNAC esters (*rac*)-**4**, (*rac*)-**6** and (*rac*)-**7** and the aromatic compound (*rac*)-**9** yielded 34%–44% of the 3L-hydroxy-SNAC esters of high enantiomeric purity (between 89% and almost 100% ee, Fig. S10–S13†). In contrast, the saturated compound 3-hydroxyhexanoyl-SNAC ((*rac*)-**17**) reacted sluggishly, which consequently resulted in a higher recovery of the starting material (59%) that was of low enantiomeric purity (19% ee, Fig. S14†). Chain elongation as in the substrate 3-hydroxydecanoyl-SNAC ((*rac*)-**18**) gave worse results (70% recovery, 3% ee, Fig. S15†), showing the limitations of this method.

## Conclusions

Seven dehydratases from the biosynthetic machineries for borrelidin, fostriecin, rifamycin and caprazamycin, besides two dehydratases putatively involved in the biosynthesis of ADEP1 and FabZ from fatty acid biosynthesis, were screened for their conversion rates towards a panel of 3-hydroxy-SNAC esters. All enzymes only accepted 3D-configured alcohols, while 3L-configured alcohols were not dehydrated. The enzymes selected for this study naturally act on structurally very different biosynthetic intermediates towards the synthesis of distinct natural products, which makes it difficult to understand why some of the investigated enzymes such as BorDH2, FosDH1 and FosDH2 show a broad substrate scope, while other dehydratases show lower substrate acceptance. In consequence, only a broad substrate screening as presented in this study can be performed to identify useful biocatalysts for synthetic approaches. In this work, BorDH2 from borrelidin biosynthesis exhibited the highest substrate tolerance and catalytic efficiency. As we have shown, this enzyme is suitable for the preparation of 3L-hydroxy-SNAC esters of high enantiomeric purity through kinetic resolutions. Good results were obtained for  $\gamma,\delta$ -unsaturated and aromatic 3-hydroxy-SNAC esters, while the application of this method to saturated substrates, especially with longer chains, is not suitable. After the discovery of the principle by Pasteur through fermentation of racemic ammonium tartrate with *Penicillium glaucum*, and the introduction of lipase to kinetic resolutions<sup>38</sup> that is today

widely used in synthetic applications,<sup>39</sup> BorDH2 has been added to the portfolio of suitable enzymes for kinetic resolutions.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft DFG (TRR 261, project ID 398967434). We thank Andreas Schneider for HPLC separation.

## References

- C. R. Hutchinson, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 3336.
- E. J. N. Helfrich and J. Piel, *Nat. Prod. Rep.*, 2016, **33**, 231.
- K. J. Weissman, *Beilstein J. Org. Chem.*, 2017, **13**, 348.
- P. Caffrey, *ChemBioChem*, 2003, **4**, 654.
- D. Gay, Y.-O. You, A. Keatinge-Clay and D. E. Cane, *Biochemistry*, 2013, **52**, 8916.
- C. R. Valenzano, Y.-O. You, A. Garg, A. Keatinge-Clay, C. Koshla and D. E. Cane, *J. Am. Chem. Soc.*, 2010, **132**, 14697.
- X. Guo, T. Liu, C. R. Valenzano, Z. Deng and D. E. Cane, *J. Am. Chem. Soc.*, 2010, **132**, 14694.
- J. M. Schwab, A. Habib and J. B. Klassen, *J. Am. Chem. Soc.*, 1986, **108**, 5304.
- Z. Yin and J. S. Dickschat, *Nat. Prod. Rep.*, 2021, **38**, 1445.
- J. W. Labonte and C. A. Townsend, *Chem. Rev.*, 2013, **113**, 2182.
- Z. Yin, E. Liebhart, E. Stegmann, H. Brötz-Oesterhelt and J. S. Dickschat, *Org. Chem. Front.*, 2022, **9**, 2714.
- D. D. Shah, Y.-O. You and D. E. Cane, *J. Am. Chem. Soc.*, 2017, **139**, 14322.
- A. C. Murphy, H. Hong, S. Vance, R. W. Broadhurst and P. F. Leadlay, *Chem. Commun.*, 2016, **52**, 8373.
- M. Schröder, T. Ross, F. Hemmerling and F. Hahn, *ACS Chem. Biol.*, 2022, **17**, 1030.
- H.-M. Ge, T. Huang, J. D. Rudolf, J. R. Lohman, S.-X. Huang, X. Guo and B. Shen, *Org. Lett.*, 2014, **16**, 3958.
- J. Wunderlich, T. Ross, M. Schröder and F. Hahn, *Org. Lett.*, 2020, **22**, 4955.
- F. Del Vecchio, H. Petkovic, S. G. Kendrew, L. Low, B. Wilkinson, R. Lill, J. Cortes, B. A. M. Rudd, J. Staunton and P. F. Leadlay, *J. Ind. Microbiol. Biotechnol.*, 2003, **30**, 489.
- I. Koyakina, J. B. McArthur, M. W. Draelos and G. J. Williams, *Org. Biomol. Chem.*, 2013, **11**, 4449.
- K. Bravo-Rodriguez, S. Klopries, K. R. M. Koopmans, U. Sundermann, S. Yahiaoui, J. Arens, S. Kushnir, F. Schulz and E. Sanchez-Garcia, *Chem. Biol.*, 2015, **22**, 1425.

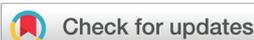
- 20 Y. Li, W. Zhang, H. Zhang, W. Tian, L. Wu, S. Wang, M. Zheng, J. Zhang, C. Sun, Z. Deng, Y. Sun, X. Qu and J. Zhou, *Angew. Chem., Int. Ed.*, 2018, **57**, 5823.
- 21 E. Kalkreuter, K. S. Bingham, A. M. Keeler, A. N. Lowell, J. J. Schmidt, D. H. Sherman and G. J. Williams, *Nat. Commun.*, 2021, **12**, 2193.
- 22 A. Pinto, M. Wang, M. Horsman and C. N. Boddy, *Org. Lett.*, 2012, **14**, 2278.
- 23 W. D. Fiers, G. J. Dodge, Y. Li, J. L. Smith, R. A. Fecik and C. C. Aldrich, *Chem. Sci.*, 2015, **6**, 5027.
- 24 Y. Li, G. J. Dodge, W. D. Fiers, R. A. Fecik, J. L. Smith and C. C. Aldrich, *J. Am. Chem. Soc.*, 2015, **137**, 7003.
- 25 A. Faille, S. Gavalda, N. Slama, C. Lherbet, L. Maveyraud, V. Guillet, F. Laval, A. Quemard, L. Mourey and J.-D. Pedelacq, *J. Mol. Biol.*, 2017, **429**, 1554.
- 26 K. Finzel, C. Nguyen, D. R. Jackson, A. Gupta, S.-C. Tsai and M. D. Burkhardt, *Chem. Biol.*, 2015, **22**, 1453.
- 27 D. A. Evans, J. Bartroli and T. L. Shih, *J. Am. Chem. Soc.*, 1981, **103**, 2127.
- 28 C. Olano, B. Wilkinson, C. Sanchez, S. J. Moss, R. Sheridan, V. Math, A. J. Weston, A. F. Brana, C. J. Martin, M. Oliynyk, C. Mendez, P. F. Leadlay and J. A. Salas, *Chem. Biol.*, 2004, **11**, 87.
- 29 O. Vergnolle, F. Hahn, A. Baerga-Ortiz, P. F. Leadlay and J. N. Andexer, *ChemBioChem*, 2011, **12**, 1011.
- 30 N. Kandziora, J. N. Andexer, S. J. Moss, B. Wilkinson, P. F. Leadlay and F. Hahn, *Chem. Sci.*, 2014, **5**, 3563.
- 31 R. Kong, X. Liu, C. Su, C. Ma, R. Qiu and L. Tang, *Chem. Biol.*, 2013, **20**, 45.
- 32 T. Schupp, C. Toupet, N. Engel and S. Goff, *FEMS Microbiol. Lett.*, 1998, **159**, 201.
- 33 D. Gay, Y. O. You, A. Keatinge-Clay and D. E. Cane, *Biochemistry*, 2013, **52**, 8916.
- 34 D. Thomy, E. Culp, M. Adamek, E. Y. Cheng, N. Ziemert, G. D. Wright, P. Sass and H. Brötz-Oesterhelt, *Appl. Environ. Microbiol.*, 2019, **85**, e01292-19.
- 35 L. Kaysser, L. Lutsch, S. Siebenberg, E. Wemakor, B. Kammerer and B. Gust, *J. Biol. Chem.*, 2009, **284**, 14987.
- 36 S. Mohan, T. M. Kelly, S. S. Eveland, C. R. H. Raetz and M. S. Anderson, *J. Biol. Chem.*, 1994, **269**, 32896.
- 37 R. J. Heath and C. O. Rock, *J. Biol. Chem.*, 1996, **271**, 27795.
- 38 H. D. Dakin, *J. Physiol.*, 1903, **30**, 253.
- 39 A. Ghanem and H. Y. Aboul-Enein, *Chirality*, 2004, **17**, 1.

## Appendix E

### **Enantioselective Synthesis of All Stereoisomers of Geosmin and of Biosynthetically Related Natural Products**

*Org. Biomol. Chem.* **2024**, *22*, 5748–5758.

[doi.org/10.1039/D4OB00934G](https://doi.org/10.1039/D4OB00934G)



Cite this: *Org. Biomol. Chem.*, 2024, **22**, 5748

# Enantioselective synthesis of all stereoisomers of geosmin and of biosynthetically related natural products†

Zhiyong Yin,<sup>a</sup> Michael Maczka,<sup>b</sup> Gregor Schnakenburg,<sup>c</sup> Stefan Schulz<sup>b</sup> and Jeroen S. Dickschat<sup>\*,a</sup>

Synthetic routes to geosmin and its enantiomer are well established, but the enantioselective synthesis of stereoisomers of geosmin is unknown. Here a stereoselective synthesis of all stereoisomers of geosmin is reported, yielding all compounds in high enantiomeric purity. Furthermore, the stereoselective synthesis of a geosmin derivative isolated from a mangrove associated streptomycete was performed, establishing the absolute configuration of the natural product. Finally, a new side product of the geosmin synthase from *Streptomyces ambofaciens* was isolated and its structure was elucidated by NMR spectroscopy. The absolute configuration of this new compound was determined through a stereoselective synthesis.

Received 3rd June 2024,  
Accepted 21st June 2024  
DOI: 10.1039/d4ob00934g

rsc.li/obc

## Introduction

Geosmin (**1**) was first isolated by Gerber and Lechevalier in 1965 from the soil bacterium *Streptomyces griseus*.<sup>1</sup> Its structure including the relative configuration was elucidated through chemical degradations and comparison to the four synthetic stereoisomers,<sup>2–4</sup> followed by determination of its absolute configuration through chemical correlation with cybullol.<sup>5</sup> One of its most obvious characteristics is the strong earthy smell that is responsible for the odour of soil.<sup>6</sup>

Today the compound is known to be very widespread and its occurrence in diverse bacteria,<sup>7–12</sup> ascomycete as well as basidiomycete fungi,<sup>13–15</sup> and amoebae<sup>16</sup> has been reported. Also plants including mosses<sup>17,18</sup> and angiosperms<sup>19,20</sup> are known sources of **1**. The small quantities present in beetroot contribute to its earthy note that many humans consider as pleasant,<sup>21</sup> while the compound is an undesired contaminant in aquatic animals especially from aquacultures.<sup>22</sup>

The biosynthesis of **1** has been extensively studied through isotopic labelling experiments.<sup>23–25</sup> After discovery of the geosmin synthase and its coding gene,<sup>26–29</sup> mechanistic

investigations through *in vitro* conversions of various isotopically substituted farnesyl pyrophosphate (FPP) isotopomers were performed.<sup>30,31</sup> Taken together, this experimental work resulted in a detailed model for geosmin biosynthesis (Scheme 1). Starting from FPP, a 1,10-cyclisation to the (*E,E*)-germacradienyl cation (**A**) is followed by a stereospecific deprotonation to isolepidozene (**2**).<sup>30</sup> A reprotonation from the *Re* side at C4<sup>31</sup> induces ring opening to **B** and capture with water leads to (1(10)*E,5E*)-germacradien-11-ol (**3**), a second major product of geosmin synthase. Upon reprotonation from the *Si* side at C1<sup>31</sup> and ring closure to **C** a retro-Prins fragmentation leads to the octalin **4**<sup>32,33</sup> that is again reprotonated from the *Re* side at C2<sup>31</sup> to result in **D**. A 1,2-hydride shift to **E** and capture with water ultimately yield **1**. A side product of geosmin biosynthesis is germacrene **D** (**5**) that is formed through a stereospecific 1,3-hydride shift<sup>30</sup> from **A** to **F** and deprotonation.

