# Understanding and Engineering the Biosynthesis of Tropolone Sesquiterpenoids in Fungi

**Von der Naturwissenschaftlichen Fakultät der**
**Gottfried Wilhelm Leibniz Universität Hannover**

**zur Erlangung des Grades**
**Doktor der Naturwissenschaften (Dr. rer. nat.)**
**genehmigte Dissertation**

**von**
**Carsten Schotte, M.Sc.**

**2021**

Referent: Prof. Dr. Russell J. Cox
Korreferent: Prof. Dr. rer. nat. Jakob Franke
Tag der Promotion: 10.09.2021

# Abstract

Keywords*:* tropolone sesquiterpenoid · biosynthesis · pathway engineering · humulene · fungi

Tropolone sesquiterpenoids (TS) are meroterpenoid natural products that share the conserved structural feature of a polyketide-derived tropolone nucleus connected to a humulene-derived macrocycle *via* a bridging dihydropyran ring.

Here, the biosynthesis of the TS xenovulene A **1** was investigated using a combination of heterologous gene expression in the fungal host *Aspergillus oryzae* NSAR1 and in *E. coli* BL21. Heterologous expression experiments validated a minimal xenovulene A **1** biosynthetic gene set encoding eight dedicated enzymes involved in tropolone formation, humulene formation and DIELS-ALDER chemistry. Reconstitution of key enzymatic steps *in vitro* identified a new type of class I terpene cyclase (AsR6), that catalyzes the stereoselective formation of α-humulene **2** from farnesyl pyrophosphate **3** or either enantiomer of nerolidyl pyrophosphate **4**.

The biosynthesis of the structurally related bistropolones eupenifeldin **5** and noreupenifeldin B **6** in *Phaeosphaeriaceae* sp. CF-150626 was also investigated. Isotopic labelling studies identified an unusual oxidative ring contraction that putatively converts **5** into **6**. Through a combined genetic and chemical approach, a candidate biosynthetic gene cluster for **5**-biosynthesis was identified (the *eup2* BGC). With EupR3 a homologue of AsR6 was characterized that stereoselectively produces 2*Z*-humulene **7**, a geometric isomer of **2**.

In cooperation with the Helmholtz Institute for Infection Research (Braunschweig, Germany) the crystal structure of AsR6 in the unliganded state and in complex with thiolo-*S*-diphosphate **8** and an *in crystallo* cyclized reaction product was obtained. A new pyrophosphate binding site was identified that consists of a binuclear magnesium cluster and a conserved lysine residue. Site-directed mutagenesis validated the motif and identified a key amino acid residue, L/M285, that drives the stereoselective formation of either **2** or **7**.

New-to-nature TS natural products were produced through heterologous expression of different combinations of biosynthetic enzymes from the xenovulene A **1**, eupenifeldin **5** and pycnidione **9** pathways. The rational design of expression experiments resulted in the formation and characterization of seven new derivatives. The obtained non-natural products differ in the nature of the polyketide moiety, the substitution pattern of the humulene macrocycle and the degree of hydroxylation.

# Zusammenfassung

Schlagwörter: Tropolonsesquiterpenoid · Biosynthese · Pathway Engineering · Humulen · Pilz

Tropolonsesquiterpenoide sind Meroterpenoide, die als konserviertes Strukturmerkmal einen von Polyketiden abgeleiteten Tropolonring aufweisen. Dieser ist über einen verbrückenden Dihydropyranring mit einem von Humulen abgeleiteten Makrocyclus verbunden.

In dieser Arbeit wurde die Biosynthese von Xenovulene A **1** durch heterologe Genexpression in *Aspergillus oryzae* NSAR1 und *E. coli* BL21 untersucht. Heterologe Expressionsexperimente identifizierten eine minimale Anzahl an Enzymen, die an der Tropolonbildung, Humulenbildung und einer zentralen DIELS-ALDER Reaktion beteiligt sind. Die Rekonstitution der wichtigsten enzymatischen Schritte *in vitro* identifizierte einen neuen Typ von Klasse-I-Terpencyclase (AsR6), der die stereoselektive Bildung von α-Humulen **2** aus Farnesylpyrophosphat **3** oder einem der Enantiomere von Nerolidylpyrophosphat **4** katalysiert.

Die Biosynthese der strukturell verwandten Bistropolone Eupenifeldin **5** und Noreupenifeldin B **6** in *Phaeosphaeriaceae* sp. CF-150626 wurde ebenfalls untersucht. Isotopenmarkierungsstudien identifizierten eine ungewöhnliche oxidative Ringkontraktion, die mutmaßlich **5** in **6** umwandelt. Durch einen kombinierten genetischen und chemischen Ansatz wurde ein potentieller Biosynthesegencluster für die Biosynthese von **5** identifiziert (der *eup2* BGC). Mit EupR3 wurde ein Homolog von AsR6 charakterisiert, das stereoselektiv 2*Z*-Humulen **7**, ein geometrisches Isomer von **2**, produziert.

In Zusammenarbeit mit dem Helmholtz-Institut für Infektionsforschung (Braunschweig, Deutschland) wurde die Kristallstruktur von AsR6 im ligandfreien Zustand und im Komplex mit Thiolo-*S*-diphosphat **8** und einem im Kristall zyklisierten Reaktionsprodukt erhalten. Eine neue Pyrophosphatbindungsstelle wurde identifiziert, die aus einem zweikernigen Magnesiumcluster und einem konservierten Lysin besteht. Die ortsgerichtete Mutagenese validierte das Motiv und identifizierte eine wichtige Aminosäure, L/M285, die die stereoselektive Bildung von **2** oder **7** maßgeblich bestimmt.

Neue Tropolonsesquiterpenoide wurden durch die heterologe Expression verschiedener Kombinationen von Biosyntheseenzymen hergestellt. Die rationale Gestaltung von Expressionsexperimenten führte zur Bildung und Charakterisierung von sieben neuen Derivaten. Die erhaltenen nicht natürlichen Produkte unterscheiden sich in der Art der Polyketideinheit, dem Substitutionsmuster des Humulen-Makrocyclus und dem Hydroxylierungsgrad.

# Acknowledgement

I would first like to thank Prof. Russell J. Cox for the excellent supervision in these last four years and for his guidance through this fascinating project. Thank you for this opportunity, the room to explore new ideas, the constant support and the many inspiring discussions we had in all those years. I thoroughly enjoyed working under your supervision and I will always be grateful for the great time I had here in Hannover!

Especially, I also like to kindly thank Prof. Jakob Franke and Prof. Alison Narayan to take the time and effort to co-referee my PhD thesis. I am also very grateful to PD Dr. Carsten Zeilinger for being the chair of the examination board and examiner.

Countless thanks go to Raissa Schor who worked flawlessly on this project before me and who introduced me into the world of fungal natural products. Thank you for the great support in the first months of my PhD! I am also very grateful for the support and friendship I found in Lukas Kahlert, Verena Hantke, Karen Lebe, Oliver Piech and Henrike Heinemann. Thank you so much, also for the great times we enjoyed outside the lab!

Science is a team effort and thus I would like to thank the cooperation partners that engaged on this project with me. Primarily I am very grateful for working with Prof. Wulf Blankenfeldt and Dr. Peer Lukat from the Helmholtz Centre for Infection Research in Braunschweig. Thanks also to Dr. Daniel Wibberg at CeBiTEC Bielefeld, Prof. Olga Genilloud and Víctor González Menéndez for the gift of fungal strains used in this study and to Lei Li, who worked on part of this project with me. Vanessa Harms is thanked for many helpful discussions about the GCMS and terpene cyclases. I also appreciate the contributions of all the great students I had the pleasure to supervise, in particular Adrian Deuschmann, Dave Biedermann and Gia Nam Nguyen.

A special thanks goes out to the fantastic team that runs the service facilities at the OCI (in particular Dagmar and Monica) and the technical staff in the BMWZ: Thank you Katja, Tjorven, Annika, Serge, Birgit and Doreen for always going out of your way to help!

I would also like to thank all current and former members of the Cox and Franke group for creating such a great atmosphere to work in! In particular these were Chongqing, Katja, Dongsong, Mary, Annika, Jin, Eric, Sen, Slawik, Hao, Yingwen, Ling, Liz, Steffen, Francesco, Maurice and our beloved Magnifica S coffee maker.

Finally, I would like to thank my family for their constant support throughout my studies.

# Abbreviations and Units

| | |
|---|---|
| aa | amino acid |
| ACP | acyl carrier protein |
| ADP | adenosine diphosphate |
| amu | atomic mass unit |
| *A. nidulans* | *Aspergillus nidulans* |
| *A. oryaze* | *Aspergillus oryzae* |
| *A. strictum* | *Acremonium strictum* |
| ATP | adenosine triphosphate |
| Asp | aspartic acid |
| BbS | bisabolene synthase |
| BLAST | basic local alignment search tool |
| bp | basepair |
| BGC | biosynthetic gene cluster |
| Carb | carbenicillin |
| *C*-MeT | *C*-methyltransferase domain |
| CoA | coenzyme A |
| COSY | correlation spectroscopy |
| d | days |
| DAD | diode array detector |
| DDP | decalin-containing diterpenoid pyrone |
| DFT | density-functional theory |
| DH | dehydratase |
| DHN | dihydronaphthalene |
| DMAPP | dimethylallyl pyrophosphate |
| DMOA | dimethylorsellinic acid |
| DNA | deoxyribonucleic acid |
| E, Glu | glutamic acid |
| *E. coli* | *Escherichia coli* |
| EDTA | ethylenediaminetetraacetic acid |
| EI | electron impact |
| EIC | extracted ion chromatogram |
| ELSD | evaporative light scattering detector |
| EPO | erythropoietin |
| ER | enoylreductase |
| FAD | flavin adenine dinucleotide |
| FMO | flavin dependent monooxygenase |
| FPP | farnesyl diphosphate/ pyrophosphate |
| GABA | gamma-aminobutyric acid |
| gDNA | genomic DNA |
| GOI | gene of interest |
| GRAS | generally recognized as safe |
| h | hours |
| hDA | hetero DIELS-ALDERase |
| HMBC | heteronuclear multiple bond correlation |
| HMG | 3-hydroxy-3-methyl-glutaryl |
| HPLC | high performance liquid chromatography |
| HRMS | high resolution mass spectrometry |
| HR PKS | highly reducing polyketide synthase |
| HSQC | heteronuclear single quantum correlation |
| Hz | Hertz |
| $IC_{50}$ | half maximal inhibitory concentration |
| IMAC | immobilized metal affinity chromatography |
| IPP | isopentenyl pyrophosphate |
| ITS | internal transcribed spacer |
| Kan | kanamycin |
| kDa | kilo Dalton |
| KO | knockout |
| KR | ketoreductase |
| KS | ketosynthase |
| L | Litre |
| LCMS | liquid chromatography mass spectrometry |
| Leu | leucine |
| µL | microliter |
| M | molar, methionine |
| MAT | malonyl-CoA acyl transferase |
| MEP | 2-*C*-D-erythritol 4-phosphate |
| mg | milligram |
| *M. grisea* | *Magnaporthe grisea* |
| min | minute |
| mL | millilitre |
| MS | mass spectrometry |
| MVA | mevalonic acid |
| *m/z* | mass to charge ratio |
| N, Asn | asparagine |
| NADPH | nicotinamide adenine dinucleotide phosphate |
| NHI | non-heme iron dioxygenase |
| nm | nanometer |
| nM | nanomolar |
| NMR | nuclear magnetic resonance |
| nOe | nuclear Oberhauser effect |
| NPP | nerolidyl pyrophosphate |
| NRPS | non-ribosomal peptide synthetase |
| NR PKS | non reducing polyketide synthase |
| ORF | open reading frame |
| P450 | cytochrome P450 |
| PAGE | polyacrylamide gel electrophoresis |
| PCR | polymerase chain reaction |
| PE | peak enhancement |
| PEG | polyethylene glycol |
| PKS | polyketide synthase |
| PPant | phosphopantetheinyl |
| PPtase | phosphopantetheinyl transferase |
| $PP_i$ | inorganic pyrophosphate |
| PR PKS | partially reducing polyketide synthase |
| ppm | parts per million |
| R | reductive release domain |
| RNA | ribonucleic acid |
| rRNA | ribosomal RNA |
| rpm | revolutions per minute |
| S | serine |
| *S. cerevisiae* | *Saccharomyces cerevisiae* |
| SAD | single-wavelength anomalous dispersion |
| SAM | *S*-adenosyl methionine |
| SAT | starter unit:acyl transferase |
| SDR | short-chain dehydrogenase |
| SDS | sodium dodecyl phosphate |
| TAE | tris-acetate-EDTA |
| TC | terpene cyclase |
| TE | thiolesterase |
| TEMED | tetramethylethylenediamine |
| TIC | total ion current |
| $t_R$ | retention time |
| TS | tropolone sesquiterpenoid |
| *T. stipitatus* | *Talaromyces stipitatus* |
| UV | ultraviolet |
| WT | wild-type |