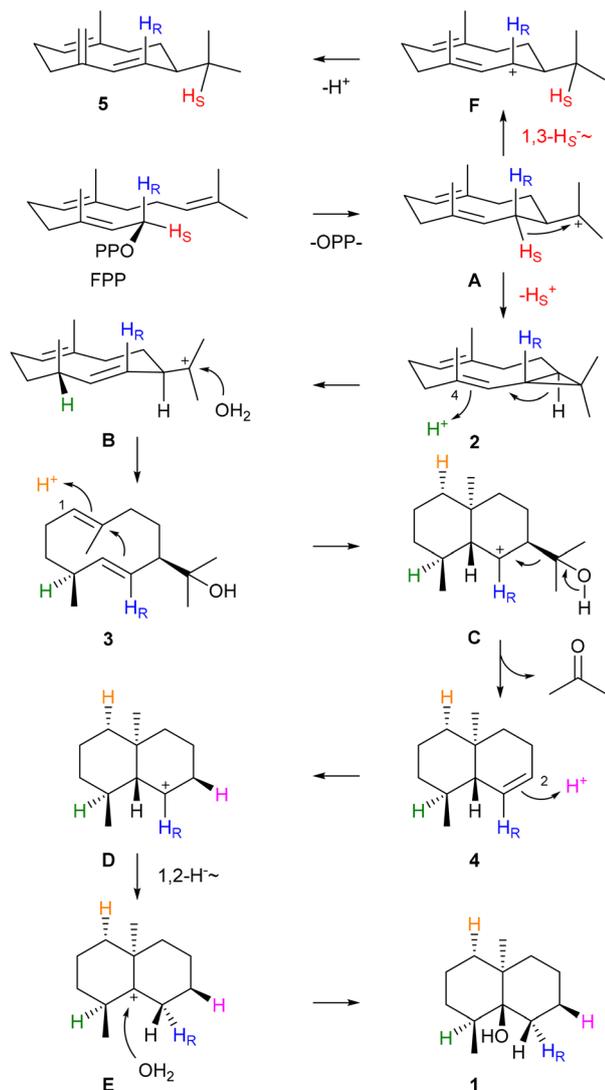
Several naturally occurring geosmin derivatives are known that putatively arise through oxidation (Scheme 2). This includes the hydroxylated compounds cybullol (**6**) from the bird's nest fungus *Cyathus bulleri*,<sup>5</sup> the diols **7–9** from *Streptomyces anulatus*, a streptomycete that was isolated from *Giraffa camelopardalis* feces,<sup>34</sup> the diol **10** from a mangrove associated streptomycete,<sup>35</sup> and the diols **11** and **12** as well as the triol **13** from *Streptomyces albolongus* isolated from *Elephas maximus* feces.<sup>36</sup> The ketones **14** and **15** have been reported as biodegradation products of **1** from *Pseudomonas* sp.,<sup>37</sup> in addition to the occurrence of **15** in the above mentioned mangrove endophyte.<sup>35</sup> Furthermore, dehydrogeosmin (**16**) is known from different species from of *Cactaceae*<sup>38</sup> and has been identified in headspace extracts of different streptomy-

<sup>a</sup>Kekulé-Institute for Organic Chemistry and Biochemistry, University of Bonn, Gerhard-Domagk-Strasse 1, 53121 Bonn, Germany. E-mail: dickschat@uni-bonn.de

<sup>b</sup>Institute for Organic Chemistry, TU Braunschweig, Hagenring 30, 38106 Braunschweig, Germany

<sup>c</sup>Institute for Inorganic Chemistry, University of Bonn, Gerhard-Domagk-Strasse 1, 53121 Bonn, Germany

†Electronic supplementary information (ESI) available. CCDC 2356914 (**10**), 2356915 (**18**), 2356916 (*ent-18*) and 2356917 (*ent-19*). For ESI and crystallographic data in CIF or other electronic format see DOI: <https://doi.org/10.1039/d4ob00934g>



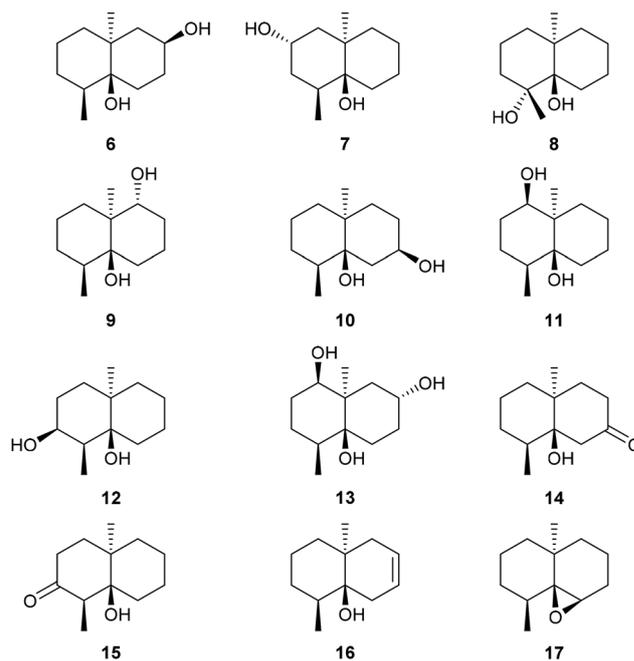
Scheme 1 Biosynthesis of geosmin (1).

cetes,<sup>39</sup> while the epoxide 17 was isolated from the liverwort *Lophocolea bidentata*.<sup>40</sup> For 13 a crystal structure has been reported,<sup>41</sup> and compounds (*rac*)-6,<sup>42</sup> (*rac*)-10,<sup>38</sup> 11,<sup>43</sup> and 16,<sup>38,44</sup> have been synthesised. Also efforts in the enantioselective synthesis of natural (–)-1<sup>45–47</sup> and its enantiomer (+)-1<sup>44,45,48</sup> have been made. Here we report on the enantioselective synthesis of all stereoisomers of 1, of the geosmin derivative 10, and on the identification and synthesis of a previously unrecognised geosmin synthase side product.

## Results and discussion

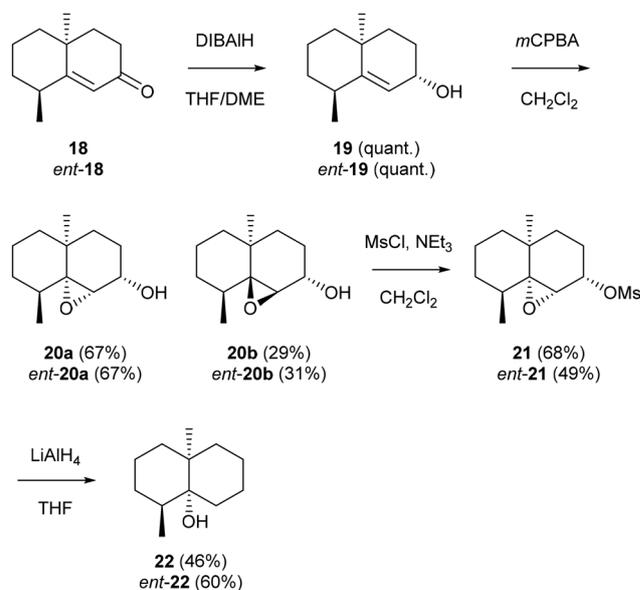
### Enantioselective synthesis of geosmin stereoisomers

A key intermediate in the synthesis of (–)-1 is the octalone 18 that can be prepared in a stereoselective Robinson annelation of 2,6-dimethylcyclohexanone and methyl vinyl ketone using (*R*)-(+)-methylbenzylamine as a chiral auxiliary following



Scheme 2 Known naturally occurring derivatives of geosmin (1).

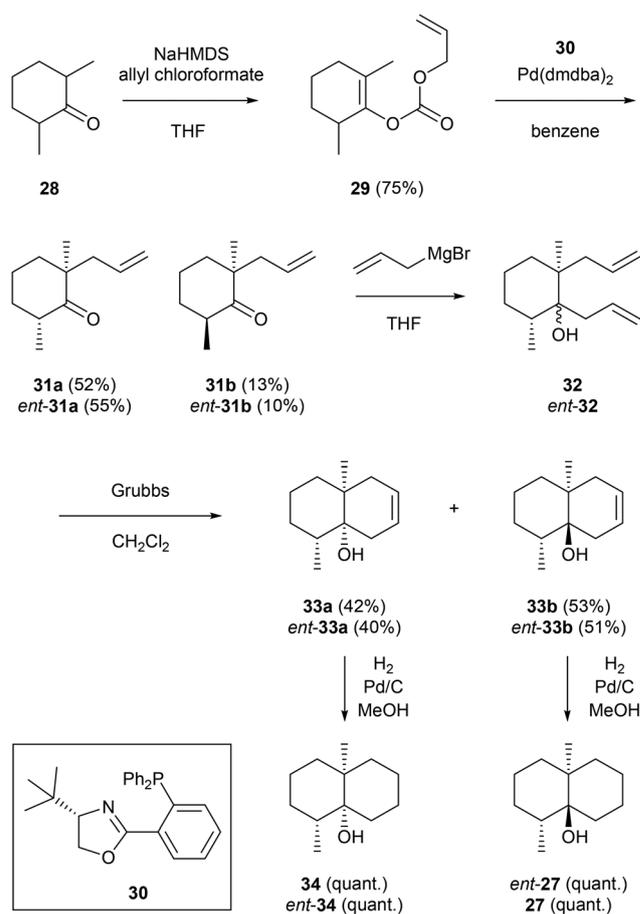
Reviel's procedure.<sup>45</sup> After repeated recrystallisation 18 is obtained in enantiomerically pure form (Fig. S1†).<sup>45</sup> This compound was also the starting point for the synthesis of 5-*epi*-geosmin (22, Scheme 3). A highly diastereoselective reduction of 18 with DIBALH in THF/DME results in allyl alcohol 19.<sup>46</sup> Epoxidation with *m*CPBA yields the epoxides 20a and 20b (dr  $\approx$  2 : 1) in favour of the desired diastereomer. After chromatographic separation 20a is converted into the mesylate 21 that upon treatment with  $\text{LiAlH}_4$  gives access to 5-*epi*-geosmin (22).

Scheme 3 Synthesis of 5-*epi*-geosmin (22) and *ent*-5-*epi*-geosmin (*ent*-22).

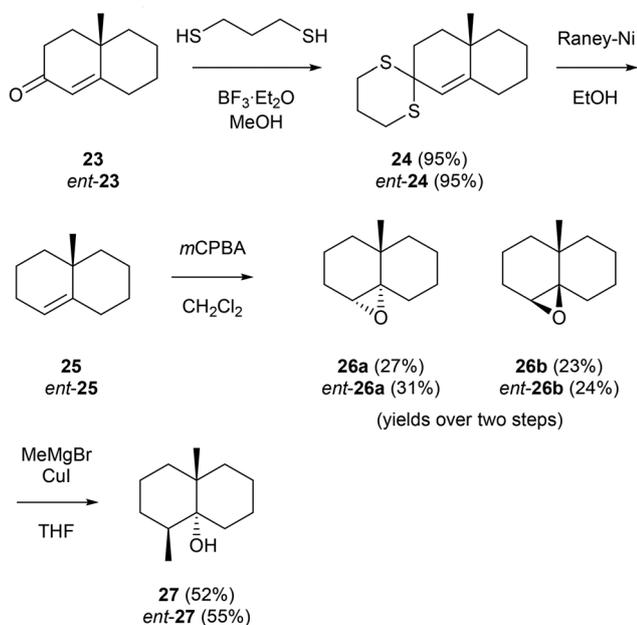
With (*S*)-(+)-methylbenzylamine in the stereoselective Robinson annelation *ent*-**18** can be obtained that was converted through analogous steps into *ent*-**22**. Taken together, **18** (*ent*-**18**) was converted into **22** (*ent*-**22**) with a yield of 21% (20%) over four steps. In this study, we also present the anomalous Cu-K $\alpha$  X-ray structures of **18**, *ent*-**18** and *ent*-**19** (Fig. S3, S4 and Tables S1–S3 $\dagger$ ).

The synthesis of 5,10-diepi-geosmin (**27**) commenced from enantiomerically pure octalone **23** obtained through stereoselective Robinson annelation and recrystallisation according to Reviel (Scheme 4 and Fig. S5 $\dagger$ ).<sup>45</sup> Conversion into the dithiane **24** with 1,3-propanedithiol and BF<sub>3</sub>·Et<sub>2</sub>O was followed by defunctionalisation to **25** with RANEY® nickel. Treatment with *m*CPBA furnished the known epoxides **26a** and **26b**<sup>49</sup> (dr  $\approx$  1.2 : 1) that were separated by column chromatography. The CuI catalysed addition of MeMgBr gave access to **27**. The same route was used to convert *ent*-**23**, accessible through Reviel's protocol using (*S*)-(+)-methylbenzylamine,<sup>45</sup> into *ent*-5,10-diepi-geosmin (*ent*-**27**). The overall yield of **27** (*ent*-**27**) from **23** (*ent*-**23**) is 13% (16%) over four steps.

The synthesis of 4,5-diepi-geosmin (**34**) started from 2,6-dimethylcyclohexanone (**28**) (Scheme 5). Generation of the sodium enolate with NaHMDS and subsequent treatment with allyl chloroformate resulted in the carbonate ester **29**. The Stoltz variation of the Tsuji–Trost allylation<sup>50</sup> that makes use of the chiral ligand **30** resulted in enantiomerically enriched **31a** and **31b** (dr  $\approx$  4 : 1) in favour of the desired diastereoisomer (as discussed below, the enantiomeric excess was determined from the final product). After separation by column chromatography, **31a** was converted with allylmagnesium bromide into **32** representing an inseparable mixture of dia-



Scheme 5 Synthesis of 4,5-diepi-geosmin (**34**) and *ent*-4,5-diepi-geosmin (*ent*-**34**).

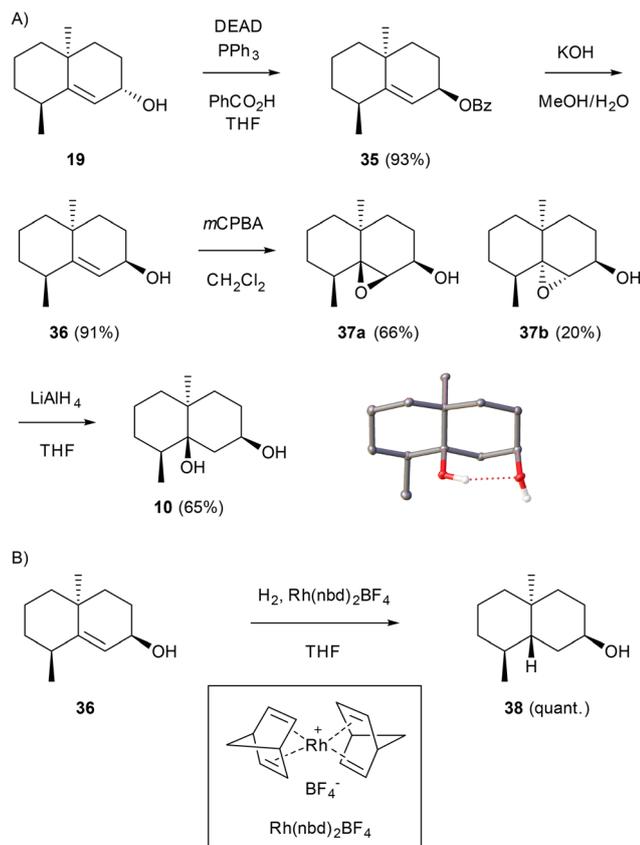


Scheme 4 Synthesis of 5,10-diepi-geosmin (**27**) and *ent*-5,10-diepi-geosmin (*ent*-**27**).

stereoisomers (dr  $\approx$  1 : 1.2). Cyclisation using the Grubbs I catalyst<sup>51</sup> resulted in the stereoisomers of dehydrogeosmin, **33a** and **33b**. Following their chromatographic separation, **33a** underwent catalytic hydrogenation to yield the target compound **34**, while **33b** yielded *ent*-**27**. Using the same approach, **29** was converted into *ent*-**34** employing the chiral ligand *ent*-**30** in the Stoltz allylation to *ent*-**31ab**. Taken together, the synthesis of **34** (*ent*-**34**) from **29** was performed with a yield of 16% (17%) over five steps.

The enantiomeric purities of synthetic **34** and *ent*-**34** were determined by gas chromatography on a chiral stationary phase, revealing an enantiomeric composition of 55% ee for **34** and of 54% ee for *ent*-**34** (Fig. S6 $\dagger$ ). The enantiomers of **34** were separable through HPLC on a chiral stationary phase, allowing to isolate both enantiomers in pure form without any noticeable presence of the other enantiomer (Fig. S7 $\dagger$ ).

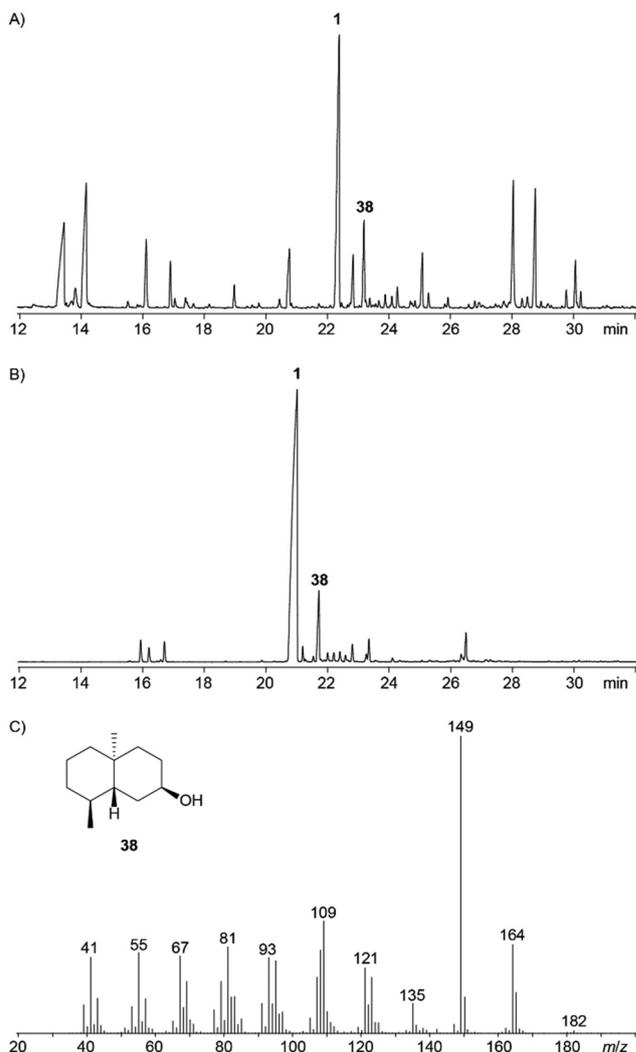
Besides the geosmin stereoisomers its hydroxylated derivative **10** was synthesised starting from **19** (Scheme 6A). As reported previously for the racemic material, Mitsunobu inversion of **19** to the benzoate ester **35** and saponification gave the alcohol **36**.<sup>52</sup> Its treatment with *m*CPBA yielded the epoxides **37a** and **37b** (dr  $\approx$  3 : 1) that were separable by column chromatography. The major stereoisomer **37a** underwent epoxide



**Scheme 6** Synthesis of **10** and the newly identified side product **38** of geosmin synthase.

opening to **10** with  $\text{LiAlH}_4$ . The synthetic material showed identical NMR data to the compound isolated from the mangrove endophyte<sup>35</sup> (Fig. S8–S15 and Table S4†) and the optical rotation of synthetic **10** ( $[\alpha]_{\text{D}}^{25} = -7.3$ ,  $c$  0.13, MeOH) in comparison to the optical rotation of natural **10** ( $[\alpha]_{\text{D}}^{22} = -6.2$ ,  $c$  0.1, MeOH)<sup>35</sup> established its absolute configuration as (4*S*,5*S*,7*R*,10*S*)-**10**. The absolute configuration of synthetic **10** reported herein was further confirmed through X-ray analysis using Cu- $\text{K}\alpha$  irradiation (Scheme 6 and Fig. S16, Table S5†).

As part of our continuous work on bacterial volatiles,<sup>53</sup> we noticed the presence of unknown constituents in headspace extracts from *Chondromyces crocatus* CmC5,<sup>9</sup> from *Streptomyces ambofaciens* ATCC 23877<sup>39</sup> and from *Streptomyces reticuli* DSM 101095 obtained using a closed-loop stripping apparatus (Fig. 1).<sup>54</sup> Their mass spectra showed a molecular ion at  $m/z$  182 and a fragment ion for the neutral loss of water ( $m/z$  164), suggesting that these compounds could be geosmin isomers. In our previous study this assumption was confirmed by the synthesis of a mixture of four diastereomers of 6,10-dimethylbicyclo[4.4.0]decan-3-ol, but a stereochemical assignment remained open.<sup>9</sup> To gain further insights into the structures of these compounds, the gene coding for the geosmin synthase in *S. ambofaciens* was cloned into the expression vector pYE-Express<sup>55</sup> and expressed in *Escherichia coli*. The purified recombinant



**Fig. 1** Detection of isogeosmin (**38**) in bacterial headspace extracts. Total ion chromatogram of headspace extracts from (A) *Streptomyces ambofaciens* and (B) *S. reticuli*. (C) EI mass spectrum of isogeosmin (**38**).

protein (Fig. S17†) indeed converted FPP into geosmin (**1**), several known geosmin synthase side products, and one of the unknown putative geosmin isomers (Fig. S18 and Table S6†). Compound isolation and structure elucidation by NMR spectroscopy (Fig. S19–S26 and Table S7†) revealed the structure of decalol **38**, for which we propose the name isogeosmin. An enantioselective synthesis of this compound was possible from **36** through catalytic hydrogenation, requiring the special  $\text{Rh}(\text{nbd})_2\text{BF}_4$  catalyst and a high  $\text{H}_2$  pressure (65 bar). A comparison of the optical rotations of **38** isolated from the enzymatic reaction ( $[\alpha]_{\text{D}}^{25} = +13.3$ ,  $c$  0.03,  $\text{CH}_2\text{Cl}_2$ ) and of synthetic **38** ( $[\alpha]_{\text{D}}^{25} = +12.3$ ,  $c$  0.13,  $\text{CH}_2\text{Cl}_2$ ) established the absolute configuration of (4*S*,5*R*,7*R*,10*S*)-**38**.

GC/MS analysis of geosmin and its stereoisomers revealed that the three compounds **1**, **22** and **34** show very similar EI mass spectra and GC retention indices ( $I$ ), while **27** elutes later and exhibits pronounced differences in the mass spectrum

with a base peak at  $m/z$  149 instead of  $m/z$  112 as observed for the other three stereoisomers (Fig. S27†). We assumed, if **22** or **34** would be side products of the geosmin synthase from *S. ambifaciens*, these compounds may be covered by the large peak of **1** in the GC/MS analysis of crude extracts from an enzyme reaction with FPP. However, a TLC analysis of the crude enzyme product in comparison to the reference standards of all for geosmin stereoisomers only confirmed the presence of **1**, but excluded the presence of the other three stereoisomers (Fig. S28†).

## Experimental

### Synthesis of **20a** and **20b**

To a solution of **19** (180 mg, 1.00 mmol) in  $\text{CH}_2\text{Cl}_2$  (18 mL) was added *m*-CPBA (232 mg, 70%, 1.34 mmol) at room temperature. The mixture was stirred for 1 h, followed by the addition of sat. aq. solutions of  $\text{Na}_2\text{SO}_3$  (10 mL) and of  $\text{NaHCO}_3$  (10 mL) and stirring for another 1 h. The reaction was quenched with sat. aq.  $\text{NH}_4\text{Cl}$  (30 mL) and extracted with  $\text{Et}_2\text{O}$  ( $3 \times 30$  mL). The combined organic layers were washed with brine (30 mL), dried over anhydrous  $\text{MgSO}_4$ , and concentrated under reduced pressure to give the crude product. Purification by silica gel column chromatography using petrol ether/ $\text{EtOAc}$  (10:1) afforded **20a** (131 mg, 0.67 mmol, 67%) and **20b** (57 mg, 0.29 mmol, 29%) as colourless needles.

Following the same procedure, *ent*-**19** (120 mg, 0.66 mmol) was converted into *ent*-**20a** (86 mg, 0.44 mmol, 67%) and *ent*-**20b** (40 mg, 0.20 mmol, 31%).

**Compound 20a.**  $^1\text{H}$  NMR (500 MHz,  $\text{C}_6\text{D}_6$ ):  $\delta_{\text{H}}$  3.81 (m, 1H), 3.14 (d,  $^3J_{\text{H,H}} = 4.9$  Hz, 1H), 1.93 (dq,  $^3J_{\text{H,H}} = 12.7$  Hz,  $^3J_{\text{H,H}} = 6.7$  Hz,  $^3J_{\text{H,H}} = 3.8$  Hz, 1H), 1.52–1.40 (m, 3H), 1.39–1.25 (m, 4H), 1.07–1.03 (m, 1H), 1.00 (s, 3H), 0.76–0.72 (m, 1H), 0.69–0.66 (m, 1H), 0.61 (d,  $^3J_{\text{H,H}} = 6.5$  Hz, 3H) ppm (Fig. S29†).  $^{13}\text{C}$  NMR (126 MHz,  $\text{C}_6\text{D}_6$ ):  $\delta_{\text{C}}$  69.5 (C), 63.9 (CH), 57.3 (CH), 36.4 ( $\text{CH}_2$ ), 34.8 ( $\text{CH}_2$ ), 34.2 (C), 31.7 (CH), 31.3 ( $\text{CH}_2$ ), 26.7 ( $\text{CH}_2$ ), 23.4 ( $\text{CH}_3$ ), 21.7 ( $\text{CH}_2$ ), 14.5 ( $\text{CH}_3$ ) ppm (Fig. S30†). HRMS (APCI):  $[\text{M} + \text{H}]^+$  calculated for  $\text{C}_{12}\text{H}_{21}\text{O}_2^+$   $m/z$  197.1536; found  $m/z$  197.1539. Optical rotation:  $[\alpha]_{\text{D}}^{25} = +18.3$  ( $c$  0.12,  $\text{CH}_2\text{Cl}_2$ ).

**Compound 20b.**  $^1\text{H}$  NMR (500 MHz,  $\text{C}_6\text{D}_6$ ):  $\delta_{\text{H}}$  3.81 (dd,  $^3J_{\text{H,H}} = 9.5$  Hz,  $^3J_{\text{H,H}} = 7.9$  Hz, 1H), 2.96 (br s, 1H), 1.89–1.79 (m, 1H), 1.64–1.58 (m, 1H), 1.53–1.37 (m, 6H), 1.29 (dddd,  $^2J_{\text{H,H}} = 14.8$  Hz,  $^3J_{\text{H,H}} = 13.2$  Hz,  $^3J_{\text{H,H}} = 9.6$  Hz,  $^3J_{\text{H,H}} = 3.5$  Hz, 1H), 1.21–1.17 (m, 1H), 0.93 (s, 3H), 0.75 (dt,  $^2J_{\text{H,H}} = 13.1$  Hz,  $^3J_{\text{H,H}} = 3.5$  Hz, 1H), 0.61 (d,  $^3J_{\text{H,H}} = 6.5$  Hz, 3H) ppm (Fig. S31†).  $^{13}\text{C}$  NMR (126 MHz,  $\text{C}_6\text{D}_6$ ):  $\delta_{\text{C}}$  67.5 (C), 66.5 (CH), 60.6 (CH), 37.7 ( $\text{CH}_2$ ), 34.2 (C), 33.4 ( $\text{CH}_2$ ), 32.3 ( $\text{CH}_2$ ), 30.1 (CH), 26.9 ( $\text{CH}_2$ ), 22.0 ( $\text{CH}_2$ ), 21.0 ( $\text{CH}_3$ ), 14.5 ( $\text{CH}_3$ ) ppm (Fig. S32†). HRMS (APCI):  $[\text{M} + \text{H}]^+$  calculated for  $\text{C}_{12}\text{H}_{21}\text{O}_2^+$   $m/z$  197.1536; found  $m/z$  197.1538. Optical rotation:  $[\alpha]_{\text{D}}^{25} = +47.2$  ( $c$  0.28,  $\text{CHCl}_3$ ), lit.  $[\alpha]_{\text{D}}^{22} = +48.3$  ( $c$  2.00,  $\text{CHCl}_3$ ).<sup>46</sup>

**Compound ent-20a.** HRMS (APCI):  $[\text{M} + \text{H}]^+$  calculated for  $\text{C}_{12}\text{H}_{21}\text{O}_2^+$   $m/z$  197.1536; found  $m/z$  197.1538. Optical rotation:  $[\alpha]_{\text{D}}^{25} = -17.4$  ( $c$  0.14,  $\text{CH}_2\text{Cl}_2$ ).

**Compound ent-20b.** HRMS (APCI):  $[\text{M} + \text{H}]^+$  calculated for  $\text{C}_{12}\text{H}_{21}\text{O}_2^+$   $m/z$  197.1536; found  $m/z$  197.1534. Optical rotation:  $[\alpha]_{\text{D}}^{25} = -49.1$  ( $c$  0.23,  $\text{CHCl}_3$ ).

### Synthesis of mesylate **21**

To a solution of **20a** (115 mg, 0.59 mmol) and  $\text{Et}_3\text{N}$  (119 mg, 1.18 mmol) in  $\text{CH}_2\text{Cl}_2$  (4 mL) was slowly added methanesulfonyl chloride (81 mg, 0.70 mmol, in 1 mL  $\text{CH}_2\text{Cl}_2$ ) at  $-20$  °C under Ar. The mixture was stirred for 2 h, then diluted with  $\text{Et}_2\text{O}$  (30 mL) and washed with sat.  $\text{NH}_4\text{Cl}$  (5 mL) and brine (5 mL). The organic layer was separated, dried over  $\text{MgSO}_4$ , and concentrated *in vacuo* to obtain the crude product. Purification by silica gel column chromatography using petrol ether/ $\text{EtOAc}$  (20:1) afforded **21** as a colourless oil (120 mg, 0.40 mmol, 68%).

Following the same procedure *ent*-**20a** (80 mg, 0.41 mmol) was converted into *ent*-**21** that was obtained as a colourless oil (55 mg, 0.20 mmol, 49%).