# Table of Contents

# 1. Introduction

For thousands of years mankind has turned to mother nature as a source for medication, to treat injuries and to reduce pain. Remarkable examples of naturally produced medications include the broad spectrum antibiotic penicillin G **10**, lovastatin **11**, a competitive inhibitor of HMG-CoA reductase used to treat hypercholesterolemia, and artemisinin **12**, a crucial antimalarial agent.[1,2] Powerful analgesics such as morphine **13** have revolutionized modern medicine and are indispensable in many clinical procedures (Scheme 1.1 A). The high value of these remedies has inspired generations of scientists to isolate, purify and characterize the active ingredients in nature that are responsible for the medical properties of the host. These active ingredients are referred to as natural products – in the broadest definition these are all the molecules produced by any living organism. Natural products are categorized into primary and secondary metabolites. Primary metabolites, such as amino acids, fatty acids or sugars, have specific physiological functions and are necessary for the host survival. Secondary metabolites do not fulfil these specific functions. DAVIES and RYAN define them as natural products that are *not* ubiquitously found in different species, that are *not* formed during all stages of host development and that are *not* required for host survival.[3]

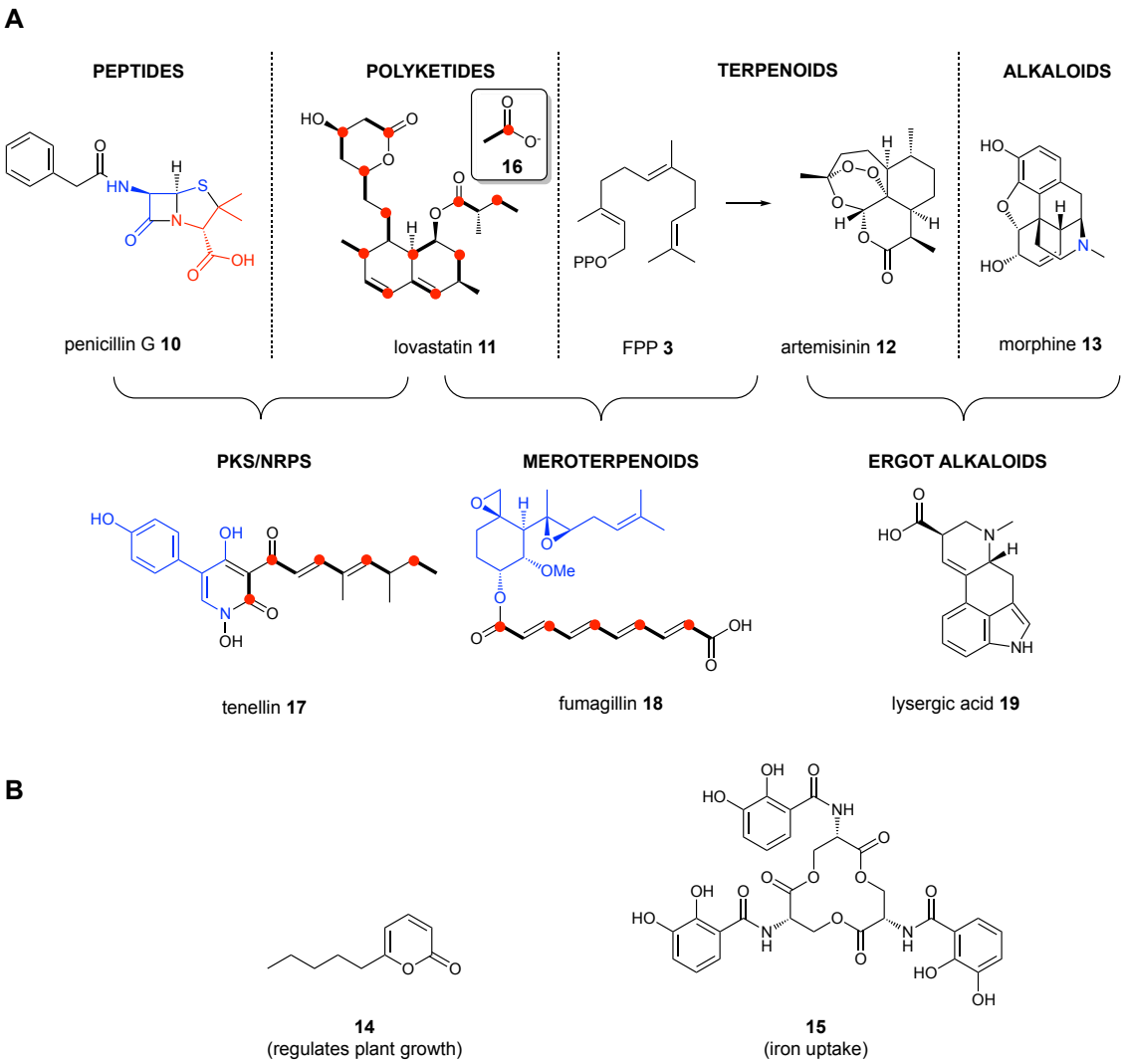
Many secondary metabolites are of low molecular weight and display potent bioactivities that humans exploit in the industrial, agricultural and pharmaceutical sector.[4] The ecological purpose of these metabolites remains, in many cases, elusive. However, secondary metabolites are reasoned to provide their hosts with an advantage in competitive communities through antibiotic or virulence activity, modulation of gene transcription or intercellular signalling.[5]

## 1.1. Fungal Secondary Metabolites

Fungi are a treasure trove for the isolation of bioactive secondary metabolites. The oldest fungal fossils are roughly one billion years old.[6] Over time, evolution has honed highly specialized fungal metabolites that mediate the interactions with other organisms: One such example is the pyran-2-one **14**, that acts as a plant-growth regulator.[7,8] Another example, enterobactin **15**, participates in fungal iron uptake (Scheme 1.1 B)[8,9]

The diversity and capacity of fungi to produce secondary metabolites is enormous. Conservative estimations from 2017 suggested the existence of at least 2.2 million fungal species - of which less than 10% are characterized![10] However, fungi represent both an opportunity and a danger to humankind alike. Toxic compounds like aflatoxins can be contaminants of human food supplies

that can cause hepatotoxicity, teratogenicity and immunotoxicity.[11] Fungal infections still represent a major threat to plant and animal health and relatively few options of treatment exist.[12]




**Scheme 1.1** Representative examples of secondary metabolites: **A,** classification of secondary metabolites according to their biosynthetic origin; bold bonds = intact acetate **16** units; red sphere = C-1 carbon of acetate; **B,** selected fungal metabolites with known ecological purpose.

In fungi (and other microorganisms) the metabolic pathways leading to formation of these secondary metabolites are encoded in biosynthetic gene clusters (BGC). BGC are physically clustered groups of genes, that together are sufficient for the production of the corresponding metabolite.[13,14] Typically, the genes cluster around a core gene that encodes the generation of the carbon backbone (*e.g.* a polyketide synthase). Based on the carbon backbone and mode of biosynthesis four major classes of natural products are distinguished: **1),** peptides (*e.g.* penicillin G **10**) are derived from the condensation of amino acid building blocks; **2),** polyketides (*e.g.* lovastatin **11**) are derived from the iterative condensation of acetate **16** units; **3),** terpenoids (*e.g.* artemisinin **12**) are derived from universal isoprenoid precursors (*e.g.* farnesyl diphosphate **3**

[FPP]); and **4,** alkaloids (*e.g.* morphine **13**), which are nitrogenous organic compounds containing a basic nitrogen atom usually derived from amino acids (Scheme 1.1A).

Additionally, encoded genes within a BGC are typically involved in tailoring/functionalization of the carbon backbone; regulation of BGC transcription; transport of metabolites; and self-resistance.[13]

Besides the four major classes of natural products hybrid compounds are biosynthetically derived from mixed biosynthetic pathways. Tenellin **17** for example, a 2-pyridone isolated from the insect pathogenic fungus *Beauveria bassiana*, is derived from mixed non-ribosomal peptide/polyketide origin. A single hybrid polyketide synthase/non-ribosomal peptide synthetase (PKS/NRPS) furnishes the carbon backbone from tyrosine, acetate, malonate and methionine.[15,16] Fumagillin **18**, isolated from *Aspergillus fumigatus* and best known for its antiangiogenic activity, is derived in similar fashion from mixed polyketide/terpene biosynthesis.[17,18] The ergot alkaloid lysergic acid **19** is of mixed (di)terpenoid/alkaloid origin (Scheme 1.1A).[19]

Natural products that are partially derived from terpenes are referred to as meroterpenoids and in this thesis a focus will be on meroterpenoids of mixed terpene/polyketide biosynthetic origin.[20] The principles of polyketide and terpene biosynthesis will be discussed below.

## 1.2. Polyketide Biosynthesis

Polyketides are a major family of natural products that are of particular importance to human medicine. Members of this family are clinically used as antibiotics (*e.g.* erythromycin **20**), antifungals (*e.g.* amphotericin B), insecticides (*e.g.* spinosyn A) or immunosuppressants (*e.g.* rapamycin).[21] The broad spectrum of biological activities is mirrored in the structural diversity of polyketides. The class of polyketides includes relatively small metabolites such as the fungicidal metabolite strobilurin A **21**, but also macrolides such as SCH-642305 **22** or more complicated structures such as squalestatin S1 **23**, a picomolar inhibitor of squalene synthase.[22–24] Surprisingly, the structural diversity is not mirrored in the biosynthetic origin of polyketides, as these are all derived from short-chain carboxylic acids such as acetate **16** and malonate **24a** (Figure 1.1).[25]

Already in the 1950s ARTHUR J. BIRCH proposed that polyketide biosynthesis is related to the biosynthesis of fatty acids and that the head-to-tail condensation of acetate **16** units gives rise to the polyketide carbon backbone:[26] In a revelatory publication in 1955 BIRCH *et al*. reported on the incorporation of four units of [1-$^{14}$C]-acetate into 6-methylsalicylic acid **25**, produced by the fungus *Penicillium patulum* (Figure 1.1).[27]

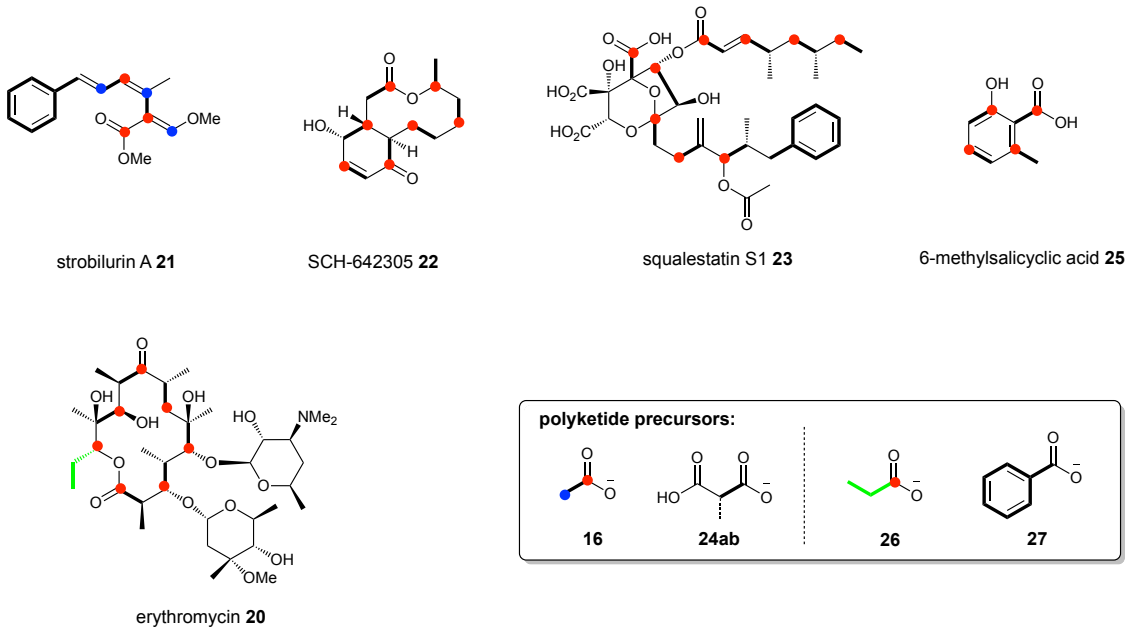



**Figure 1.1:** Representative examples of polyketides and their common biosynthetic origin from acetate **16** and (methyl)malonate **24a(b)**; the use of different starter units (*e.g.*, proprionate **26**, benzoate **27**) expands the structural diversity of polyketides.

More recent work, together with the exponential growth of genomic data, revealed that polyketides are biosynthesized by very large multifunctional enzymes called polyketide synthases.[21] The highly programmed class of polyketide synthases can be classified into three types of enzymes, based on their three-dimensional architectures. Type I PKS are giant multifunctional enzymes in which individual catalytic domains are covalently linked. Type I systems can be further divided into modular and iterative PKS.[25,28] Type I modular PKS are structurally organized into distinct modules and each module contains a distinct set of (covalently linked) domains, that catalyze the incorporation and modification of one extender unit.[28] For example, the bacterial DEBS1-3 enzymes, responsible for the biosynthesis of erythromycin **20**, comprise seven modules that catalyze the successive incorporation of seven carboxylic acids (propionate **26** and methylmalonate **24b**; Figure 1.1 A).[29] Fungal PKS almost exclusively belong to the class of monomodular *iterative* type I PKS, and they use a single set of catalytic domains (one module) repeatedly.[25] The three types of PKS are complemented by type II PKS, that consist of multi-enzyme complexes of non-covalently linked monofunctional enzymes, and type III PKS, that work in an ACP-independent fashion (*vide infra*).[28]

### 1.2.1. Fungal Iterative PKS

Monomodular fungal iterative type I PKS catalyze the formation of complex polyketides using a single set of domains in an iterative fashion.[25] In general, polyketide biosynthesis can be divided into three stages: **1)**, loading of the PKS with appropriate starter/extender units; **2)**, chain extension; and **3)**, optional β-keto processing.[25] The minimal domains required to fulfil loading

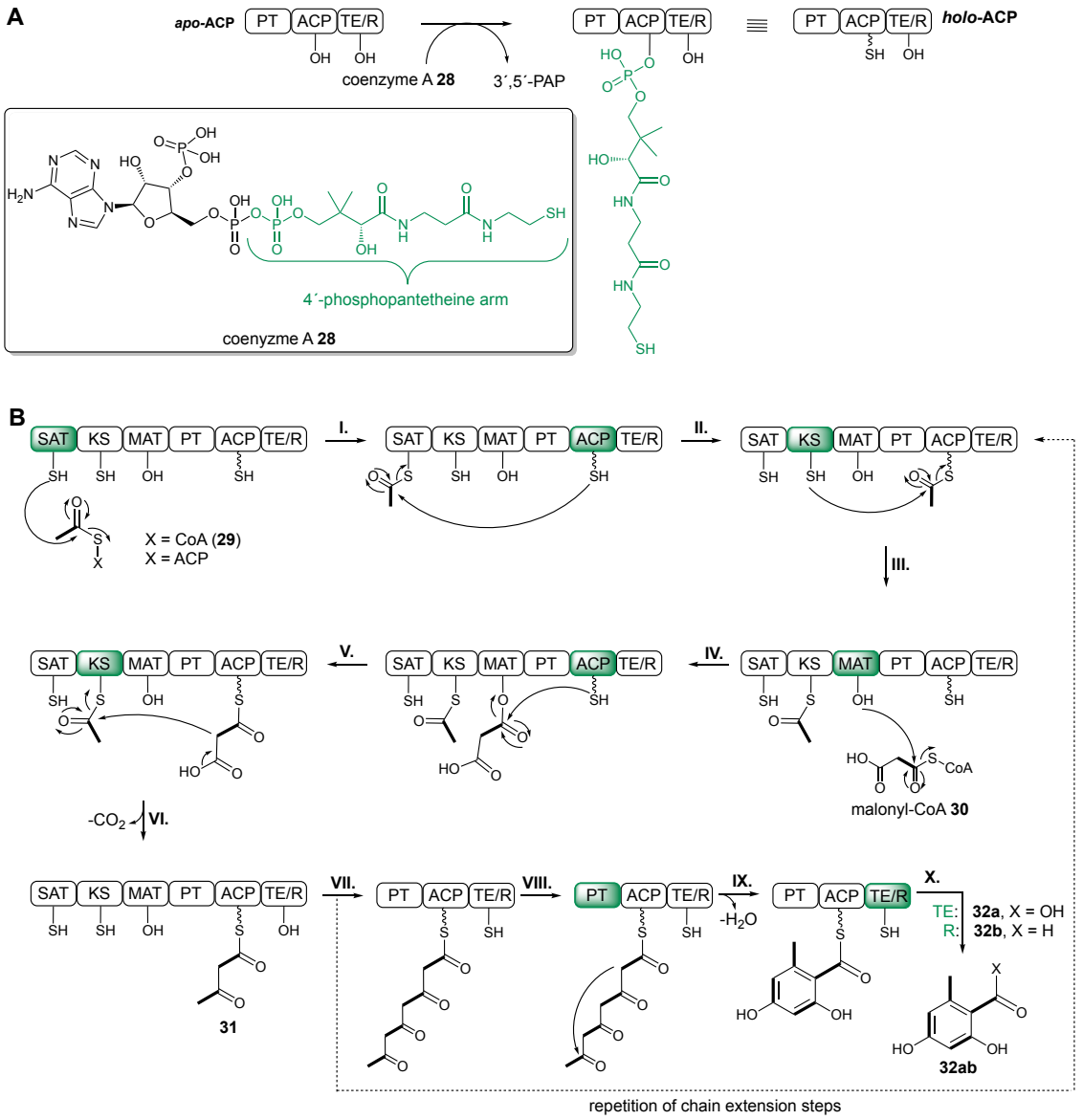
and chain extension are the β-ketoacyl synthase (KS), malonyl-CoA transacylase (MAT) and acyl-carrier protein (ACP).[30]

ACP domains are multifunctional domains responsible for **a),** accepting starter/extender units from the MAT-domain; **b),** collaborating with the KS domain to mediate chain extension; and **c),** transport of the nascent polyketide chain between β-processing domains.[31] Prior to catalysis the *apo*-ACP protein requires post-translational attachment of a 4′-phosphopantetheine arm (PPant arm).[31] This modification is catalyzed by a phosphopantetheinyl transferase (PPTase), that mediates the transfer of the PPant arm from coenzyme A **28** to a conserved serine residue of the acyl carrier protein (Scheme 1.2 A).[32] This ~20 Å "prosthetic" arm then covalently binds the nascent polyketide chain *via* its terminal sulfhydryl group as a reactive thiolester intermediate.[32]

Depending on the presence or absence of β-keto processing domains (ketoreductase [KR], dehydratase [DH] and enoylreductase [ER]) fungal type I PKS are further classified as non-reducing (NR PKS), highly-reducing (HR PKS) and partially-reducing PKS (PR PKS).[25]

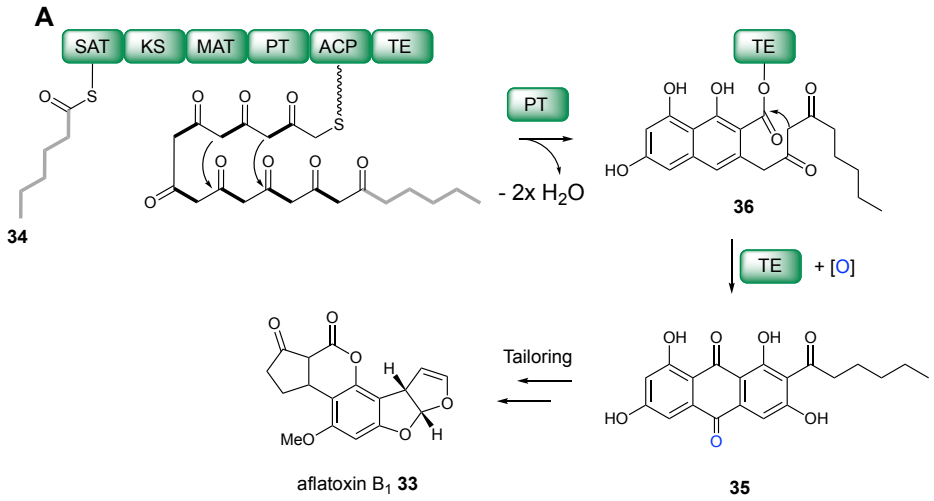
NR PKS possess the minimal set of domains required for chain elongation (MAT, KS, ACP) and additionally a starter unit:ACP transacylase (SAT) domain, product template (PT) domain and a release domain (often a thiolesterase domain [TE]).[30] Starter units (usually acetate **16**) and extender units (mostly malonate **24a**) are activated as coenzyme A (CoA) thiolesters. The SAT-domain catalyzes the initial transfer of a starter acyl unit (often acetyl-CoA **29**)[33] to the ACP domain (Scheme 1.1 B; Step I + II).[30] After loading, the ACP domain primes the KS-domain with the starter unit (Scheme 1.1 B; Step III).[30] Loading and transfer of the first extender unit (typically malonyl-CoA **30**) to the ACP is catalyzed by the MAT domain (Scheme 1.2 B; Step IV+V). The successive chain extension is catalyzed by the KS domain: Decarboxylative Claisen condensation affords the newly formed acyl thiolester **31**, that has undergone net addition of one $C_2$ unit (Scheme 1.2 B; Step VI).[25] Successive cycles of chain extension lead to the mature polyketide chain (Scheme 1.2 B; Step VII).

The PT domain is then involved in catalysis of intramolecular formation of cyclic/fused polyketide products (*e.g.* through regioselective intramolecular aldol reactions; Scheme 1.2 B; Step VIII and IX).[34] At the final stage of biosynthesis the TE domain mediates release of the mature polyketide, *e.g.* by hydrolysis to the acid **32a** or an R domain can catalyze reductive release leading to the aldehyde **32b** (Scheme 1.2 B; Step X).[35]




**Scheme 1.2** Overview over the biosynthesis of fungal polyketides: **A,** post-translational modification of the ACP domain; 3´,5´-PAP = 3´,5´-phosphoadenosine phosphate; **B,** mechanism of polyketide biosynthesis; SAT = starter unit:acyl transferase; KS = ketosynthase; MAT = malonyl-CoA transacylase; PT = product template; ACP = acyl carrier protein; TE = thiolesterase; R = reductive release domain.