**Compound 21.**  $^1\text{H}$  NMR (500 MHz,  $\text{C}_6\text{D}_6$ ):  $\delta_{\text{H}}$  4.89 (ddd,  $^3J_{\text{H,H}} = 5.3$  Hz,  $^3J_{\text{H,H}} = 4.2$  Hz,  $^3J_{\text{H,H}} = 4.2$  Hz, 1H), 3.29 (d,  $^3J_{\text{H,H}} = 3.6$  Hz, 1H), 2.38 (s, 3H), 1.84 (dq,  $^3J_{\text{H,H}} = 12.6$  Hz,  $^3J_{\text{H,H}} = 6.7$  Hz,  $^3J_{\text{H,H}} = 4.0$  Hz, 1H), 1.66 (dddd,  $^2J_{\text{H,H}} = 14.4$  Hz,  $^3J_{\text{H,H}} = 7.2$  Hz,  $^3J_{\text{H,H}} = 4.4$  Hz,  $^3J_{\text{H,H}} = 3.2$  Hz, 1H), 1.48 (ddd,  $^2J_{\text{H,H}} = 14.1$  Hz,  $^3J_{\text{H,H}} = 11.3$  Hz,  $^3J_{\text{H,H}} = 3.1$  Hz, 1H), 1.43–1.38 (m, 1H), 1.35 (dddd,  $^2J_{\text{H,H}} = 11.2$  Hz,  $^3J_{\text{H,H}} = 11.2$  Hz,  $^3J_{\text{H,H}} = 5.6$  Hz,  $^3J_{\text{H,H}} = 3.0$  Hz, 1H), 1.28–1.19 (m, 2H), 1.06–0.99 (m, 2H), 0.95 (s, 3H), 0.74–0.68 (dddd,  $^2J_{\text{H,H}} = 12.7$ ,  $^3J_{\text{H,H}} = 12.7$ ,  $^3J_{\text{H,H}} = 12.0$ ,  $^3J_{\text{H,H}} = 5.3$ , 1H), 0.63 (ddd,  $^2J_{\text{H,H}} = 14.0$  Hz,  $^3J_{\text{H,H}} = 6.9$  Hz,  $^3J_{\text{H,H}} = 3.4$  Hz, 1H), 0.55 (d,  $^3J_{\text{H,H}} = 6.7$  Hz, 3H) ppm (Fig. S33†).  $^{13}\text{C}$  NMR (126 MHz,  $\text{C}_6\text{D}_6$ ):  $\delta_{\text{C}}$  75.9 (CH), 68.3 (C), 54.7 (CH), 38.6 ( $\text{CH}_3$ ), 37.1 ( $\text{CH}_2$ ), 34.4 (C), 33.7 ( $\text{CH}_2$ ), 32.9 ( $\text{CH}_2$ ), 31.7 (CH), 23.9 ( $\text{CH}_2$ ), 23.0 ( $\text{CH}_2$ ), 21.5 ( $\text{CH}_3$ ), 14.4 ( $\text{CH}_3$ ) ppm (Fig. S34†). Optical rotation:  $[\alpha]_{\text{D}}^{25} = +6.1$  ( $c$  0.05,  $\text{CH}_2\text{Cl}_2$ ).

**Compound ent-21.** Optical rotation:  $[\alpha]_{\text{D}}^{25} = -7.3$  ( $c$  0.11,  $\text{CH}_2\text{Cl}_2$ ).

### Synthesis of 5-*epi*-geosmin (**22**)

A solution of **21** (100 mg, 0.36 mmol) in THF (1 mL) was added dropwise to a suspension of lithium aluminum hydride (37 mg, 0.97 mmol) in THF (5 mL) at 0 °C under Ar. The mixture was refluxed for 3 h. After completion of the reaction, water (0.1 mL) was added dropwise, followed by 2 N NaOH (0.1 mL) and water (2 mL). The mixture was extracted with  $\text{Et}_2\text{O}$  ( $3 \times 10$  mL). The combined organic layers were washed with brine (10 mL), dried over  $\text{MgSO}_4$ , and concentrated *in vacuo* to afford the crude product. Purification by silica gel column chromatography using pentane/ $\text{Et}_2\text{O}$  (10:1) yielded **22** (30 mg, 0.16 mmol, 46%) as a colourless oil.

Following the same procedure, *ent*-**21** (50 mg, 0.18 mmol) was converted into *ent*-**22** that was obtained as a colourless oil (20 mg, 0.11 mmol, 60%).

**Compound 22.**  $^1\text{H}$  NMR (700 MHz,  $\text{C}_6\text{D}_6$ ):  $\delta_{\text{H}}$  1.76–1.67 (m, 3H), 1.59–1.52 (m, 1H), 1.47 (dddd,  $^2J_{\text{H,H}} = 13.5$  Hz,  $^3J_{\text{H,H}} = 13.5$  Hz,  $^3J_{\text{H,H}} = 13.5$  Hz,  $^3J_{\text{H,H}} = 4.3$  Hz,  $^3J_{\text{H,H}} = 4.3$  Hz, 1H), 1.45–1.31 (m, 6H), 1.28–1.21 (m, 1H), 1.04–0.99 (m, 2H), 0.94

(s, 3H), 0.89 (dddd,  $^2J_{\text{H,H}} = 13.6$  Hz,  $^3J_{\text{H,H}} = 4.2$  Hz,  $^3J_{\text{H,H}} = 2.2$  Hz,  $^3J_{\text{H,H}} = 2.2$  Hz, 1H), 0.79 (d,  $^3J_{\text{H,H}} = 6.8$  Hz, 3H) ppm (Fig. S35<sup>†</sup>).  $^{13}\text{C}$  NMR (176 MHz,  $\text{C}_6\text{D}_6$ ):  $\delta_{\text{C}}$  74.6 (C), 38.2 (C), 37.6 (CH), 37.3 ( $\text{CH}_2$ ), 34.3 ( $\text{CH}_2$ ), 32.6 ( $\text{CH}_2$ ), 26.2 ( $\text{CH}_2$ ), 22.4 ( $\text{CH}_3$ ), 22.0 ( $\text{CH}_2$ ), 21.8 ( $\text{CH}_2$ ), 21.5 ( $\text{CH}_3$ ), 15.1 ( $\text{CH}_3$ ) ppm (Fig. S36<sup>†</sup>). HRMS (APCI):  $[\text{M} - \text{H}_2\text{O} + \text{H}]^+$  calculated for  $\text{C}_{12}\text{H}_{21}^+$   $m/z$  165.1638; found  $m/z$  165.1634. Optical rotation:  $[\alpha]_{\text{D}}^{25} = +37.5$  ( $c$  0.04,  $\text{CH}_2\text{Cl}_2$ ).

**Compound ent-22.** HRMS (APCI):  $[\text{M} - \text{H}_2\text{O} + \text{H}]^+$  calculated for  $\text{C}_{12}\text{H}_{21}^+$   $m/z$  165.1638; found  $m/z$  165.1638. Optical rotation:  $[\alpha]_{\text{D}}^{25} = -35.1$  ( $c$  0.13,  $\text{CH}_2\text{Cl}_2$ ).

### Synthesis of 24

A mixture of **23**<sup>45</sup> (120 mg, 0.73 mmol), propane-1,3-dithiol (389 mg, 3.60 mmol), and  $\text{BF}_3 \cdot \text{Et}_2\text{O}$  (276 mg, 1.94 mmol) in 1.5 mL methanol was stirred at room temperature for 1 h. The reaction mixture was poured onto ice-cooled aq. NaOH (25 mL, 10% w/w). The aqueous layer was extracted with a 1 : 1 mixture of pentane/diethyl ether ( $3 \times 20$  mL). The combined organic layers were dried with  $\text{MgSO}_4$  and concentrated to dryness. The residue was purified through silica gel column chromatography with petrol ether/ $\text{Et}_2\text{O}$  (20 : 1) to afford **24** (177 mg, 0.70 mmol, 95%) as a colourless oil.

Following the same procedure, **ent-23** (200 mg, 1.22 mmol) was converted into **ent-24** that was obtained as a colourless oil (293 mg, 1.15 mmol, 95%).

**Compound 24.**  $^1\text{H}$  NMR (500 MHz,  $\text{C}_6\text{D}_6$ ):  $\delta_{\text{H}}$  5.68 (dd,  $^3J_{\text{H,H}} = 1.5$  Hz,  $^3J_{\text{H,H}} = 1.5$  Hz, 1H), 2.62 (ddd,  $^2J_{\text{H,H}} = 14.0$  Hz,  $^3J_{\text{H,H}} = 10.9$  Hz,  $^3J_{\text{H,H}} = 2.8$  Hz, 1H), 2.56 (ddd,  $^2J_{\text{H,H}} = 14.3$  Hz,  $^3J_{\text{H,H}} = 11.2$  Hz,  $^3J_{\text{H,H}} = 3.1$  Hz, 1H), 2.41–2.34 (m, 2H), 2.27 (dddd,  $^2J_{\text{H,H}} = 14.3$  Hz,  $^3J_{\text{H,H}} = 5.6$  Hz,  $^3J_{\text{H,H}} = 3.2$  Hz,  $^4J_{\text{H,H}} = 0.9$  Hz, 1H), 2.17 (ddd,  $^2J_{\text{H,H}} = 13.7$  Hz,  $^3J_{\text{H,H}} = 12.8$  Hz,  $^3J_{\text{H,H}} = 2.9$  Hz, 1H), 2.02 (dddd,  $^2J_{\text{H,H}} = 13.7$  Hz,  $^3J_{\text{H,H}} = 13.7$  Hz,  $^3J_{\text{H,H}} = 4.8$  Hz,  $^4J_{\text{H,H}} = 2.0$  Hz, 1H), 1.91 (ddd,  $^2J_{\text{H,H}} = 13.3$  Hz,  $^3J_{\text{H,H}} = 13.3$  Hz,  $^3J_{\text{H,H}} = 3.5$  Hz, 1H), 1.88–1.83 (m, 1H), 1.67–1.59 (m, 1H), 1.55–1.47 (m, 2H), 1.42 (ddd,  $^2J_{\text{H,H}} = 13.4$  Hz,  $^3J_{\text{H,H}} = 5.2$  Hz,  $^2J_{\text{H,H}} = 2.9$  Hz, 1H), 1.42–1.33 (m, 3H), 1.26–1.19 (m, 1H), 1.17–1.06 (m, 1H), 0.98 (s, 3H) ppm (Fig. S37<sup>†</sup>).  $^{13}\text{C}$  NMR (126 MHz,  $\text{C}_6\text{D}_6$ ):  $\delta_{\text{C}}$  147.6 (C), 123.0 (CH), 49.7 (C), 42.1 ( $\text{CH}_2$ ), 37.0 ( $\text{CH}_2$ ), 35.7 (C), 33.1 ( $\text{CH}_2$ ), 32.7 ( $\text{CH}_2$ ), 28.4 ( $\text{CH}_2$ ), 27.2 ( $\text{CH}_2$ ), 26.6 ( $\text{CH}_2$ ), 25.3 ( $\text{CH}_2$ ), 24.0 ( $\text{CH}_3$ ), 22.5 ( $\text{CH}_2$ ) ppm (Fig. S38<sup>†</sup>). HRMS (APCI):  $[\text{M} + \text{H}]^+$  calculated for  $\text{C}_{14}\text{H}_{23}\text{S}_2^+$   $m/z$  255.1236; found  $m/z$  255.1236. Optical rotation:  $[\alpha]_{\text{D}}^{25} = +250.4$  ( $c$  0.05,  $\text{CH}_2\text{Cl}_2$ ).

**Compound ent-24.** HRMS (APCI):  $[\text{M} + \text{H}]^+$  calculated for  $\text{C}_{14}\text{H}_{23}\text{S}_2^+$   $m/z$  255.1236; found  $m/z$  255.1236. Optical rotation:  $[\alpha]_{\text{D}}^{25} = -242.1$  ( $c$  0.12,  $\text{CH}_2\text{Cl}_2$ ).

### Synthesis of 26a and 26b

To a solution of **24** (170 mg, 0.67 mmol) in EtOH (10 mL) was added RANEY®-nickel (1.0 g, in 1 mL  $\text{H}_2\text{O}$ ). The mixture was heated to reflux for 4 h. The solids were filtered off and the filtrate was diluted with water (30 mL), followed by extraction with pentane ( $3 \times 15$  mL). The combined organic layers were washed with sat. brine (10 mL), dried with  $\text{MgSO}_4$  and concentrated to dryness. The residue was

purified through silica gel column chromatography with pentane to afford **25** (82 mg). Using the same procedure as for the epoxidation of **19**, the obtained material was converted into **26a** (29 mg, 0.18 mmol, 27% over two steps) and **26b** (26 mg, 0.15 mmol, 23% over two steps) that were obtained as colourless oils.

Following the same procedure, **ent-24** (290 mg, 1.14 mmol) was converted into **ent-26a** (59 mg, 0.35 mmol, 31% over two steps) and **ent-26b** (45 mg, 0.27 mmol, 24% over two steps) that were obtained as a colourless oils.

**Compound 26a.**<sup>49</sup>  $^1\text{H}$  NMR (700 MHz,  $\text{C}_6\text{D}_6$ ):  $\delta_{\text{H}}$  2.70 (d,  $^3J_{\text{H,H}} = 5.0$  Hz, 1H), 1.99 (ddd,  $^2J_{\text{H,H}} = 13.6$  Hz,  $^3J_{\text{H,H}} = 13.6$  Hz,  $^3J_{\text{H,H}} = 4.3$  Hz, 1H), 1.71 (ddd,  $^2J_{\text{H,H}} = 15.2$  Hz,  $^3J_{\text{H,H}} = 11.8$  Hz,  $^3J_{\text{H,H}} = 7.0$  Hz, 1H), 1.60 (dddd,  $^2J_{\text{H,H}} = 15.2$  Hz,  $^3J_{\text{H,H}} = 6.7$  Hz,  $^3J_{\text{H,H}} = 5.0$  Hz,  $^3J_{\text{H,H}} = 2.3$  Hz,  $^3J_{\text{H,H}} = 1.4$  Hz, 1H), 1.58–1.53 (m, 1H), 1.51–1.44 (m, 2H), 1.38–1.32 (m, 2H), 1.22 (dddd,  $^2J_{\text{H,H}} = 13.6$  Hz,  $^3J_{\text{H,H}} = 13.6$  Hz,  $^3J_{\text{H,H}} = 11.5$  Hz,  $^3J_{\text{H,H}} = 6.6$  Hz,  $^3J_{\text{H,H}} = 3.2$  Hz, 1H), 1.16–1.07 (m, 3H), 1.11 (s, 3H), 0.92 (ddd,  $^2J_{\text{H,H}} = 13.2$  Hz,  $^3J_{\text{H,H}} = 3.2$  Hz,  $^3J_{\text{H,H}} = 3.2$  Hz, 1H), 0.84–0.81 (m, 1H) ppm (Fig. S39<sup>†</sup>).  $^{13}\text{C}$  NMR (176 MHz,  $\text{C}_6\text{D}_6$ ):  $\delta_{\text{C}}$  63.5 (C), 60.9 (CH), 36.0 ( $\text{CH}_2$ ), 35.4 ( $\text{CH}_2$ ), 33.8 (C), 32.4 ( $\text{CH}_2$ ), 26.3 ( $\text{CH}_2$ ), 24.9 ( $\text{CH}_2$ ), 23.4 ( $\text{CH}_3$ ), 22.0 ( $\text{CH}_2$ ), 16.8 ( $\text{CH}_2$ ) ppm (Fig. S40<sup>†</sup>). HRMS (APCI):  $[\text{M} + \text{H}]^+$  calculated for  $\text{C}_{11}\text{H}_{19}\text{O}^+$   $m/z$  167.1430; found  $m/z$  167.1431. Optical rotation:  $[\alpha]_{\text{D}}^{25} = +10.0$  ( $c$  0.03,  $\text{CH}_2\text{Cl}_2$ ).

**Compound 26b.**<sup>49</sup>  $^1\text{H}$  NMR (700 MHz,  $\text{C}_6\text{D}_6$ ):  $\delta_{\text{H}}$  2.66 (d,  $^3J_{\text{H,H}} = 4.1$  Hz, 1H), 1.86–1.76 (m, 3H), 1.71 (ddd,  $^2J_{\text{H,H}} = 13.6$  Hz,  $^3J_{\text{H,H}} = 13.6$  Hz,  $^3J_{\text{H,H}} = 3.9$  Hz, 1H), 1.66–1.58 (m, 1H), 1.58–1.46 (m, 4H), 1.38–1.28 (m, 1H), 1.27–1.23 (m, 1H), 1.20–1.12 (m, 1H), 0.91 (s, 3H), 0.89–0.85 (m, 1H), 0.78 (ddd,  $^2J_{\text{H,H}} = 12.9$  Hz,  $^3J_{\text{H,H}} = 12.9$  Hz,  $^3J_{\text{H,H}} = 3.7$  Hz, 1H) ppm (Fig. S41<sup>†</sup>).  $^{13}\text{C}$  NMR (176 MHz,  $\text{C}_6\text{D}_6$ ):  $\delta_{\text{C}}$  64.6 (C), 59.8 (CH), 38.1 ( $\text{CH}_2$ ), 34.6 ( $\text{CH}_2$ ), 33.5 (C), 30.6 ( $\text{CH}_2$ ), 24.3 ( $\text{CH}_2$ ), 22.9 ( $\text{CH}_2$ ), 22.0 ( $\text{CH}_2$ ), 20.6 ( $\text{CH}_3$ ), 16.1 ( $\text{CH}_2$ ) ppm (Fig. S42<sup>†</sup>). HRMS (APCI):  $[\text{M} + \text{H}]^+$  calculated for  $\text{C}_{11}\text{H}_{19}\text{O}^+$   $m/z$  167.1430; found  $m/z$  167.1427. Optical rotation:  $[\alpha]_{\text{D}}^{25} = +73.6$  ( $c$  0.15,  $\text{CH}_2\text{Cl}_2$ ).

**Compound ent-26a.** HRMS (APCI):  $[\text{M} + \text{H}]^+$  calculated for  $\text{C}_{11}\text{H}_{19}\text{O}^+$   $m/z$  167.1430; found  $m/z$  167.1427. Optical rotation:  $[\alpha]_{\text{D}}^{25} = -10.1$  ( $c$  0.11,  $\text{CH}_2\text{Cl}_2$ ).

**Compound ent-26b.** HRMS (APCI):  $[\text{M} + \text{H}]^+$  calculated for  $\text{C}_{11}\text{H}_{19}\text{O}^+$   $m/z$  167.1430; found  $m/z$  167.1427. Optical rotation:  $[\alpha]_{\text{D}}^{25} = -74.3$  ( $c$  0.14,  $\text{CH}_2\text{Cl}_2$ ).

### Synthesis of 5,10-diepi-geosmin (27)

To a solution of methylmagnesium bromide (0.42 mL, 3 M in  $\text{Et}_2\text{O}$ , 1.26 mmol) were added THF (3 mL) and CuI (8 mg, 0.04 mmol) under Ar. A solution of **26a** (7 mg, 0.04 mmol) in THF (0.5 mL) was added dropwise. The mixture was refluxed for 2 h and then quenched by the addition of sat. aq.  $\text{NH}_4\text{Cl}$  (5 mL). The mixture was extracted with  $\text{Et}_2\text{O}$  ( $3 \times 10$  mL). The combined organic layers were washed with brine (10 mL), dried over  $\text{MgSO}_4$ , and concentrated to give the crude product. Purification by silica gel column chromatography with pentane/ $\text{Et}_2\text{O}$  (10 : 1) afforded **27** (4 mg, 0.02 mmol, 52%) as a colourless oil.

Following the same procedure *ent-26a* (10 mg, 0.06 mmol) was converted into *ent-27* that was obtained as a colourless oil (6 mg, 0.03 mmol, 55%).

**Compound 27.**  $^1\text{H}$  NMR (700 MHz,  $\text{C}_6\text{D}_6$ ):  $\delta_{\text{H}}$  2.12 (dddd,  $^2J_{\text{H,H}} = 13.7$  Hz,  $^3J_{\text{H,H}} = 13.7$  Hz,  $^3J_{\text{H,H}} = 5.1$  Hz,  $^3J_{\text{H,H}} = 5.1$  Hz, 1H), 1.85 (ddd,  $^2J_{\text{H,H}} = 13.0$  Hz,  $^3J_{\text{H,H}} = 13.0$  Hz,  $^3J_{\text{H,H}} = 4.5$  Hz, 1H), 1.81–1.72 (m, 3H), 1.71–1.65 (m, 1H), 1.50–1.42 (m, 3H), 1.39–1.33 (m, 2H), 1.22 (br d,  $^2J_{\text{H,H}} = 13.2$  Hz, 1H), 0.97 (s, 3H), 0.95–0.92 (m, 1H), 0.91 (d,  $^3J_{\text{H,H}} = 7.9$  Hz, 3H), 0.90–0.82 (m, 2H) ppm (Fig. S43<sup>†</sup>).  $^{13}\text{C}$  NMR (176 MHz,  $\text{C}_6\text{D}_6$ ):  $\delta_{\text{C}}$  74.8 (C), 41.6 (CH), 38.5 ( $\text{CH}_2$ ), 37.3 ( $\text{CH}_2$ ), 35.6 ( $\text{CH}_2$ ), 32.8 ( $\text{CH}_2$ ), 28.6 ( $\text{CH}_2$ ), 21.8 ( $\text{CH}_2$ ), 21.7 ( $\text{CH}_3$ ), 21.4 ( $\text{CH}_2$ ), 17.6 ( $\text{CH}_2$ ), 16.9 ( $\text{CH}_3$ ) ppm (Fig. S44<sup>†</sup>). HRMS (APCI):  $[\text{M} - \text{H}_2\text{O} + \text{H}]^+$  calculated for  $\text{C}_{12}\text{H}_{21}^+$   $m/z$  165.1638; found  $m/z$  165.1638. Optical rotation:  $[\alpha]_{\text{D}}^{25} = +76.0$  ( $c$  0.15,  $\text{CH}_2\text{Cl}_2$ ).

**Compound *ent-27*.** HRMS (APCI):  $[\text{M} - \text{H}_2\text{O} + \text{H}]^+$  calculated for  $\text{C}_{12}\text{H}_{21}^+$   $m/z$  165.1638; found  $m/z$  165.1636. Optical rotation:  $[\alpha]_{\text{D}}^{25} = -73.5$  ( $c$  0.24,  $\text{CH}_2\text{Cl}_2$ ).

### Synthesis of 29

A solution of 2,6-dimethylcyclohexan-1-one (511 mg, 4.05 mmol) in THF (7 mL) was added dropwise to a solution of NaHMDS (4.9 mmol) in THF (30 mL) at 0 °C under Ar atmosphere. The mixture was stirred for 10 min and the cooled to –78 °C. A solution of allyl chloroformate (634 mg, 5.26 mmol) in THF (15 mL) was added dropwise. The colour of the solution turned to deep red. The reaction mixture was slowly warmed to room temperature overnight. The resulting yellow suspension was quenched by the addition of water (10 mL),  $\text{NH}_4\text{Cl}$  solution (sat. 10 mL) and hexane (10 mL). The biphasic mixture was stirred for 30 min before the layers were separated. The aqueous layer was extracted with  $\text{Et}_2\text{O}$  (3 × 20 mL) and the combined organic layers were washed with brine (60 mL), dried over  $\text{MgSO}_4$ , filtrated and concentrated *in vacuo*. Column chromatography on silica gel with petrol ether/ $\text{Et}_2\text{O}$  (100 : 1) produced **23** (624 mg, 3.02 mmol, 75%) as a colourless oil.

**Compound 29.**<sup>56</sup>  $^1\text{H}$  NMR (500 MHz,  $\text{C}_6\text{D}_6$ ):  $\delta_{\text{H}}$  5.70 (ddt,  $^3J_{\text{H,H}} = 17.2$  Hz,  $^3J_{\text{H,H}} = 10.5$  Hz,  $^3J_{\text{H,H}} = 5.6$  Hz, 1H), 5.14 (ddt,  $^2J_{\text{H,H}} = 17.2$  Hz,  $^4J_{\text{H,H}} = 1.5$  Hz,  $^2J_{\text{H,H}} = 1.5$  Hz, 1H), 4.94 (ddt,  $^2J_{\text{H,H}} = 10.5$  Hz,  $^4J_{\text{H,H}} = 1.5$  Hz,  $^2J_{\text{H,H}} = 1.5$  Hz, 1H), 4.44 (dm,  $^3J_{\text{H,H}} = 5.6$  Hz, 2H), 2.59 (m, 1H), 1.80–1.75 (m, 2H), 1.66 (dddd,  $^2J_{\text{H,H}} = 12.8$  Hz,  $^3J_{\text{H,H}} = 8.8$  Hz,  $^3J_{\text{H,H}} = 5.8$  Hz,  $^3J_{\text{H,H}} = 3.1$  Hz, 1H), 1.57 (dt,  $^5J_{\text{H,H}} = 1.9$  Hz,  $^4J_{\text{H,H}} = 1.0$  Hz, 3H), 1.47–1.38 (m, 1H), 1.38–1.29 (m, 1H), 1.26–1.19 (m, 1H), 1.07 (d,  $^3J_{\text{H,H}} = 7.0$  Hz, 3H) ppm (Fig. S45<sup>†</sup>);  $^{13}\text{C}$  NMR (126 MHz,  $\text{C}_6\text{D}_6$ ):  $\delta_{\text{C}}$  153.6 (C), 146.6 (C), 132.2 (CH), 121.0 (C), 118.1 ( $\text{CH}_2$ ), 68.3 ( $\text{CH}_2$ ), 32.0 ( $\text{CH}_2$ ), 31.8 (CH), 30.8 ( $\text{CH}_2$ ), 20.3 ( $\text{CH}_2$ ), 18.4 ( $\text{CH}_3$ ), 16.2 ( $\text{CH}_3$ ) ppm (Fig. S46<sup>†</sup>).

### Synthesis of 31a and 31b

Bis(3,5,3',5'-dimethoxydibenzylideneacetone)palladium(0) ( $\text{Pd}(\text{dmdba})_2$ , 19 mg, 0.02 mmol) and (*S*)-4-*tert*-butyl-2-[2-(diphenylphosphino)phenyl]-2-oxazoline ((*S*)-*t*-Bu-PHOX, 11 mg, 0.03 mmol) were dissolved in benzene (6 mL, degassed by an argon stream for 1.5 h) under argon atmosphere at room temperature.

After stirring the brown-yellow suspension for 30 min at room temperature and cooling to 10 °C, a solution of **29** (49 mg, 0.23 mmol) in degassed benzene (0.5 mL) was added dropwise. The resulting mixture was stirred for 15 h at 10 °C and then concentrated under reduced pressure. The residue was purified by silica gel column chromatography with petroleum ether/ $\text{Et}_2\text{O}$  (200 : 1) to yield **31a** (20 mg, 0.12 mmol, 52%) and **31b** (5 mg, 0.03 mmol, 13%) as colourless oil.

Following the same procedure with ((*R*)-*t*-Bu-PHOX), **29** (49 mg, 0.23 mmol) was converted into *ent-31a* (21 mg, 0.13 mmol, 55%) and *ent-31b* (4 mg, 0.02 mmol, 10%) that were obtained as a colourless oils.

**Compound 31a.**<sup>57</sup>  $^1\text{H}$  NMR (500 MHz,  $\text{C}_6\text{D}_6$ ):  $\delta_{\text{H}}$  5.52 (dddd,  $^3J_{\text{H,H}} = 17.1$  Hz,  $^3J_{\text{H,H}} = 10.2$  Hz,  $^3J_{\text{H,H}} = 7.6$  Hz,  $^3J_{\text{H,H}} = 7.0$  Hz, 1H), 4.95 (dddd,  $^3J_{\text{H,H}} = 10.1$  Hz,  $^2J_{\text{H,H}} = 2.1$  Hz,  $^4J_{\text{H,H}} = 1.1$  Hz,  $^4J_{\text{H,H}} = 1.1$  Hz, 1H), 4.92 (dddd,  $^3J_{\text{H,H}} = 17.0$  Hz,  $^2J_{\text{H,H}} = 2.1$  Hz,  $^4J_{\text{H,H}} = 1.4$  Hz,  $^4J_{\text{H,H}} = 1.4$  Hz, 1H), 2.26 (dd,  $^2J_{\text{H,H}} = 14.1$  Hz,  $^3J_{\text{H,H}} = 7.0$  Hz, 1H), 2.19 (dq,  $^3J_{\text{H,H}} = 12.9$  Hz,  $^3J_{\text{H,H}} = 6.5$  Hz,  $^3J_{\text{H,H}} = 5.6$  Hz, 1H), 1.87 (dddd,  $^2J_{\text{H,H}} = 14.1$  Hz,  $^3J_{\text{H,H}} = 7.6$  Hz,  $^4J_{\text{H,H}} = 1.3$  Hz,  $^4J_{\text{H,H}} = 1.3$  Hz, 1H), 1.66–1.57 (m, 1H), 1.56–1.41 (m, 2H), 1.24–1.16 (m, 2H), 1.08 (s, 3H), 1.03 (d,  $^3J_{\text{H,H}} = 6.5$  Hz, 3H), 1.05–0.99 (m, 1H) ppm (Fig. S47<sup>†</sup>).  $^{13}\text{C}$  NMR (126 MHz,  $\text{C}_6\text{D}_6$ ):  $\delta_{\text{C}}$  213.7 (C), 133.8 (CH), 117.7 ( $\text{CH}_2$ ), 48.6 (C), 41.9 ( $\text{CH}_2$ ), 41.3 (CH), 40.0 ( $\text{CH}_2$ ), 36.7 ( $\text{CH}_2$ ), 23.0 ( $\text{CH}_3$ ), 21.3 ( $\text{CH}_2$ ), 15.5 ( $\text{CH}_3$ ) ppm (Fig. S48<sup>†</sup>). HRMS (APCI):  $[\text{M} + \text{H}]^+$  calculated for  $\text{C}_{11}\text{H}_{19}\text{O}^+$   $m/z$  167.1430; found  $m/z$  167.1428. Optical rotation:  $[\alpha]_{\text{D}}^{25} = -68.1$  ( $c$  0.12,  $\text{CH}_2\text{Cl}_2$ ).