A `classic´ example of a NR PKS is PksA, the polyketide synthase catalyzing the first committed step of aflatoxin $B_1$ **33** biosynthesis in *Aspergillus parasiticus*.[36] PksA accepts an unusual starter unit, hexanoyl **34** from a dedicated fatty acid synthase, and seven rounds of iterative chain extension (with malonyl CoA **30**) afford the first isolatable intermediate, norsolorinic acid **35**. During **35**-biosynthesis the PT domain catalyzes two successive aldol cyclizations that afford the bicyclic intermediate **36**. Finally, the TE domain catalyzes a Claisen/Dieckmann cyclization that leads to release of norsolorinic acid **35** (Scheme 1.3).[36]

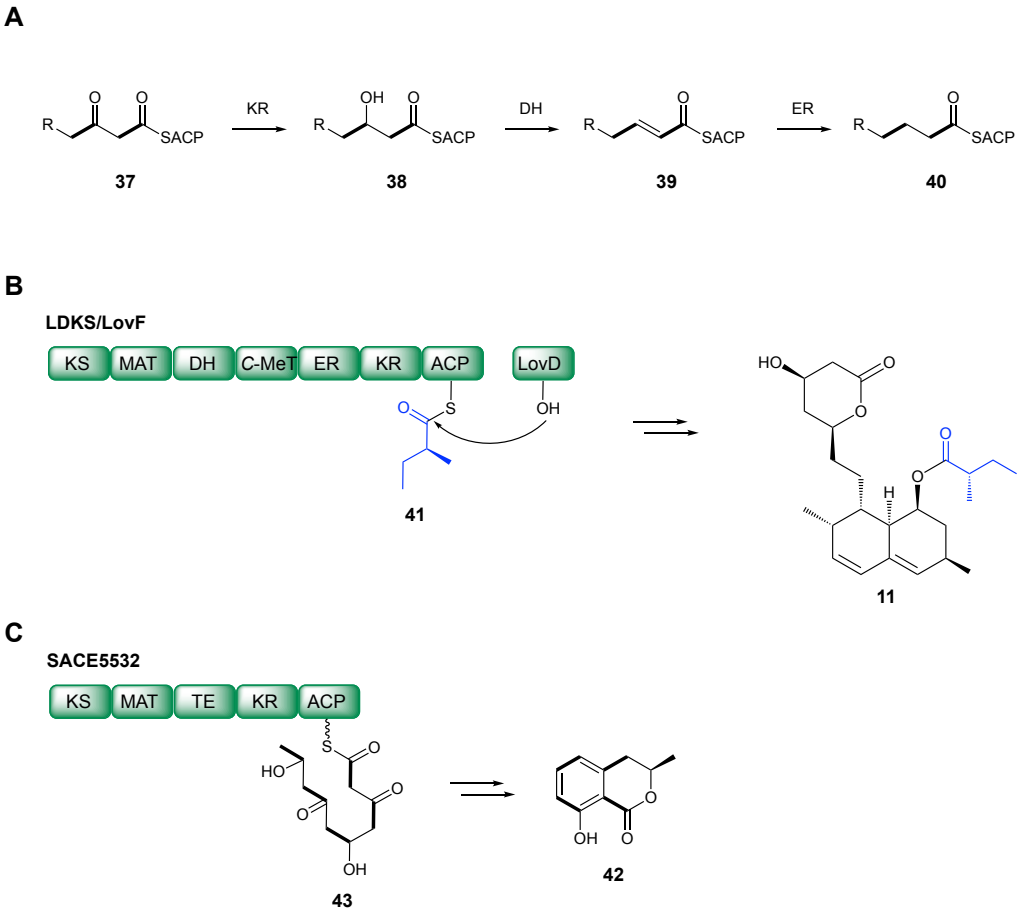



**Scheme 1.3** Biosynthesis of norsolorinic acid **35**: SAT = starter unit:acyl transferase; KS = ketosynthase; MAT = malonyl-CoA transacylase; PT = product template; ACP = acyl carrier protein; TE = thiolesterase.

HR PKS contain all additional domains (KR, DH, ER) required for β-keto functionalization and are highly programmed enzymes, as the degree of β-keto functionalization can vary in each elongation cycle.[30] After chain elongation the KR domain reduces the β-ketothiolester **37** to a secondary alcohol **38**; the DH domain then catalyzes the dehydration to the unsaturated thiolester **39** and the ER domain catalyzes enoyl reduction to afford the fully reduced thiolester **40** (Scheme 1.4 A).[25] One such example of a HR PKS is the diketide synthase LovF from the lovastatin **11** biosynthetic pathway in *Aspergillus terreus*.[37] LovF synthesizes the α-*S*-methyl butyrate **41**, using all of the previously described β-keto processing domains (Scheme 1.4 B). It also harbors an additional methyltransferase domain (*C*-MeT), that installs the α-methyl group before keto-reduction. The *trans* acting acyltransferase LovD is ultimately responsible for off-loading **41** from the PKS.[30,37]

Partially reducing polyketide synthases do not catalyze the full extent of β-keto processing reactions. Only few examples have been reported in the literature.[30] One such example is the PR PKS SACE5532 in *Saccharopolyspora erythraea*, responsible for the formation of (*R*)-mellein **42**. SACE5532 catalyzes formation of the pentaketide **43**, that has undergone selective ketoreduction by the KR domain at the stage of the diketide and tetraketide intermediates (Scheme 1.4 C).[38]

The different types of polyketide synthases, the varying degree of β-ketoreduction, the variety in starter units, optional methylation and different modes of polyketide release are the reasons for the extraordinary diversity observed in this class of secondary metabolites.




**Scheme 1.4** Highly reducing and partially reducing polyketide synthases: **A,** catalytic reactions of the ketoreductase, dehydratase and enoylreductase domain; **B,** domain architecture and catalytic activity of LovF in the biosynthesis of lovastatin **11**; **C,** domain architecture and catalytic activity of SACE5532 in the biosynthesis of (*R*)-mellein **42**; KS = ketosynthase; MAT = malonyl-CoA transacylase; ACP = acyl carrier protein; TE = thiolesterase; KR = ketoreductase; DH = dehydratase; ER = enoylreductase; *C*-MeT = *C*-methyltransferase.

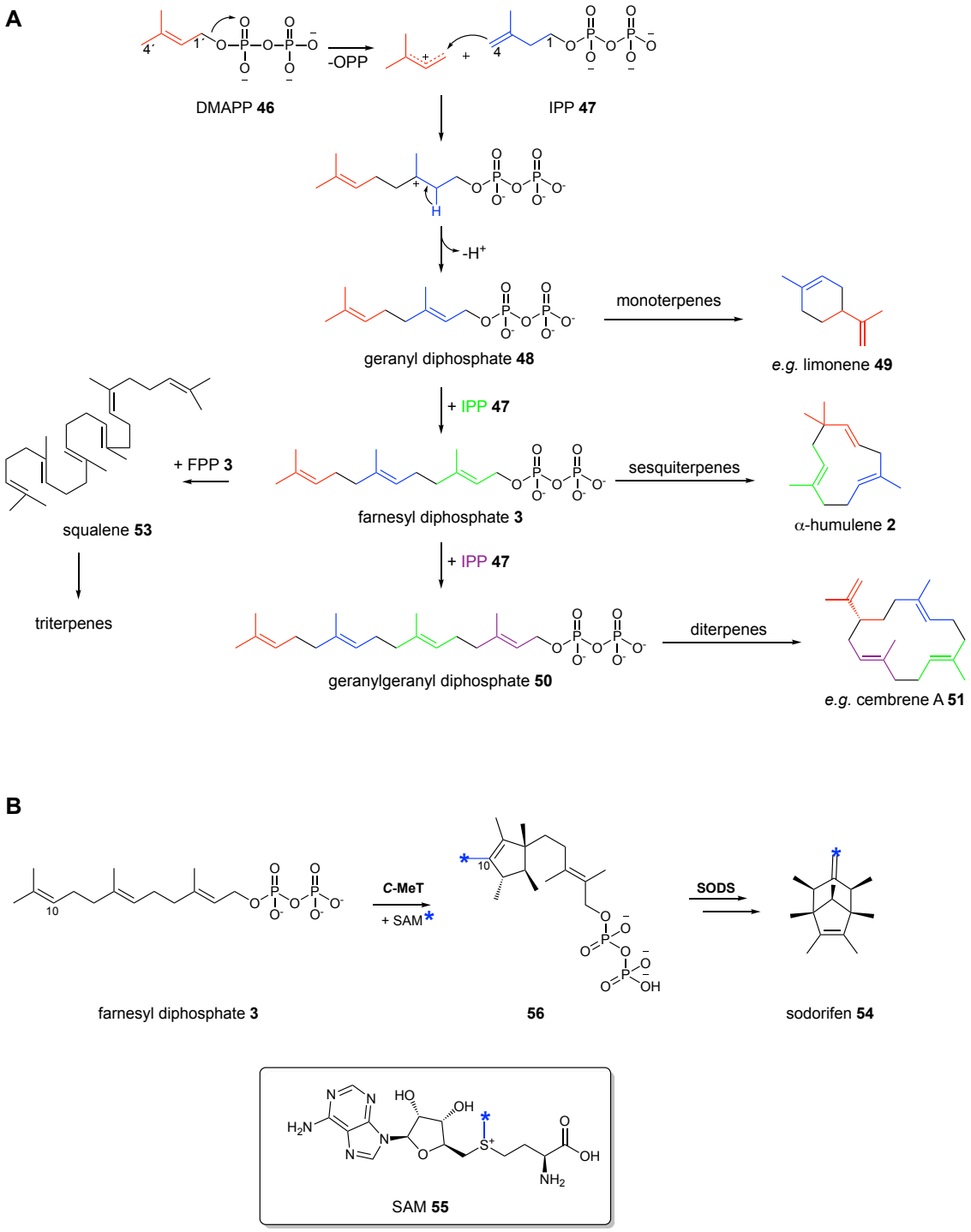
## 1.3. Terpene Biosynthesis


geosmin **44** (±)-linalool **45**

**Figure 1.2:** Chemical structures of geosmin **44** and (±)-linalool **45**.

Terpenes hold a unique position among natural products as they represent the largest group of secondary metabolites. As of 2017 more than 80 000 terpenes have been reported in the literature.[39] Together with the biosynthetically related steroids and carotenoids, the terpenome thus accounts for approximately one third of all secondary metabolites.[39] Terpenes have been isolated from all kingdoms of life and they have been widely used as pharmaceuticals, flavors/fragrances, biofuels or pesticides.[40] A prominent example is geosmin **44**, that causes the odor of moist soil typically encountered after rainfall.[41] (±)-Linalool **45** is a key ingredient of

many plant essential oils and worldwide consumption exceeds 1000 tons per annum (Figure 1.2).[42]




**Scheme 1.5** Terpene biosynthesis: **A,** successive condensation of DMAPP **46** and IPP **47** gives rise to the isoprenoid precursors of mono-, sesqui-, di- and triterpenes; **B,** atypical formation of sodorifen **54** from pre-sodorifen **56**; *C*-MeT = *C*-methyltransferase; SODS = sodorifen synthase; SAM = *S*-adenosylmethionine.

The structural diversity of terpenes is rooted in simple biosynthetic building blocks. All terpenes are derived from the universal precursors dimethylallyl diphosphate **46** (DMAPP) and isopentenyl diphosphate **47** (IPP; Scheme 1.5 A).[39] Precursors **46** and **47** are 5-carbon building blocks. Linear 1´-4 head-to-tail coupling of **46** and **47** affords achiral $C_{5n}$ isoprenoid diphosphates ($n$ = 2, 3, ..).[43] The latter are the direct precursors of all terpenes: Condensation of DMAPP **46** with one molecule of IPP **47** affords geranyl diphosphate **48** ($C_{10}$; precursor of monoterpenes such as limonene **49**); successive condensation with IPP **47** affords farnesyl diphosphate **3** ($C_{15}$; FPP; precursor of sesquiterpenes such as α-humulene **2**), geranylgeranyl diphosphate **50** ($C_{20}$; precursor of diterpenes such as cembrene A **51**) and geranylfarnesyl diphosphate **52** ($C_{25}$; precursor of sesterterpenes; Scheme 1.5 A).[43] Additional diversity is generated by the non-canonical condensation of individual building blocks. For example, the 1´-1 head-to-head coupling of two molecules of FPP **3** affords squalene **53**, the biosynthetic precursor of the triterpenoids (Scheme 1.5 A).[43] There are also a limited number of examples where the isoprenoid (*e.g.* FPP **3**) is further modified prior to the subsequent terpene cyclase reaction. In the biosynthesis of sodorifen **54**, for example, FPP **3** first undergoes C-10 methylation-induced cyclization (catalyzed by a dedicated *S*-adenosyl methionine **55** (SAM)-dependent *C*-methyltransferase) before the corresponding terpene cyclase, SODS, catalyzes **54** formation (Scheme 1.5 B).[44,45]

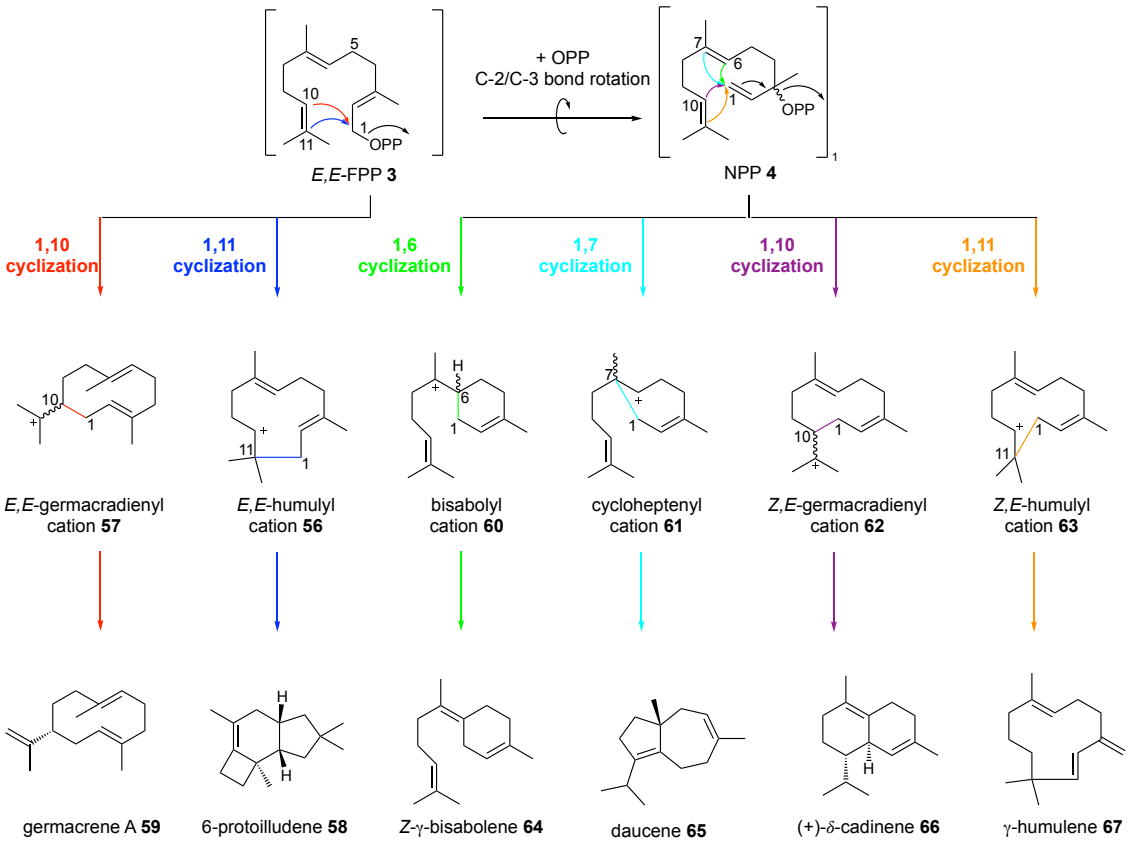
Terpenes such as limonene **49**, α-humulene **2** and cembrene A **51** are formed from geranyl diphosphate **48**, farnesyl diphosphate **3** and geranylgeranyl diphosphate **50** respectively by enzymes that are referred to as terpene cyclases. Terpene cyclases clade into two large families of enzymes, referred to as class I and class II terpene cyclases.[39]

Class I terpene cyclases typically contain a trinuclear metal cluster that is involved in coordination of the diphosphate group. The metal cluster is coordinated by characteristic aspartate-rich motifs that are located on the upper walls of their respective active sites.[39] The highly conserved motifs consist of a **D**Dxx(x)**E/D** and (**N**,**D**)D(L,I,V)x(**S**,**T**)xxx**E** region, the latter typically being referred to as the `NSE/DTE´ motif/triad (boldface indicates metal binding).[39]

Sesquiterpene cyclases, which are the object of research in this study, use FPP **3** as the universal isoprenoid substrate. Initial cleavage of the diphosphate group is thought to result in the formation of a highly reactive allylic carbocation. For FPP **3**, initial carbocation formation can follow seven possible trajectories, depending on the first C-C bond forming reaction.[43] Direct cyclization of FPP **3** *via* a C-1/C-11 bond forming reaction gives rise to the *E,E*-humulyl cation **56** (Scheme 1.6). Alternatively, initial C-1/C-10 bond forming reaction affords the germacradienyl cation **57**. Subsequent cyclisation cascades, rearrangements and hydride migrations give rise to more

structurally complex sesquiterpenes such as 6-protoilludene **58** and germacrene A **59** (Scheme 1.6).[46,47]

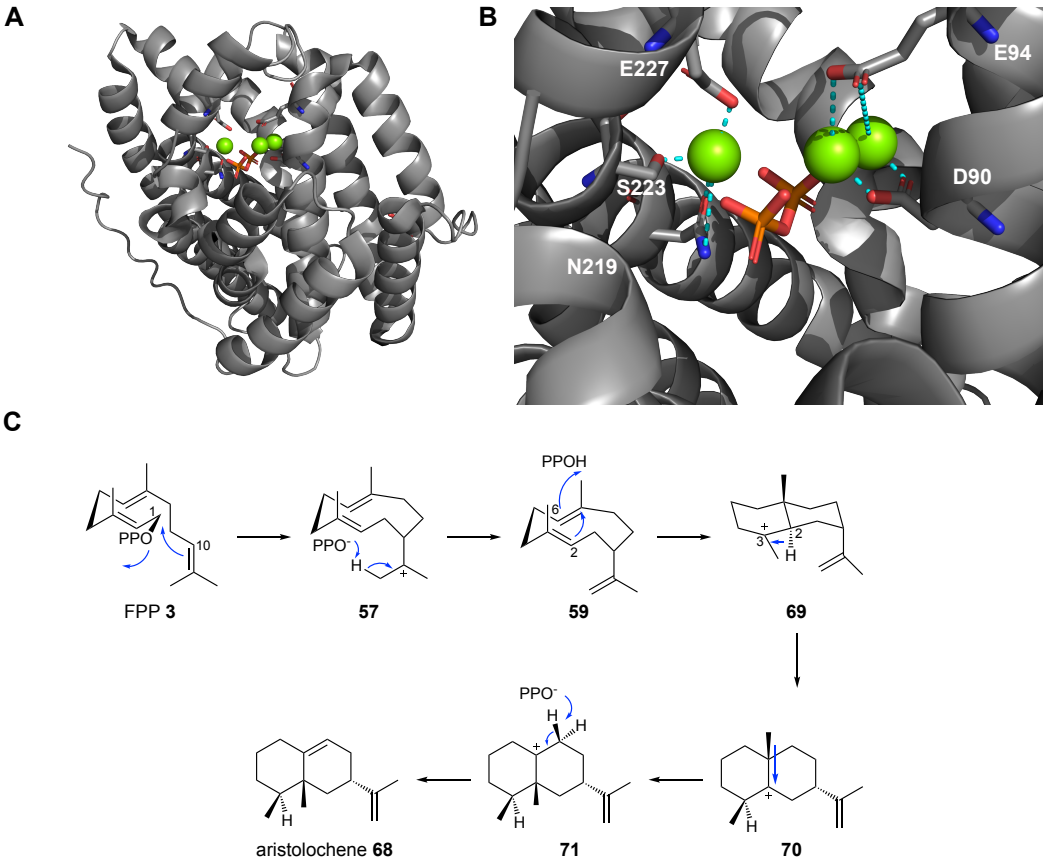
The scope of possible FPP-cyclizations is broadened by the possibility of initial *trans-cis* isomerization of the C-2/C-3- alkene in FPP **3**, resulting in the formation of nerolidyl diphosphate (NPP) **4**.[48] **4** serves as a further template for the formation of the bisabolyl cation **60** (*via* C-1/C-6 bond forming reaction), the cycloheptenyl cation **61** (*via* C-1/C-7 bond forming reaction), the *Z*,*E*-germacradienyl cation **62** (*via* C-1/C-10 bond forming reaction) or the *Z*,*E*-humulyl cation **63** (*via* C-1/C-11 bond forming reaction).[43] Cyclization of **60**-**63** *e.g.* gives rise to *Z*-γ-bisabolene **64**, daucene **65**, (+)-*δ*-cadinene **66** and γ-humulene **67**.[49–52] The latter four trajectories with NPP **4** afford terpenes with a *Z*-configured alkene.




**Scheme 1.6:** Possible cyclizations of FPP **3** during sesquiterpene formation.

The diverse carbocation chemistry involved in terpene formation is illustrated for the formation of aristolochene **68** (Scheme 1.7). Crystal structures of aristolochene synthase have been obtained from both *Penicillium roqueforti* and *Aspergillus terreus*.[53,54] CHRISTIANSON et al. reported the *A. terreus* aristolochene synthase crystal structure at 2.15 Å resolution, in complex with the typical trinuclear $Mg^{2+}$-cluster and inorganic pyrophosphate ($PP_i$; Scheme 1.7 A).[53] Aristolochene synthase crystallized in an α-helical bundle fold that is typical for class I terpene cyclases (Scheme 1.7 A). The conserved **D90**Dxx**E94** motif (boldface indicates ligand binding)

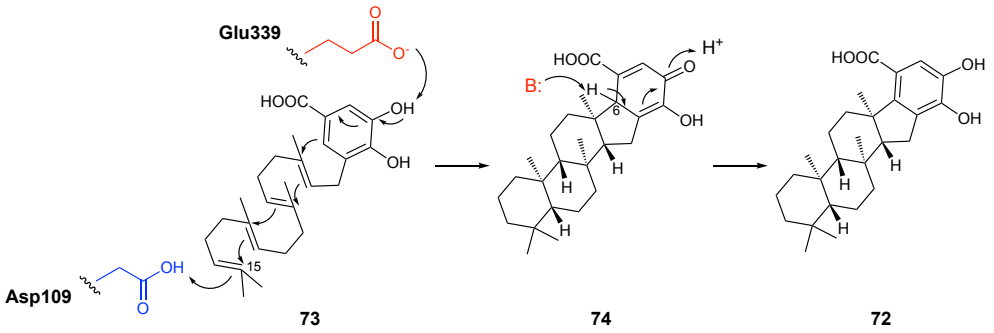
involved in FPP **3** coordination was identified on helix D. The carboxylate of the aspartate residue D90 and glutamate residue E94 coordinate two of the three $Mg^{2+}$ ions, $Mg^{2+}_A$ and $Mg^{2+}_C$, in *syn*,*syn*-bidentate geometry (Scheme 1.7 B).[53] The third $Mg^{2+}$ ion, $Mg^{2+}_B$, is coordinated by the three amino acid residues **N219**, **S223** and **E227**, that complement the second metal binding motif (NSE). Multiple metal interactions are responsible for the coordination of the pyrophosphate $PP_i$. In the absence of pyrophosphate in *apo*-aristolochene synthase the trinuclear magnesium cluster was not coordinated and substantial structural changes were observed, indicating that binding of the $Mg^{2+}_{A-C}$-$PP_i$-complex initiates the closing of the active site and triggers ionization of the diphosphate.[53]