**Compound 31b.**<sup>58</sup>  $^1\text{H}$  NMR (500 MHz,  $\text{C}_6\text{D}_6$ ):  $\delta_{\text{H}}$  5.92 (dddd,  $^3J_{\text{H,H}} = 17.1$  Hz,  $^3J_{\text{H,H}} = 10.1$  Hz,  $^3J_{\text{H,H}} = 7.8$  Hz,  $^3J_{\text{H,H}} = 7.0$  Hz, 1H), 5.04 (dddd,  $^3J_{\text{H,H}} = 10.1$  Hz,  $^2J_{\text{H,H}} = 2.4$  Hz,  $^4J_{\text{H,H}} = 1.0$  Hz,  $^4J_{\text{H,H}} = 1.0$  Hz, 1H), 5.01 (dddd,  $^3J_{\text{H,H}} = 17.1$  Hz,  $^2J_{\text{H,H}} = 2.4$  Hz,  $^4J_{\text{H,H}} = 1.4$  Hz,  $^4J_{\text{H,H}} = 1.4$  Hz, 1H), 2.35 (dddd,  $^2J_{\text{H,H}} = 13.8$  Hz,  $^3J_{\text{H,H}} = 7.0$  Hz,  $^4J_{\text{H,H}} = 1.0$  Hz,  $^4J_{\text{H,H}} = 1.0$  Hz, 1H), 2.29 (dddd,  $^2J_{\text{H,H}} = 13.8$  Hz,  $^3J_{\text{H,H}} = 7.8$  Hz,  $^4J_{\text{H,H}} = 1.0$  Hz,  $^4J_{\text{H,H}} = 1.0$  Hz, 1H), 2.19 (dq,  $^3J_{\text{H,H}} = 12.9$  Hz,  $^3J_{\text{H,H}} = 6.5$  Hz,  $^3J_{\text{H,H}} = 5.5$  Hz, 1H), 1.60–1.55 (m, 1H), 1.44–1.32 (m, 3H), 1.28–1.23 (m, 1H), 1.01 (d,  $^3J_{\text{H,H}} = 6.5$  Hz, 3H), 0.99–0.93 (m, 1H), 0.86 (s, 3H) ppm (Fig. S49<sup>†</sup>).  $^{13}\text{C}$  NMR (126 MHz,  $\text{C}_6\text{D}_6$ ):  $\delta_{\text{C}}$  214.2 (C), 135.9 (CH), 117.5 ( $\text{CH}_2$ ), 47.8 (C), 43.6 ( $\text{CH}_2$ ), 41.1 (CH), 38.8 ( $\text{CH}_2$ ), 36.5 ( $\text{CH}_2$ ), 23.1 ( $\text{CH}_3$ ), 21.5 ( $\text{CH}_2$ ), 15.4 ( $\text{CH}_3$ ) ppm (Fig. S50<sup>†</sup>). HRMS (APCI):  $[\text{M} + \text{H}]^+$  calculated for  $\text{C}_{11}\text{H}_{19}\text{O}^+$   $m/z$  167.1430; found  $m/z$  167.1428. Optical rotation:  $[\alpha]_{\text{D}}^{25} = +11.3$  ( $c$  0.14,  $\text{CH}_2\text{Cl}_2$ ).

**Compound *ent-31a*.** HRMS (APCI):  $[\text{M} + \text{H}]^+$  calculated for  $\text{C}_{11}\text{H}_{19}\text{O}^+$   $m/z$  167.1430; found  $m/z$  167.1428. Optical rotation:  $[\alpha]_{\text{D}}^{25} = +71.4$  ( $c$  0.22,  $\text{CH}_2\text{Cl}_2$ ).

**Compound *ent-31b*.** HRMS (APCI):  $[\text{M} + \text{H}]^+$  calculated for  $\text{C}_{11}\text{H}_{19}\text{O}^+$   $m/z$  167.1430; found  $m/z$  167.1429. Optical rotation:  $[\alpha]_{\text{D}}^{25} = -13.8$  ( $c$  0.25,  $\text{CH}_2\text{Cl}_2$ ).

### Synthesis of 33a and 33b

A solution of **31a** (8 mg, 0.05 mmol) in THF (2 mL) was added allylmagnesium bromide (0.06 mL, 1.0 M in  $\text{Et}_2\text{O}$ , 0.06 mmol) at 40 °C under Ar. The mixture was stirred for 1 h. The reaction was quenched with sat.  $\text{NH}_4\text{Cl}$  (3 mL) and extracted with  $\text{Et}_2\text{O}$  (3 × 5 mL). The combined organic layers were washed with

brine (3 mL), dried over anhydrous MgSO<sub>4</sub>, and concentrated under reduced pressure to give the crude product. Purification by silica gel column chromatography using petrol ether/Et<sub>2</sub>O (20 : 1) afforded **32** as a mixture of diastereoisomers (11 mg, dr ≈ 1 : 1.2). The diastereomeric mixture of **32** was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (4 mL), followed by the dropwise addition of Grubbs I catalyst (5 mg, 0.01 mmol) dissolved in CH<sub>2</sub>Cl<sub>2</sub> (0.5 mL). The mixture was stirred for 4 h and then concentrated to dryness. The residue was purified through column chromatography on AgNO<sub>3</sub> impregnated silica gel with petrol ether/Et<sub>2</sub>O (30 : 1–5 : 1) to afford **33a** (4 mg, 0.02 mmol, 42%) and **33b** (5 mg, 0.03 mmol, 53%) as colourless oils.

Following the same procedure, *ent*-**31a** (20 mg, 0.12 mmol) was converted into *ent*-**33a** (7 mg, 0.04 mmol, 40%) and *ent*-**33b** (9 mg, 0.05 mmol, 51%) that were obtained as a colourless oils.

**Compound 33a.** <sup>1</sup>H NMR (700 MHz, C<sub>6</sub>D<sub>6</sub>): δ<sub>H</sub> 5.51–5.46 (m, 1H), 5.44–5.40 (m, 1H), 2.36 (br d, <sup>2</sup>J<sub>H,H</sub> = 18.4 Hz, 1H), 2.16 (ddm, <sup>2</sup>J<sub>H,H</sub> = 17.7 Hz, <sup>3</sup>J<sub>H,H</sub> = 5.6 Hz, 1H), 1.91 (dddd, <sup>2</sup>J<sub>H,H</sub> = 17.6 Hz, <sup>3</sup>J<sub>H,H</sub> = 7.1 Hz, <sup>4</sup>J<sub>H,H</sub> = 2.4 Hz, <sup>5</sup>J<sub>H,H</sub> = 2.4 Hz, <sup>5</sup>J<sub>H,H</sub> = 2.4 Hz, 1H), 1.70 (ddd, <sup>2</sup>J<sub>H,H</sub> = 13.2 Hz, <sup>3</sup>J<sub>H,H</sub> = 13.2 Hz, <sup>3</sup>J<sub>H,H</sub> = 4.7 Hz, 1H), 1.61–1.54 (m, 1H), 1.48–1.33 (m, 4H), 1.23–1.19 (m, 1H), 0.97 (d, <sup>4</sup>J<sub>H,H</sub> = 0.9 Hz, 3H), 0.96–0.93 (m, 1H), 0.80 (d, <sup>3</sup>J<sub>H,H</sub> = 6.6 Hz, 3H) ppm (Fig. S51†). <sup>13</sup>C NMR (176 MHz, C<sub>6</sub>D<sub>6</sub>): δ<sub>C</sub> 125.7 (CH), 124.3 (CH), 73.3 (C), 37.0 (C), 35.7 (CH<sub>2</sub>), 34.0 (CH<sub>2</sub>), 33.4 (CH<sub>2</sub>), 32.6 (CH), 30.2 (CH<sub>2</sub>), 23.7 (CH<sub>3</sub>), 21.6 (CH<sub>2</sub>), 15.9 (CH<sub>3</sub>) ppm (Fig. S52†). HRMS (APCI): [M + H]<sup>+</sup> calculated for C<sub>12</sub>H<sub>21</sub>O<sup>+</sup> *m/z* 181.1587; found *m/z* 181.1584. Optical rotation: [α]<sub>D</sub><sup>25</sup> = –6.5 (c 0.13, CH<sub>2</sub>Cl<sub>2</sub>).

**Compound 33b.** <sup>1</sup>H NMR (700 MHz, C<sub>6</sub>D<sub>6</sub>): δ<sub>H</sub> 5.57–5.53 (m, 1H), 5.44–5.41 (m, 1H), 2.33 (dddd, <sup>2</sup>J<sub>H,H</sub> = 17.8 Hz, <sup>3</sup>J<sub>H,H</sub> = 6.8 Hz, <sup>4</sup>J<sub>H,H</sub> = 2.3 Hz, <sup>5</sup>J<sub>H,H</sub> = 2.3 Hz, <sup>5</sup>J<sub>H,H</sub> = 2.3 Hz, 1H), 2.31–2.25 (m, 1H), 2.09 (br d, <sup>2</sup>J<sub>H,H</sub> = 17.8 Hz, 1H), 1.80–1.72 (m, 2H), 1.59 (dddd, <sup>2</sup>J<sub>H,H</sub> = 13.8 Hz, <sup>3</sup>J<sub>H,H</sub> = 13.8 Hz, <sup>3</sup>J<sub>H,H</sub> = 13.8 Hz, <sup>3</sup>J<sub>H,H</sub> = 3.5 Hz, <sup>3</sup>J<sub>H,H</sub> = 3.5 Hz, 1H), 1.47 (br dd, <sup>2</sup>J<sub>H,H</sub> = 17.8 Hz, <sup>3</sup>J<sub>H,H</sub> = 4.0 Hz, 1H), 1.43–1.35 (m, 2H), 1.31–1.28 (m, 1H), 1.08–1.04 (m, 1H), 0.95 (d, <sup>3</sup>J<sub>H,H</sub> = 7.3 Hz, 3H), 0.92 (s, 3H) ppm (Fig. S53†). <sup>13</sup>C NMR (176 MHz, C<sub>6</sub>D<sub>6</sub>): δ<sub>C</sub> 126.6 (CH), 124.3 (CH), 73.8 (C), 41.1 (C), 39.2 (CH), 36.5 (C), 36.3 (CH<sub>2</sub>), 35.4 (CH<sub>2</sub>), 29.2 (CH<sub>2</sub>), 22.5 (CH<sub>3</sub>), 17.7 (CH<sub>2</sub>), 17.5 (CH<sub>3</sub>) ppm (Fig. S54†). HRMS (APCI): [M + H]<sup>+</sup> calculated for C<sub>12</sub>H<sub>21</sub>O<sup>+</sup> *m/z* 181.1587; found *m/z* 181.1588. Optical rotation: [α]<sub>D</sub><sup>25</sup> = –4.3 (c 0.08, CH<sub>2</sub>Cl<sub>2</sub>).

**Compound ent-33a.** HRMS (APCI): [M – H<sub>2</sub>O + H]<sup>+</sup> calculated for C<sub>12</sub>H<sub>19</sub><sup>+</sup> *m/z* 163.1481; found *m/z* 163.1479. Optical rotation: [α]<sub>D</sub><sup>25</sup> = +5.6 (c 0.22, CH<sub>2</sub>Cl<sub>2</sub>).

**Compound ent-33b.** HRMS (APCI): [M – H<sub>2</sub>O + H]<sup>+</sup> calculated for C<sub>12</sub>H<sub>19</sub><sup>+</sup> *m/z* 163.1481; found *m/z* 163.1480. Optical rotation: [α]<sub>D</sub><sup>25</sup> = +3.3 (c 0.25, CH<sub>2</sub>Cl<sub>2</sub>).

### Synthesis of 4,5-diepi-geosmin (**34**)

A mixture of **33a** (3 mg, 0.02 mmol) and Pd/C (0.5 mg, 5% Pd) in CH<sub>3</sub>OH (2 mL) was stirred in a H<sub>2</sub> atmosphere (10 bar) for 2 h. At the end of the reaction the catalyst was removed by filtration and the solvents were evaporated to obtain the product 4,5-diepi-geosmin (**34**) as a colourless oil (3 mg, 0.02 mmol, 100%).

Following the same procedure, (*ent*)-**33a** (3 mg, 0.02 mmol) was converted into *ent*-**34** that was obtained as a colourless oil (3 mg, 0.02 mmol, 100%), and **33b** (8 mg, 0.04 mmol) was converted into *ent*-**27** (8 mg, 0.04 mmol, 100%).

**Compound 34.** <sup>59</sup> <sup>1</sup>H NMR (700 MHz, C<sub>6</sub>D<sub>6</sub>): δ<sub>H</sub> 1.80–1.69 (m, 3H), 1.55–1.51 (m, 1H), 1.49–1.41 (m, 3H), 1.40–1.31 (m, 4H), 1.29–1.26 (m, 1H), 1.18–1.10 (m, 1H), 1.03–1.00 (m, 1H), 0.97 (s, 3H), 0.90–0.87 (m, 1H), 0.78 (d, <sup>3</sup>J<sub>H,H</sub> = 6.7 Hz, 3H) ppm (Fig. S55†). <sup>13</sup>C NMR (176 MHz, C<sub>6</sub>D<sub>6</sub>): δ<sub>C</sub> 74.3 (C), 38.3 (C), 37.1 (CH<sub>2</sub>), 33.5 (CH<sub>2</sub>), 32.0 (CH), 31.8 (CH<sub>2</sub>), 30.9 (CH<sub>2</sub>), 23.3 (CH<sub>2</sub>), 22.9 (CH<sub>3</sub>), 21.9 (CH<sub>2</sub>), 21.8 (CH<sub>2</sub>), 15.3 (CH<sub>3</sub>) ppm (Fig. S56†). HRMS (APCI): [M – H<sub>2</sub>O + H]<sup>+</sup> calculated for C<sub>12</sub>H<sub>21</sub><sup>+</sup> *m/z* 165.1634; found *m/z* 165.1638. Optical rotation: [α]<sub>D</sub><sup>25</sup> = –5.0 (c 0.05, CH<sub>2</sub>Cl<sub>2</sub>).