**Scheme 1.7** Formation of aristolochene **68** in *Aspergillus terreus*: **A,** overview of aristolochene synthase in complex with $Mg^{2+}$ and inorganic pyrophosphate (PDB: 2OA6); **B,** active site illustrating the coordination of the trinuclear metal cluster and $PP_i$; green spheres = $Mg^{2+}$ ions; **C,** proposed mechanism of aristolochene **68** formation.[39]

Concerted electrophilic attack from the C-10/C-11 olefin leads to formation of the *E,E*-germacradienyl cation **57** (Scheme 1.7 C). Deprotonation by the general base/general acid $PP_i$ affords the first neutral intermediate (*S*)-germacrene A **59**. Reprotonation at the C-6 carbon is then immediately succeeded by electrophilic attack of the C-7 carbocation that leads to formation of the eudesmane cation **69**. Hydride transfer (C-2 → C-3), methyl migration (C-7 → C2) and final stereoselective deprotonation furnishes the product (+)-aristolochene **68** (Scheme 1.7 C).[39,53]

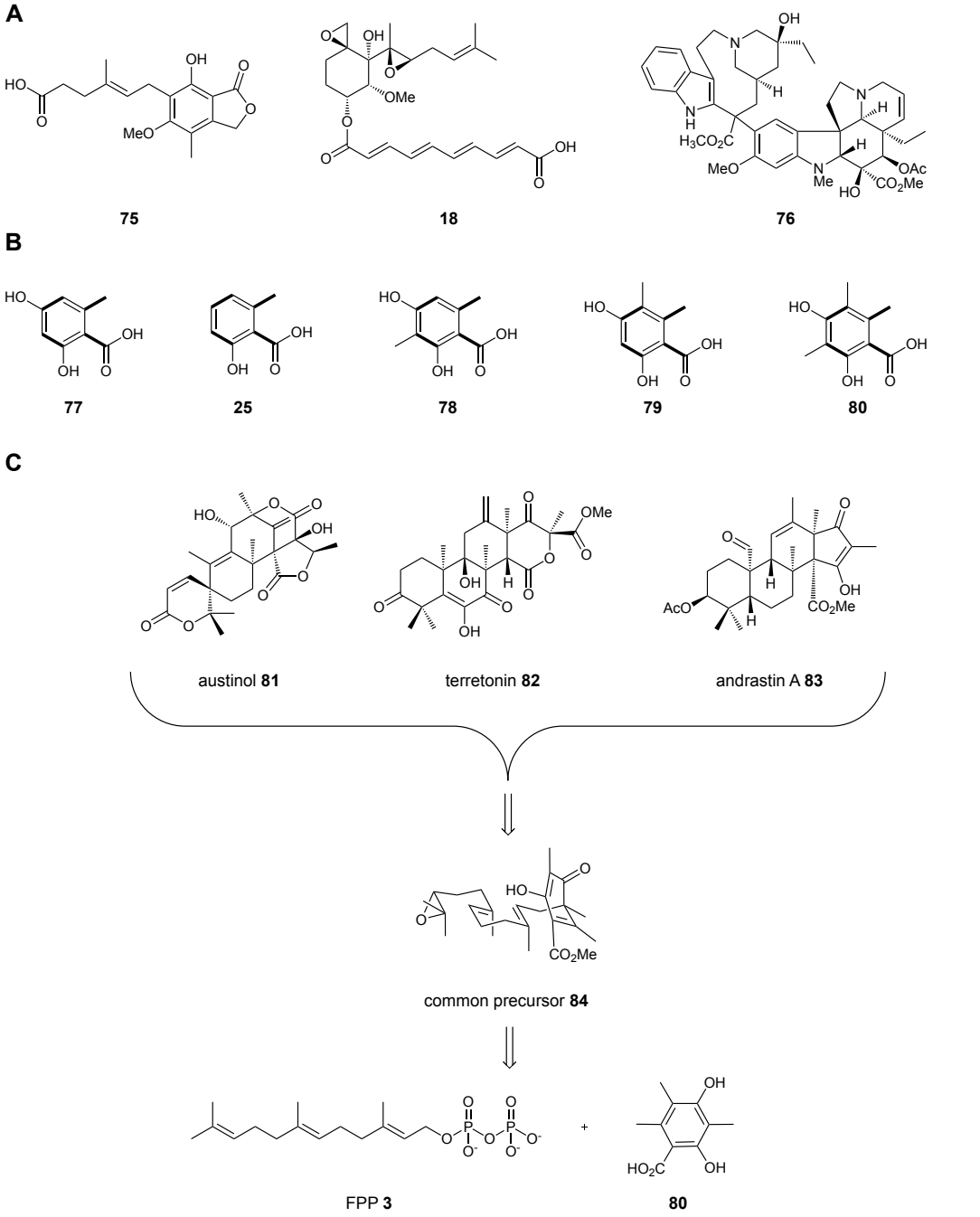
In contrast to class I terpene cyclases, canonical class II terpene cyclases catalyse cyclization of isoprenoid precursors by protonation of olefin or epoxide moieties (Scheme 1.8).[55] One such example is merosterolic acid synthase (MstE), which is responsible for formation of merosterolic acid A **72**. In the case of **72**-formation, key aspartate D109 protonates the terminal C-14/C-15 alkene in the isoprenoid precursor 3-geranylgeranyl-3,4-dihydroxybenzoate **73**. Brønsted acid catalysis initiates a cyclization cascade that drives formation of the pentacyclic core in **74**. Base (Glu339)-catalysed deprotonation of the *meta*-hydroxy group within the dihydroxybenzoate subsequently quenches the cascade reaction. Spontaneous deprotonation at C-6, followed by keto-enol tautomerism, drives reconstitution of aromaticity in the pathway product **72**.[55] Initiation of terpene cyclization by an aspartate such as D109 (typically found within a Dx**D**D motif; boldface indicates protonation) is the conserved feature of all class II canonical terpene cyclases.[39,55]




**Scheme 1.8:** Mechanism of canonical class II terpene cyclases illustrated for the biosynthesis of merosterolic acid **72**; Asp = aspartate; Glu = glutamate; B = general base.

## 1.4. Meroterpenoid Biosynthesis

Meroterpenoids are hybrid secondary metabolites of mixed biosynthetic origin.[56] The term `meroterpenoid´ was first coined by CORNFORTH in 1968, with the prefix `mero´ being derived from the Greek word *merus* (`part´/`partial´).[20] It summarizes a large number of natural products that are of partial terpene biosynthetic origin. Meroterpenoids have been isolated from all kingdoms of life, including plants, animals and microorganism - but fungi are particularly prolific producers of these metabolites.[20] A comprehensive review by CHEN *et al.* lists 1585 fungal meroterpenoids, that were isolated between 2009-2019.[56] Based on the biosynthesis of the non-terpenoid part, meroterpenoids can be divided into polyketide-derived meroterpenoids and non-polyketide-derived meroterpenoids. Polyketide derived meroterpenoids contribute to more than 45% of all isolated meroterpenoids. Other notable sources of the non-terpenoid parts are typically derived from the shikimate pathway or indoles.[56]

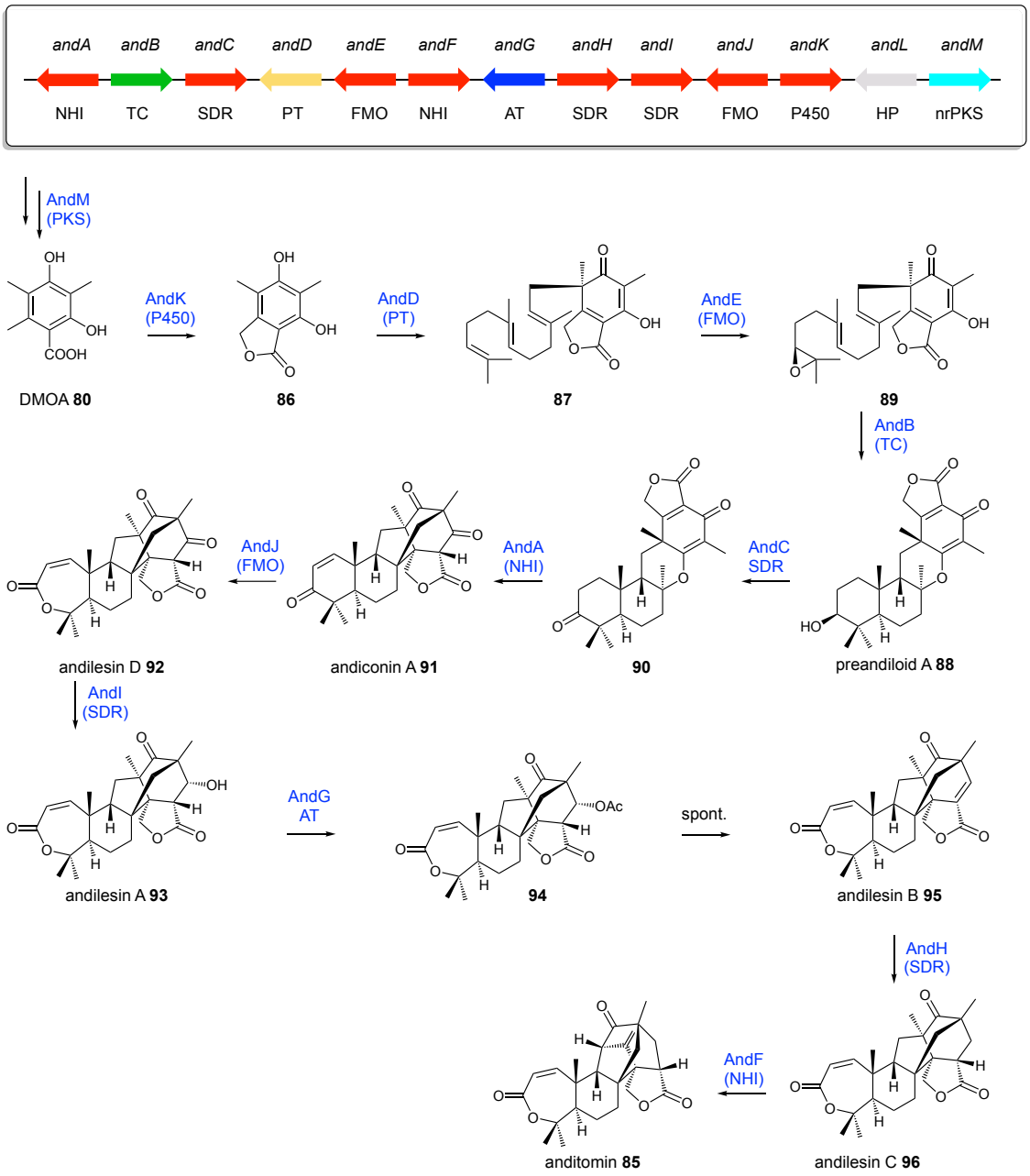



**Scheme 1.9** Meroterpenoid natural products: **A,** chemical structures of representative meroterpenoids; **B,** orsellinic acid derived tetraketides with intact acetate units highlighted; **C,** common biosynthetic origin of austinol **81**, terretonin **82** and andrastin A **83**.