**Compound ent-34.** HRMS (APCI): [M – H<sub>2</sub>O + H]<sup>+</sup> calculated for C<sub>12</sub>H<sub>21</sub><sup>+</sup> *m/z* 165.1638; found *m/z* 165.1635. Optical rotation: [α]<sub>D</sub><sup>25</sup> = +5.4 (c 0.08, CH<sub>2</sub>Cl<sub>2</sub>).

**Compound ent-27.** Optical rotation: [α]<sub>D</sub><sup>25</sup> = –70.0 (c 0.31, CH<sub>2</sub>Cl<sub>2</sub>).

### Synthesis of **35**<sup>52</sup>

A solution of **19** (66 mg, 0.37 mmol), triphenylphosphine (134 mg, 0.51 mmol), and benzoic acid (64 mg, 0.51 mmol) in dry THF (4 mL) was added to diethyl azodicarboxylate (225 μL, 40% in toluene, 0.51 mmol) in THF (1 mL). After stirring at room temperature for 2 h, the solvent was evaporated and the residue was taken up in diethyl ether (10 mL). After removal of the precipitate by filtration, the filtrate was washed with saturated NaHCO<sub>3</sub> solution, dried, and concentrated. The residue was purified through silica gel column chromatography with petrol ether/EtOAc (30 : 1–15 : 1) to afford **35** as a colourless solid (97 mg, 0.34 mmol, 93%).

<sup>1</sup>H NMR (700 MHz, C<sub>6</sub>D<sub>6</sub>): δ<sub>H</sub> 8.24–8.21 (m, 2H), 7.10–7.07 (m, 1H), 7.04–7.01 (m, 2H), 5.66–5.64 (m, 2H), 2.03–1.97 (m, 1H), 1.87–1.83 (m, 1H), 1.81–1.75 (m, 1H), 1.70 (ddd, <sup>2</sup>J<sub>H,H</sub> = 12.8 Hz, <sup>3</sup>J<sub>H,H</sub> = 12.8 Hz, <sup>3</sup>J<sub>H,H</sub> = 3.3 Hz, 1H), 1.56–1.52 (m, 1H), 1.47 (dddd, <sup>2</sup>J<sub>H,H</sub> = 13.3 Hz, <sup>3</sup>J<sub>H,H</sub> = 13.3 Hz, <sup>3</sup>J<sub>H,H</sub> = 3.7 Hz, <sup>3</sup>J<sub>H,H</sub> = 3.7 Hz, 1H), 1.41–1.35 (m, 2H), 1.32 (ddd, <sup>2</sup>J<sub>H,H</sub> = 13.0 Hz, <sup>3</sup>J<sub>H,H</sub> = 5.6 Hz, <sup>3</sup>J<sub>H,H</sub> = 3.0 Hz, 1H), 1.27 (ddd, <sup>2</sup>J<sub>H,H</sub> = 12.7 Hz, <sup>3</sup>J<sub>H,H</sub> = 12.7 Hz, <sup>3</sup>J<sub>H,H</sub> = 3.9 Hz, 1H), 0.90–0.86 (m, 1H), 0.89 (d, <sup>3</sup>J<sub>H,H</sub> = 6.6 Hz, 3H), 0.87 (s, 3H) ppm (Fig. S57†). <sup>13</sup>C NMR (176 MHz, C<sub>6</sub>D<sub>6</sub>): δ<sub>C</sub> 166.2 (C), 153.9 (C), 132.6 (CH), 131.8 (C), 130.0 (2 × CH), 128.5 (2 × CH), 115.9 (CH), 69.2 (CH), 42.1 (CH<sub>2</sub>), 37.1 (CH<sub>2</sub>), 36.0 (CH<sub>2</sub>), 35.6 (C), 33.3 (CH), 25.3 (CH<sub>2</sub>), 23.9 (CH<sub>3</sub>), 22.3 (CH<sub>2</sub>), 18.8 (CH<sub>3</sub>) ppm (Fig. S58†). HRMS (EI): [M]<sup>+</sup> calculated for C<sub>19</sub>H<sub>24</sub>O<sub>2</sub><sup>+</sup> *m/z* 284.1774; found *m/z* 284.1774. Optical rotation: [α]<sub>D</sub><sup>25</sup> = +73.3 (c 0.03, CH<sub>2</sub>Cl<sub>2</sub>).

### Synthesis of **36**<sup>52</sup>

The ester **35** (97 mg, 0.34 mmol) was saponified through a refluxing solution of KOH (96 mg, 1.71 mmol) in aqueous methanol (1 : 10, 3 mL) for 2 h. The cooled solution was concentrated, diluted with ether, and filtered. The filtrate was washed with sat. aq. NaHCO<sub>3</sub> solution. Drying with MgSO<sub>4</sub>, evaporation of the solvents and column chromatography on

silica gel with petrol ether/ethyl acetate (10 : 1) furnished **36** as a colourless solid (56 mg, 0.31 mmol, 91%).

$^1\text{H NMR}$  (500 MHz,  $\text{C}_6\text{D}_6$ ):  $\delta_{\text{H}}$  5.38 (dd,  $^3J_{\text{H,H}} = 2.0$  Hz,  $^4J_{\text{H,H}} = 1.9$  Hz, 1H), 4.10 (m, 1H), 2.02 (dqddd,  $^3J_{\text{H,H}} = 12.7$  Hz,  $^3J_{\text{H,H}} = 6.4$  Hz,  $^3J_{\text{H,H}} = 4.0$  Hz,  $^4J_{\text{H,H}} = 1.9$  Hz,  $^5J_{\text{H,H}} = 1.9$  Hz, 1H), 1.66–1.55 (m, 4H), 1.52 (dddd,  $^2J_{\text{H,H}} = 13.5$  Hz,  $^3J_{\text{H,H}} = 13.5$  Hz,  $^3J_{\text{H,H}} = 13.5$  Hz,  $^3J_{\text{H,H}} = 3.8$  Hz,  $^3J_{\text{H,H}} = 3.8$  Hz, 1H), 1.44–1.39 (m, 1H), 1.37–1.33 (m, 1H), 1.27 (ddd,  $^2J_{\text{H,H}} = 13.0$  Hz,  $^3J_{\text{H,H}} = 6.2$  Hz,  $^3J_{\text{H,H}} = 2.9$  Hz, 1H), 1.21 (ddd,  $^2J_{\text{H,H}} = 13.5$  Hz,  $^3J_{\text{H,H}} = 13.5$  Hz,  $^3J_{\text{H,H}} = 4.4$  Hz, 1H), 0.93 (d,  $^3J_{\text{H,H}} = 6.5$  Hz, 3H), 0.90 (s, 3H), 0.90–0.86 (m, 1H) ppm (Fig. S59<sup>†</sup>).  $^{13}\text{C NMR}$  (126 MHz,  $\text{C}_6\text{D}_6$ ):  $\delta_{\text{C}}$  150.7 (C), 120.3 (CH), 65.4 (CH), 41.8 (CH<sub>2</sub>), 37.4 (CH<sub>2</sub>), 35.7 (CH<sub>2</sub>), 35.6 (C), 33.0 (CH), 28.5 (CH<sub>2</sub>), 24.3 (CH<sub>3</sub>), 22.5 (CH<sub>2</sub>), 18.9 (CH<sub>3</sub>) ppm (Fig. S60<sup>†</sup>). Optical rotation:  $[\alpha]_{\text{D}}^{25} = -100.0$  ( $c$  0.08,  $\text{CH}_2\text{Cl}_2$ ).

### Synthesis of **37a** and **37b**

Using the same procedure as for the epoxidation of **19**, compound **35** (36 mg, 0.20 mmol) was converted into **37a** (26 mg, 0.67 mmol, 67%) and **37b** (8 mg, 0.67 mmol, 20%) that were obtained as colourless needles.

**Compound 37a.**  $^1\text{H NMR}$  (500 MHz,  $\text{C}_6\text{D}_6$ ):  $\delta_{\text{H}}$  3.79 (br dd,  $^3J_{\text{H,H}} = 5.1$  Hz,  $^3J_{\text{H,H}} = 5.1$  Hz, 1H), 3.13 (d,  $^3J_{\text{H,H}} = 4.3$  Hz, 1H), 2.26 (d,  $^3J_{\text{H,H}} = 10.5$  Hz, OH), 1.75–1.67 (m, 1H), 1.58 (ddd,  $^2J_{\text{H,H}} = 13.7$  Hz,  $^3J_{\text{H,H}} = 13.7$  Hz,  $^3J_{\text{H,H}} = 3.7$  Hz, 1H), 1.50–1.32 (m, 7H), 1.20–1.15 (m, 1H), 0.73 (s, 3H), 0.65 (ddd,  $^2J_{\text{H,H}} = 13.1$  Hz,  $^3J_{\text{H,H}} = 3.6$  Hz,  $^3J_{\text{H,H}} = 3.6$  Hz, 1H), 0.57 (d,  $^3J_{\text{H,H}} = 6.6$  Hz, 3H) ppm (Fig. S61<sup>†</sup>).  $^{13}\text{C NMR}$  (126 MHz,  $\text{C}_6\text{D}_6$ ):  $\delta_{\text{C}}$  69.8 (C), 63.6 (CH), 58.7 (CH), 38.0 (CH<sub>2</sub>), 34.0 (C), 33.5 (CH<sub>2</sub>), 31.5 (CH<sub>2</sub>), 30.0 (CH), 27.2 (CH<sub>2</sub>), 22.0 (CH<sub>2</sub>), 20.6 (CH<sub>3</sub>), 14.5 (CH<sub>3</sub>) ppm (Fig. S62<sup>†</sup>). HRMS (APCI):  $[\text{M} + \text{H}]^+$  calculated for  $\text{C}_{12}\text{H}_{21}\text{O}_2^+$   $m/z$  197.1536; found  $m/z$  197.1535. Optical rotation:  $[\alpha]_{\text{D}}^{25} = +48.6$  ( $c$  0.07,  $\text{CH}_2\text{Cl}_2$ ).

**Compound 37b.**  $^1\text{H NMR}$  (500 MHz,  $\text{C}_6\text{D}_6$ ):  $\delta_{\text{H}}$  3.72 (dd,  $^3J_{\text{H,H}} = 9.9$  Hz,  $^3J_{\text{H,H}} = 7.0$  Hz, 1H), 3.00 (br s, 1H), 2.01–1.92 (m, 1H), 1.61 (ddd,  $^2J_{\text{H,H}} = 13.1$  Hz,  $^3J_{\text{H,H}} = 13.1$  Hz,  $^3J_{\text{H,H}} = 4.7$  Hz, 1H), 1.54–1.48 (m, 1H), 1.46–1.22 (m, 4H), 1.32–1.24 (m, 1H), 1.07–1.02 (m, 1H), 1.04 (s, 3H), 0.91 (dddd,  $^2J_{\text{H,H}} = 13.1$  Hz,  $^3J_{\text{H,H}} = 13.1$  Hz,  $^3J_{\text{H,H}} = 13.1$  Hz,  $^3J_{\text{H,H}} = 4.1$  Hz, 1H), 0.82–0.76 (m, 1H), 0.71 (d,  $^3J_{\text{H,H}} = 6.8$  Hz, 3H) ppm (Fig. S63<sup>†</sup>).  $^{13}\text{C NMR}$  (126 MHz,  $\text{C}_6\text{D}_6$ ):  $\delta_{\text{C}}$  67.4 (C), 67.3 (CH), 60.2 (CH), 36.0 (CH<sub>2</sub>), 35.0 (CH<sub>2</sub>), 34.4 (C), 32.4 (CH<sub>2</sub>), 31.8 (CH), 26.6 (CH<sub>2</sub>), 23.1 (CH<sub>3</sub>), 22.0 (CH<sub>2</sub>), 14.6 (CH<sub>3</sub>) ppm (Fig. S64<sup>†</sup>). HRMS (APCI):  $[\text{M} + \text{H}]^+$  calculated for  $\text{C}_{12}\text{H}_{21}\text{O}_2^+$   $m/z$  197.1536; found  $m/z$  197.1535. Optical rotation:  $[\alpha]_{\text{D}}^{25} = +32.0$  ( $c$  0.05,  $\text{CH}_2\text{Cl}_2$ ).

### Synthesis of **10**

To a solution of **37a** (23 mg, 0.12 mmol) in THF (6.5 mL) was added lithium aluminum hydride (42 mg, 1.11 mmol) at 0 °C under Ar. The mixture was stirred at 0 °C for 1 h and then at room temperature overnight. Water (0.1 mL) was added dropwise, followed by 2 N NaOH (0.1 mL) and water (2 mL). The aqueous phase was extracted with Et<sub>2</sub>O (3 × 10 mL). The combined organic layers were washed with brine (10 mL), dried over MgSO<sub>4</sub>, and concentrated *in vacuo* to give the crude

product. Column chromatography on silica gel with petrol ether/Et<sub>2</sub>O (10 : 1) yielded **10** (20 mg, 0.10 mmol, 86%) as a colourless solid.