Meroterpenoids are a structurally highly diverse class of natural products that display a broad range of biological activities (Scheme 1.9 A). Prominent examples include mycophenolic acid **75**, a potent inhibitor of inosine 5´-monophosphate dehydrogenase. Its derivative, mycophenolate mofetil, is widely used as an immunosuppressant drug.[57–59] Fumagillin **18** is a clinically deployed agent used for the treatment of microsporidiosis and amebiasis.[59–61] Vinblastine **76** is used in chemotherapy as a very potent anti-cancer agent.[62]

Polyketide-derived meroterpenoids are the most well-studied class of meroterpenoids. In particular tetraketide-derived compounds have been studied intensively.[56,59,63] Tetraketide-meroterpenoids are derived from orsellinic acid **77** and its derivatives 6-methylsalicylic acid **25**, 3-methylorsellinic acid **78**, 5-methylorsellinic acid **79** and 3,5-dimethylorsellinic acid **80** (Scheme 1.9 B). As meroterpenoids are secondary metabolites of diverse biosynthetic origin there is no universal biosynthesis. However, investigations into the related meroterpenoids austinol **81**, terretonin **82** and andrastin A **83** (Scheme 1.10 C) revealed one `common´ approach of meroterpenoid formation that is also, in parts, encountered in many other examples.[64–67] In this, initial formation of a core non-reduced polyketide (3,5-dimethylorsellinic acid **80**) is followed by its prenylation. Stereoselective epoxidation at one olefin of the polyprenyl chain then affords a universal precursor **84**. Differences in the cyclization of the isoprenoid moiety give rise to different meroterpenoids. Variable oxidative modifications are frequently observed and tailor the final pathway products, leading to the structural differences observed in **81**-**83** (Scheme 1.9 C).

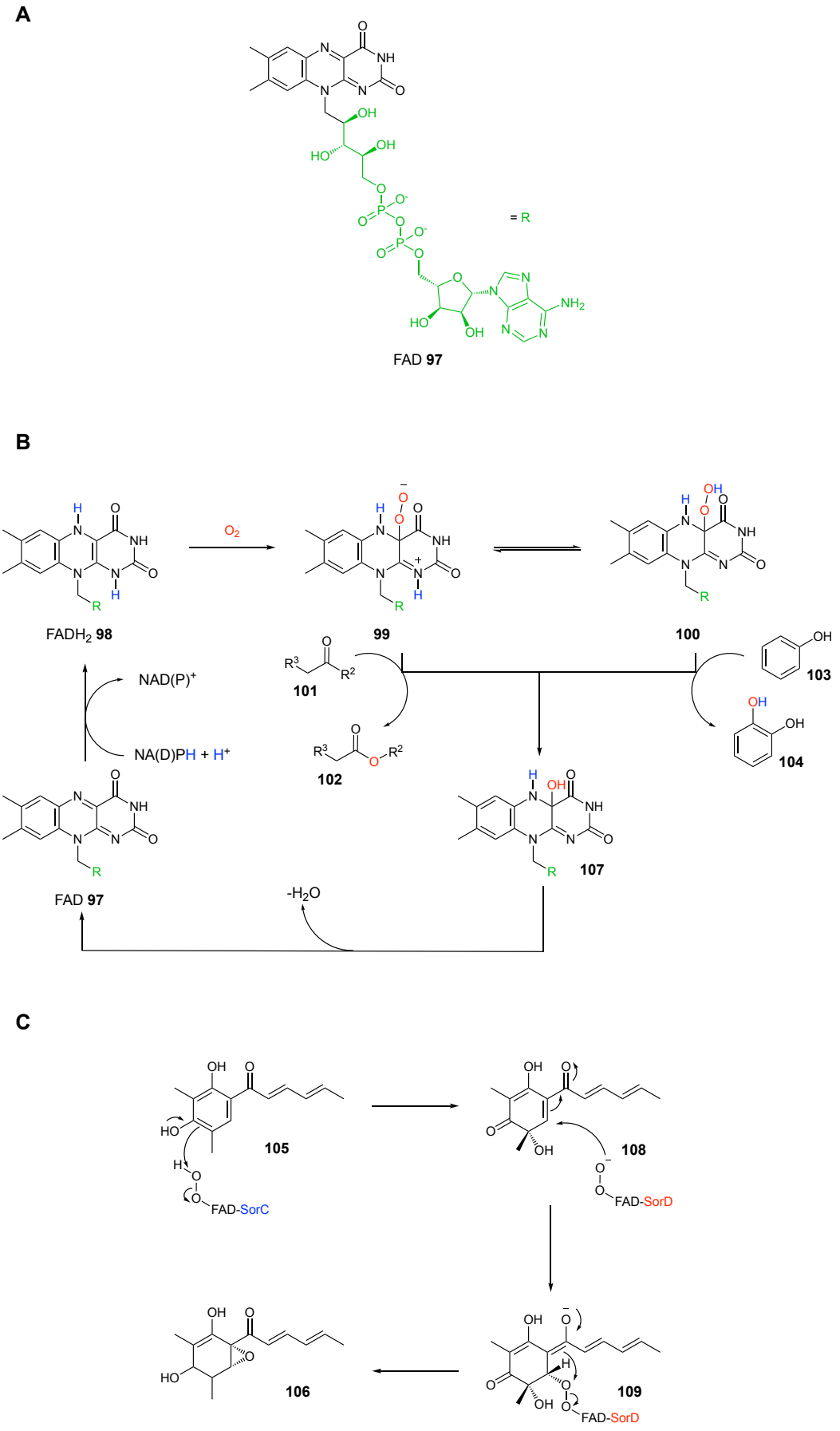
The completely elucidated biosynthetic pathway towards anditomin **85** serves to illustrate this underlying biosynthetic logic. ABE *et al*. identified the responsible BGC for formation of **85** in *Emericella variecolor* in 2014.[64] Reconstruction of the pathway in *Aspergillus oryzae* NSAR1 delineated each individual biosynthetic step. Biosynthesis starts with formation of 3,5-dimethylorsellinic acid **80** by the non-reducing PKS AndM (Scheme 1.10). Subsequent oxidation by AndK yields the pthalide 5,7-dihydroxy-4,6-dimethylphthalide **86** (DHDMP). The prenyltransferase AndD then catalyses transfer of the sesquiterpenoid FPP to **86**, generating farnesyl-DHDMP **87**. Epoxidation at the C-10/C-11 position by AndE enables the cyclization by the terpene synthase AndB and yields the intermediate preandiloid A **88**. A series of further oxidative and reductive events furnish anditomin **85** (Scheme 1.10). The biosynthesis of **85** serves to illustrate the often long and highly complex reaction sequences involved in meroterpenoid formation. In the case of **85**, 12 dedicated enzymes act sequentially and contribute to the final pathway product.




**Scheme 1.10** Complete biosynthetic pathway of anditomin **85**: Top Box = schematic representation of the *and* BGC; NHI = non-heme iron dioxygenase; TC = terpene cyclase; SDR = short-chain dehydrogenase; PT = prenyltransferase; FMO = FAD-dependent monooxygenase; AT = acetyltransferase; P450 = cytochrome P450; HP = hypothetical protein; nrPKS = non-reducing polyketide synthase; spont. = spontaneous acetate elimination.[64]

## 1.4.1. Oxidative Tailoring Enzymes

Oxidative enzymes play a pivotal role in tailoring of fungal secondary metabolites. The scope of catalyzed reactions ranges from `simple´ transformations (*e.g.,* hydroxylation, monooxygenation, dioxygenation) to complex ring expansions, ring closures and ring cleavages that go along with significant carbon skeletal rearrangements.[68] Since the reaction of atmospheric oxygen with spin-paired systems is spin-forbidden, nature has evolved numerous (co-factor dependent) strategies to activate atmospheric $O_2$.[68]




**Scheme 1.11** Catalytic cycle of FAD-dependent monooxygenases: **A,** chemical structure of FAD; **B,** mechanism for the activation of atmospheric oxygen by FAD-dependent monooxygenases;[69] **C,** proposed mechanism of oxidative dearomatization and epoxidation by SorC and SorD in *Trichoderma reesei*.[70]

*FAD-dependent monooxygenases*

Flavin adenine diphosphate (FAD **97**)-dependent monooxygenases are highly versatile oxidative enzymes in secondary metabolite formation. $FADH_2$/FAD **98**/**97** (Scheme 1.11 A) is used as prosthetic group to activate atmospheric oxygen.[69] The scope of catalyzed reactions ranges from hydroxylation, epoxidation, Baeyer-Villiger oxidation to non-oxidative carbon-hetero bond cleavage reactions.[69] The catalytic cycle of canonical FAD-dependent monooxygenases starts with reduction of the FAD **97** cofactor by NAD(P)H, resulting in the formation of the reduced $FADH_2$ **98** (Scheme 1.11 B). Binding of oxygen by **98** then leads to formation of the reactive (hydro)peroxyflavin species **99** or **100**, that are the oxygen transfering agents during catalysis.[69] Oxygenation occurs *via* nucleophilic attack of either **99** to the substrate (*e.g.* **101** to **102**) or *vice versa* by attack of the substrate to the hydroperoxyflavin **100** (*e.g.* as in **103** to **104**).[69] An intriguing example of two FAD-dependent monooxygenases are the multifunctional enzymes SorC and SorD from *Trichoderma reesei*, involved during the biosynthesis of sorbicillinoids. SKELLAM *et al*. recently demonstrated that the two enzymes catalyse the successive oxidative dearomatisation and epoxidation of sorbicillinol **105**, leading to the formation of epoxysorbicillinol **106** (Scheme 1.12 C).[70]

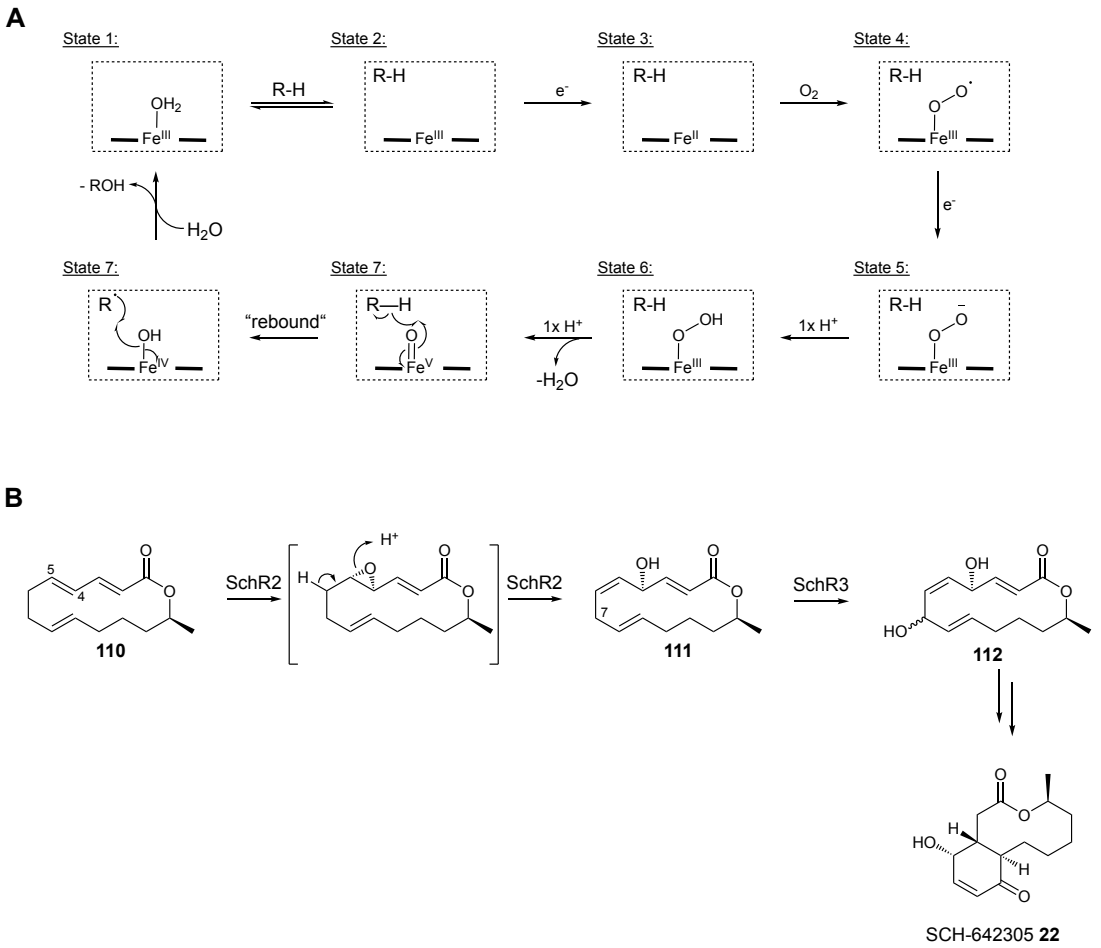
*Cytochrome P450 enzymes*

Cytochrome P450 enzymes represent another class of monooxygenases that are ubiquitously found in bacteria, fungi, plants and animals.[71] These enzymes catalyze a multitude of oxidative reactions. The scope of catalyzed reactions ranges from the hydroxylation of saturated carbon bonds to alkene epoxidation and the oxidation of aromatic compounds.[72]