$^1\text{H NMR}$  (700 MHz,  $\text{CDCl}_3$ ):  $\delta_{\text{H}}$  4.12 (br s, 1H), 1.97 (ddd,  $^2J_{\text{H,H}} = 13.6$  Hz,  $^3J_{\text{H,H}} = 13.6$  Hz,  $^3J_{\text{H,H}} = 3.9$  Hz, 1H), 1.94–1.85 (m, 2H), 1.72–1.62 (m, 4H), 1.53 (dd,  $^2J_{\text{H,H}} = 14.8$  Hz,  $^3J_{\text{H,H}} = 3.2$  Hz, 1H), 1.49–1.45 (m, 1H), 1.43–1.32 (m, 2H), 1.05–1.02 (m, 1H), 1.01 (s, 3H), 0.95 (ddd,  $^2J_{\text{H,H}} = 13.2$  Hz,  $^3J_{\text{H,H}} = 4.5$  Hz,  $^3J_{\text{H,H}} = 2.1$  Hz, 1H), 0.84 (d,  $^3J_{\text{H,H}} = 6.6$  Hz, 3H) ppm (Fig. S9<sup>†</sup>).  $^{13}\text{C NMR}$  (176 MHz,  $\text{CDCl}_3$ ):  $\delta_{\text{C}}$  76.7 (C), 68.3 (CH), 37.7 (C), 35.1 (CH<sub>2</sub>), 34.8 (CH<sub>2</sub>), 34.2 (CH), 30.5 (CH<sub>2</sub>), 29.9 (CH<sub>2</sub>), 29.0 (CH<sub>2</sub>), 20.8 (CH<sub>2</sub>), 20.2 (CH<sub>3</sub>), 15.1 (CH<sub>3</sub>) ppm (Fig. S10<sup>†</sup>). HRMS (APCI):  $[\text{M} + \text{H}]^+$  calculated for  $\text{C}_{12}\text{H}_{23}\text{O}_2^+$   $m/z$  199.1693; found  $m/z$  199.1692. Optical rotation:  $[\alpha]_{\text{D}}^{25} = -7.3$  ( $c$  0.13, MeOH), lit.  $[\alpha]_{\text{D}}^{22} = -6.2$  ( $c$  0.1, MeOH).<sup>35</sup>

## Conclusions

The enantioselective synthesis of geosmin and its enantiomer has been reported before, and here we have enantioselectively synthesised all six other stereoisomers of geosmin. The synthesis of **22** and *ent*-**22** made use of Reviel's key intermediates **18** and *ent*-**18**,<sup>45</sup> and analogously **27** and *ent*-**27** were synthesised from **23** and *ent*-**23**. For the synthesis of **34** and *ent*-**34** the Stoltz variation of the Tsuji–Trost allylation<sup>50</sup> was used, resulting in enantiomerically enriched materials (*ca.* 55% ee) that could be purified through HPLC using a chiral stationary phase. Furthermore, the previously reported compound **10** from a mangrove associated streptomycete<sup>35</sup> was synthesised enantioselectively, establishing its absolute configuration. Isogeosmin (**38**) is observed as a headspace constituent of some myxobacteria<sup>9</sup> and streptomycetes including *S. reticuli* and *S. ambofaciens*. In this study we have demonstrated that this compound is a side product of geosmin synthase and unequivocally clarified its structure including absolute configuration through an enantioselective synthesis. The formation of **38** is explainable by an alternative reprotonation of the neutral intermediate **4** at C1 and capture of water. Several aspects of the chemical ecology of geosmin are known,<sup>60</sup> *e.g.* geosmin is a constituent of the defence secretions of the polydesmid millipede *Niponia nodulosa*,<sup>61</sup> it repels the bacterial predator *Caenorhabditis elegans* and reduces its grazing on streptomycetes,<sup>62</sup> but attracts springtails feeding on bacteria which promotes spore dispersal.<sup>63</sup> In contrast, the biological functions of the side products of geosmin synthase including the newly identified compound **38** are unknown and remain a subject of future studies.

## Data availability

The data supporting this article have been included as part of the ESI.<sup>†</sup> Crystallographic data for **10**, **18**, *ent*-**18** and *ent*-**19** have been deposited at the Cambridge Crystallographic Data

Centre under accession numbers 2356914, 2356915, 2356916 and 2356917† and can be obtained from <https://www.ccdc.cam.ac.uk/datarequest/cif>.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

This work was funded by the DFG (469042295). We thank Andreas Schneider for HPLC purifications.

## Notes and references

- N. N. Gerber and H. A. Lechevalier, *Appl. Microbiol.*, 1965, **13**, 935–938.
- N. N. Gerber, *Tetrahedron Lett.*, 1968, **9**, 2971–2974.
- J. A. Marshall and A. R. Hochstetler, *J. Org. Chem.*, 1966, **31**, 1020–1025.
- J. A. Marshall and A. R. Hochstetler, *J. Org. Chem.*, 1968, **33**, 2593–2595.
- W. A. Ayer and M. G. Paice, *Can. J. Chem.*, 1976, **54**, 910–916.
- R. G. Buttery and J. A. Garibaldi, *J. Agric. Food Chem.*, 1976, **24**, 1246–1247.
- R. S. Safferman, A. A. Rosen, C. I. Mashni and M. E. Morris, *Environ. Sci. Technol.*, 1967, **1**, 25–26.
- C. E. G. Schoeller, H. Guertler, R. Pedersen, S. Molin and K. Wilkins, *J. Agric. Food Chem.*, 2002, **50**, 2615–2621.
- S. Schulz, J. Fuhlendorff and H. Reichenbach, *Tetrahedron*, 2004, **60**, 3863–3872.
- J. S. Dickschat, S. C. Wenzel, H. B. Bode, R. Müller and S. Schulz, *ChemBioChem*, 2004, **5**, 778–787.
- C. Höckelmann, P. G. Becher, S. H. von Reuss and F. Jüttner, *Z. Naturforsch., C: J. Biosci.*, 2009, **64C**, 49–55.
- C. A. Citron, J. Gleitzmann, G. Laurenzano, R. Pukall and J. S. Dickschat, *ChemBioChem*, 2012, **13**, 202–214.
- T. Kikuchi, S. Kadota, H. Suehara, A. Nishi and K. Tsubaki, *Chem. Pharm. Bull.*, 1981, **29**, 1782–1784.
- T. O. Larsen and J. C. Frisvad, *Mycol. Res.*, 1995, **99**, 1153–1166.
- P. C. Schmidberger and P. Schieberle, *J. Agric. Food Chem.*, 2017, **65**, 9287–9296.
- S. J. Hayes, K. P. Hayes and B. S. Robinson, *J. Protozool.*, 1991, **38**, 44–47.
- J. Spoerle, H. Becker, N. S. Allen and M. P. Gupta, *Z. Naturforsch., C: J. Biosci.*, 1991, **46C**, 183–188.
- Y. Saritas, M. M. Sonwa, H. Iznaguen, W. A. König, H. Muhle and R. Mues, *Phytochemistry*, 2001, **57**, 443–457.
- R. A. Flath, R. R. Forrey, J. O. John and B. G. Chan, *J. Agric. Food Chem.*, 1978, **26**, 1290–1293.
- J. Shi, G. Tong, Q. Yang, M. Huang, H. Ye, Y. Liu, J. Wu, J. Zhang, X. Sun and D. Zhao, *J. Agric. Food Chem.*, 2021, **69**, 11361–11371.
- T. E. Acree, C. Y. Lee, R. M. Butts and J. Barnard, *J. Agric. Food Chem.*, 1976, **24**, 430–431.
- P. E. Persson, *Water Res.*, 1980, **14**, 1113–1118.
- R. Bentley and R. Meganathan, *FEBS Lett.*, 1981, **125**, 220–222.
- D. Spiteller, A. Jux, J. Piel and W. Boland, *Phytochemistry*, 2002, **61**, 827–834.
- J. S. Dickschat, H. B. Bode, T. Mahmud, R. Müller and S. Schulz, *J. Org. Chem.*, 2005, **70**, 5174–5182.
- B. Gust, G. L. Challis, K. Fowler, T. Kieser and K. F. Chater, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 1541–1546.
- D. E. Cane and R. M. Watt, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 1547–1551.
- J. Jiang, X. He and D. E. Cane, *J. Am. Chem. Soc.*, 2006, **128**, 8128–8129.
- J. Jiang, X. He and D. E. Cane, *Nat. Chem. Biol.*, 2007, **3**, 711–715.
- X. He and D. E. Cane, *J. Am. Chem. Soc.*, 2004, **126**, 2678–2679.
- H. Xu and J. S. Dickschat, *ChemBioChem*, 2023, **24**, e202300101.
- J. Jiang and D. E. Cane, *J. Am. Chem. Soc.*, 2008, **130**, 428–429.
- T. Nawrath, J. S. Dickschat, R. Müller, J. Jiang, D. E. Cane and S. Schulz, *J. Am. Chem. Soc.*, 2008, **130**, 430–431.
- N. Ding, L. Han, Y. Jiang, G. Li, J. Liu, Y. Mu and X. Huang, *Magn. Reson. Chem.*, 2018, **56**, 352–359.
- L. Ding and C. Hertweck, *J. Nat. Prod.*, 2020, **83**, 2207–2211.
- N. Ding, Y. Jiang, L. Han, X. Chen, J. Ma, X. Qu, Y. Mu, J. Liu, L. Li, C. Jiang and X. Huang, *J. Nat. Prod.*, 2016, **79**, 799–805.
- R. W. Eaton and P. Sandusky, *Biodegradation*, 2010, **21**, 71–79.
- R. Kaiser and C. Nussbaumer, *Helv. Chim. Acta*, 1990, **73**, 133–139.
- C. A. Citron, J. Gleitzmann, G. Laurenzano, R. Pukall and J. S. Dickschat, *ChemBioChem*, 2012, **13**, 202–214.
- A. Rieck, N. Bülow and W. A. König, *Phytochemistry*, 1995, **40**, 847–851.
- M.-W. Sun, X.-M. Zhang, H.-L. Bi, W.-J. Li and C.-H. Lu, *Nat. Prod. Res.*, 2017, **31**, 77–83.
- W. A. Ayer, L. M. Browne and S. Fung, *Can. J. Chem.*, 1976, **54**, 3276–3282.
- Q. Liu, L. Han, B. Qin, Y. Mu, P. Guan, S. Wang and X. Huang, *Org. Chem. Front.*, 2018, **5**, 1719–1723.
- U. Huber, W. Boland, W. A. König and B. Gehrcke, *Helv. Chim. Acta*, 1993, **76**, 1949–1955.
- G. Revial, *Tetrahedron Lett.*, 1989, **30**, 4121–4124.
- A. Saito, A. Tanaka and T. Oritani, *Tetrahedron: Asymmetry*, 1996, **7**, 2923–2928.
- A. Tanaka, T. Tokuyama, A. Saito and T. Oritani, *Biosci. Biotechnol. Biochem.*, 1998, **62**, 2435–2437.

- 48 H. J. Swarts, A. A. Haaksma, B. J. M. Jansen and A. de Groot, *Tetrahedron*, 1992, **48**, 5497–5508.
- 49 A. M. Pohlit and H. M. C. Ferraz, *Magn. Reson. Chem.*, 1994, **32**, 732–738.
- 50 D. C. Behenna and B. M. Stoltz, *J. Am. Chem. Soc.*, 2004, **126**, 15044–15045.
- 51 P. Schwab, M. B. France, J. W. Ziller and R. H. Grubbs, *Angew. Chem., Int. Ed. Engl.*, 1995, **34**, 2039–2041.
- 52 T.-L. Ho and C.-Y. Su, *J. Org. Chem.*, 2000, **65**, 3566–3568.
- 53 S. Schulz and J. S. Dickschat, *Nat. Prod. Rep.*, 2007, **24**, 814–842.
- 54 K. Grob and F. Zürcher, *J. Chromatogr.*, 1976, **117**, 285–294.
- 55 J. S. Dickschat, K. A. K. Pahirulzaman, P. Rabe and T. A. Klapschinski, *ChemBioChem*, 2014, **15**, 810–814.
- 56 D. C. Behenna, J. T. Mohr, N. H. Sherden, S. C. Marinescu, A. M. Harned, K. Tani, M. Seto, S. Ma, Z. Novk, M. R. Krout, R. M. McFadden, J. L. Roizen, J. A. Enquist, D. E. White, S. R. Levine, K. V. Petrova, A. Iwashita, S. C. Virgil and B. M. Stoltz, *Chem. – Eur. J.*, 2011, **17**, 14199–14223.
- 57 N. S. Simpkins, *J. Chem. Soc., Chem. Commun.*, 1986, 88–90.
- 58 S. J. Branca and A. B. Smith, *J. Org. Chem.*, 1977, **42**, 1026–1030.
- 59 L. M. Browne, R. E. Klinck and J. B. Stothers, *Org. Magn. Reson.*, 1979, **12**, 561–568.
- 60 P. Garbeva, M. Avalos, D. Ulanova, G. P. van Wezel and J. S. Dickschat, *Environ. Microbiol.*, 2023, **25**, 1565–1574.
- 61 H. Omura, Y. Kuwahara and T. Tanabe, *J. Chem. Ecol.*, 2002, **28**, 2601–2612.
- 62 L. Zaroubi, I. Ozugergin, K. Mastronardi, A. Imfeld, C. Law, Y. Gelinias, A. Piekny and B. L. Findlay, *Appl. Environ. Microbiol.*, 2022, **88**, e0009322.
- 63 P. G. Becher, V. Verschut, M. J. Bibb, M. J. Bush, B. P. Molnar, E. Barane, M. M. Al-Bassam, G. Chandra, L. Song, G. L. Challis, M. J. Buttner and K. Flärdh, *Nat. Microbiol.*, 2020, **5**, 821–829.

## Publications

- [1] Z. Yin, H. Xu, J. S. Dickschat, *Org. Biomol. Chem.* **2024**, 22, 8714–8719.
- [2] Z. Yin, J. S. Dickschat, *Beilstein J. Org. Chem.* **2024**, 5, 734–740.
- [3] Z. Yin, E. Liebhart, E. Stegmann, H. Brötz-Oesterhelt, J. S. Dickschat, *Org. Chem. Front.* **2022**, 9, 2714–2720.
- [4] Z. Yin, D. Bär, B. Gust, J. S. Dickschat, *Org. Biomol. Chem.* **2022**, 20, 9103–9107.
- [5] Z. Yin, M. Maczka, G. Schnakenburg, S. Schulz, J. S. Dickschat, *Org. Biomol. Chem.* **2024**, 22, 5748–5758.

## **Abstract**

This cumulative doctoral thesis encompasses five research articles focused on investigations of multifunctional enzymes in natural product biosynthesis. The primary focus of this research is on exploring protein-protein and protein-substrate interactions, which play a critical role in polyketide and non-ribosomal peptide biosynthetic pathways. A part of this research is about studies of PPTase-CP interactions, ketosynthase-substrate, and dehydratase-substrate specificities.

An interesting contribution of this thesis is the detailed investigation into the development of fluorescent probes in monitoring PPTase-CP interactions. This provides a simple way to monitor the carrier proteins activation by PPTases. Additionally, this thesis investigated the substrate specificity of ketosynthase by using <sup>13</sup>C-labeled substrates.

Another key enzyme explored in this thesis is dehydratase, which plays an important role in the polyketide and fatty acid biosynthetic pathways. In order to investigate the dehydratase product configurations and substrate scope, isotopically labeled probes and various substrates were synthesized, which enabled detailed stereochemical investigations.

Moreover, this thesis also investigates the bifunctional terpene synthase, geosmin synthase, and presents the total synthesis of all isomers of geosmin, along with the characterization of the new compound, isogeosmin, which is produced by geosmin synthase.

All in all, this thesis provides a comprehensive and detailed investigations of multifunctional enzymes in natural products biosynthesis. It advances the understanding of the enzymatic processes involved in these essential biochemical pathways, offering new insights into the enzyme mechanisms that drive these reactions.