Cytochrome P450s are defined by the presence of an iron(III) protoporphyrin-IX core that catalyzes the cleavage of atmospheric oxygen with NAD(P)H-consumption.[71,72] The cleavage of $O_2$ leads to the insertion of one oxygen atom into the substrate, while the second oxygen is reduced to water.[71] Cytochrome P450s are typically coupled to a reductase enzyme that catalyzes transfer of electrons from NADPH.[72]

The catalytic cycle of cytochrome P450 enzymes starts with the coordination of the substrate in the active site, leading to the displacement of a previously coordinated water molecule (Scheme 1.12 A; State 1 + 2). After substrate coordination a one-electron transfer from NADPH of the coupled reductase system leads to formation of a ferrous Fe(II)-species (Scheme 1.12 A; State 3). This Fe(II)-species can then bind molecular oxygen (Scheme 1.12 A; State 4). A second one-electron transfer (Scheme 1.12 A; State 5) and two successive protonation steps and loss of water afford a reactive Fe(V)-oxo species that mediates insertion of oxygen into the substrate (Scheme 1.12 A; State 6-8).[71,72]

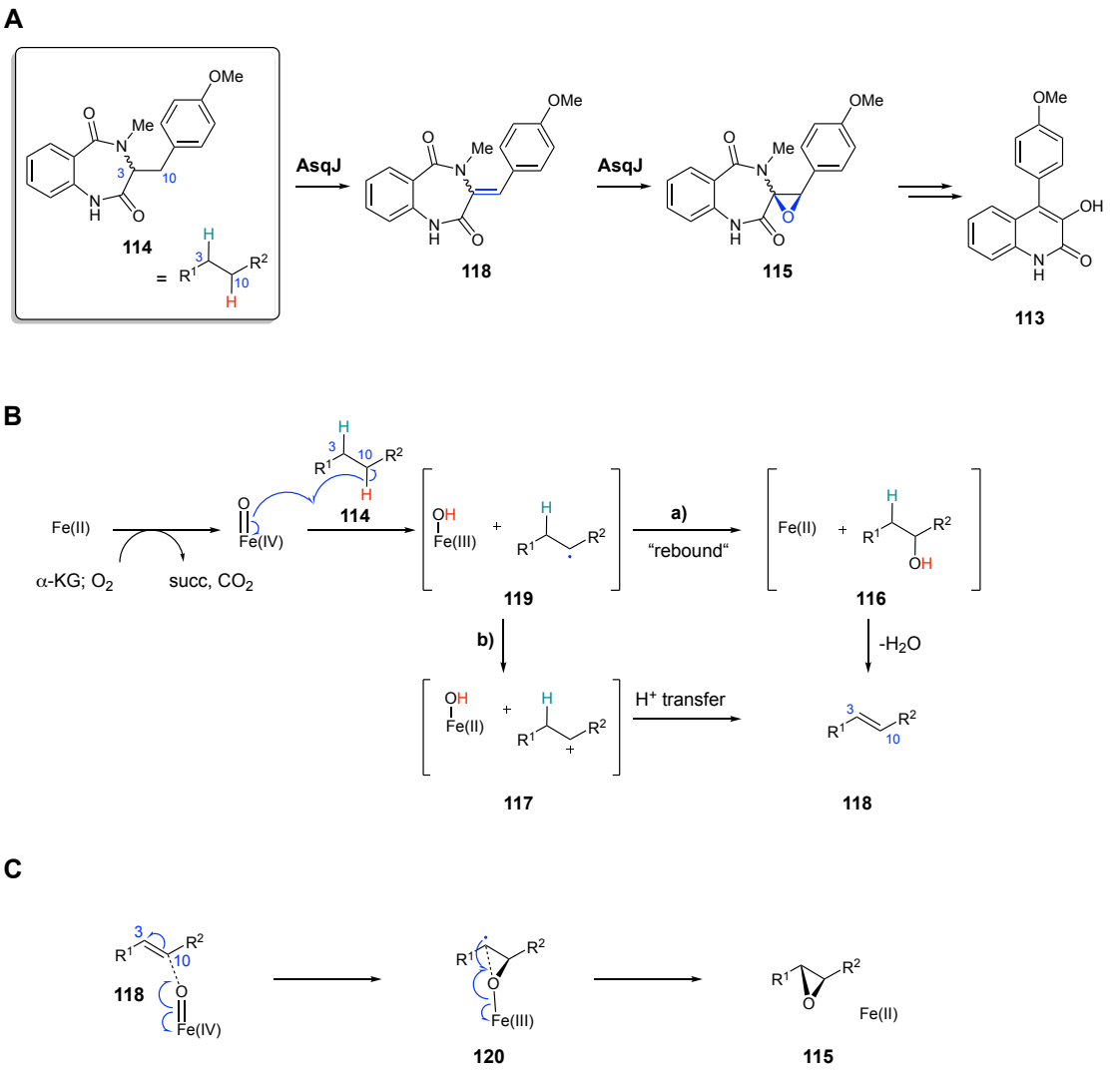
Examples of fungal cytochromes P450s are the enzymes SchR2 and SchR3 in the biosynthesis of the fungal macrolide SCH-642305 **22**.[23] SchR2 likely catalyzes epoxidation of the C-4/C-5 alkene in the polyketide precursor benquoine **110**. Epoxidation was proposed to be followed by epoxide-opening to the hydroxylated intermediate **111**. SchR3 then activates the saturated carbon C-7 and inserts a second hydroxyl moiety, yielding intermediate **112**, that is then further converted to SCH-642305 **22** (Scheme 1.12 B).[23] Likewise, a range of cytochrome P450 monooxygenases have been reported in the biosynthesis of fungal cytochalasans.[73]




**Scheme 1.12** Catalytic cycle of cytochrome P450 enzymes: **A,** catalytic cycle of the hydroxylation of hypothetical substrate R-H; R = hydrocarbon; dashed box = active site of P450; **B,** catalysed transformations in the biosynthesis of SCH-642306 **22** by the cytochrome P450 enzymes SchR2 and SchR3.[23]

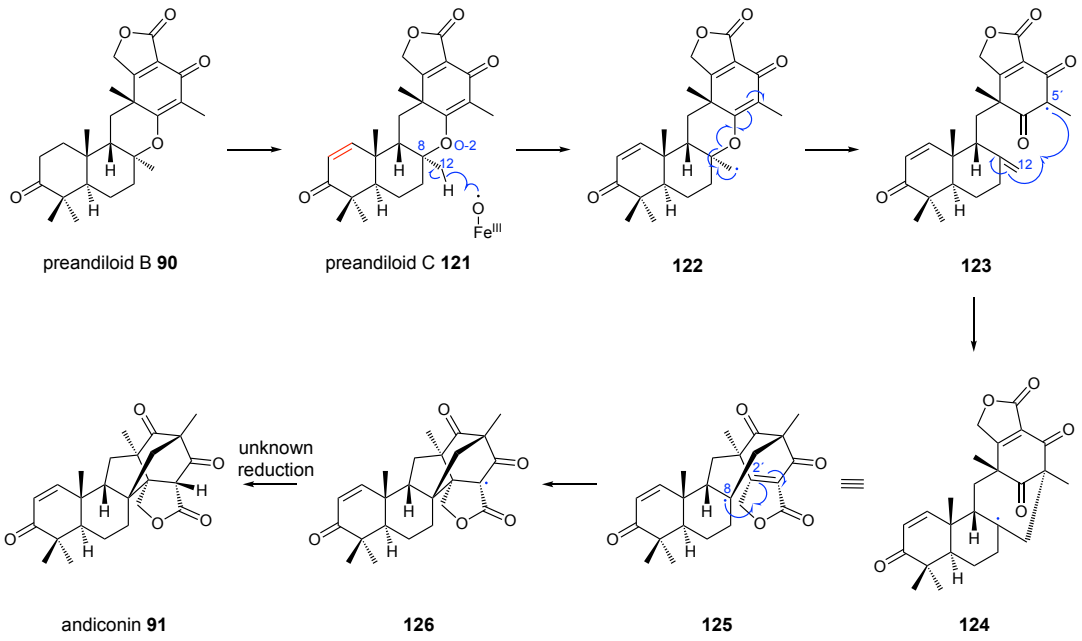
*Non-heme dependent oxygenases*

Non-heme iron dependent oxygenases (NHI) are versatile tailoring enzymes that catalyze diverse oxidative chemical reactions.[74] A major subgroup of these enzymes activates molecular oxygen and deploy α-ketoglutarate as a co-substrate.[75] One such example is the NHI enzyme AsqJ, involved in the biosynthesis of 4-methoxyviridicatin **113** in *Aspergillus nidulans*. AsqJ catalyzes the consecutive dehydrogenation and epoxidation of cyclopeptin **114**, leading to formation of **115** (Scheme 1.13 A).[76] The initial dehydrogenation proceeds through hydrogen abstraction from C-10, catalyzed by a short-lived $Fe^{IV}$-oxo species (Scheme 1.13 B). Abstraction of the second hydrogen atom then proceeds either *via* the hydroxylated intermediate **116** (Fe[II]-species; Scheme 1.13 B; path a]) or *via* the carbocation **117** (Fe[II]-species; Scheme 1.13 B; path b]).[76] A similar radical mechanism also explains the successive installation of the epoxide functionality in **115** (Scheme 1.13 C).[77]




**Scheme 1.13** Successive dehydrogenation and epoxidation during the biosynthesis of 4-methoxyviridicatin **113**: **A,** biosynthesis of **113** from cyclopeptin **114**; **B,** mechanism of dehydrogenation; **C,** mechanism of epoxidation; α-KG = α-ketoglutarate; succ = succinate.

α-Ketoglutarate dependent NHI enzymes are also implicated in the catalysis of significant structural rearrangements during meroterpenoid biosynthesis. An intriguing example is AndA from the previously discussed anditomin **85** biosynthetic pathway (Scheme 1.10).[64] AndA catalyzes the two-step reaction from preandiloid B **90** to andiconin **91** (Scheme 1.14). In a first step, dehydration of preandiloid B **90** leads to formation of the enone preandiloid C **121** (likely through a similar mechanism as reported for AsqJ). Subsequently the congested bridged-ring system in andiconin **85** is constructed.[74] Based on density functional theory (DFT) calculations ABE *et al*. proposed a radical mechanism in which hydrogen abstraction from C-12 initiates a reaction sequence that involves cleavage of the C-8/O-2 bond as well as C-12/C-5′ and C-8/C-2′ bond formation.[74]

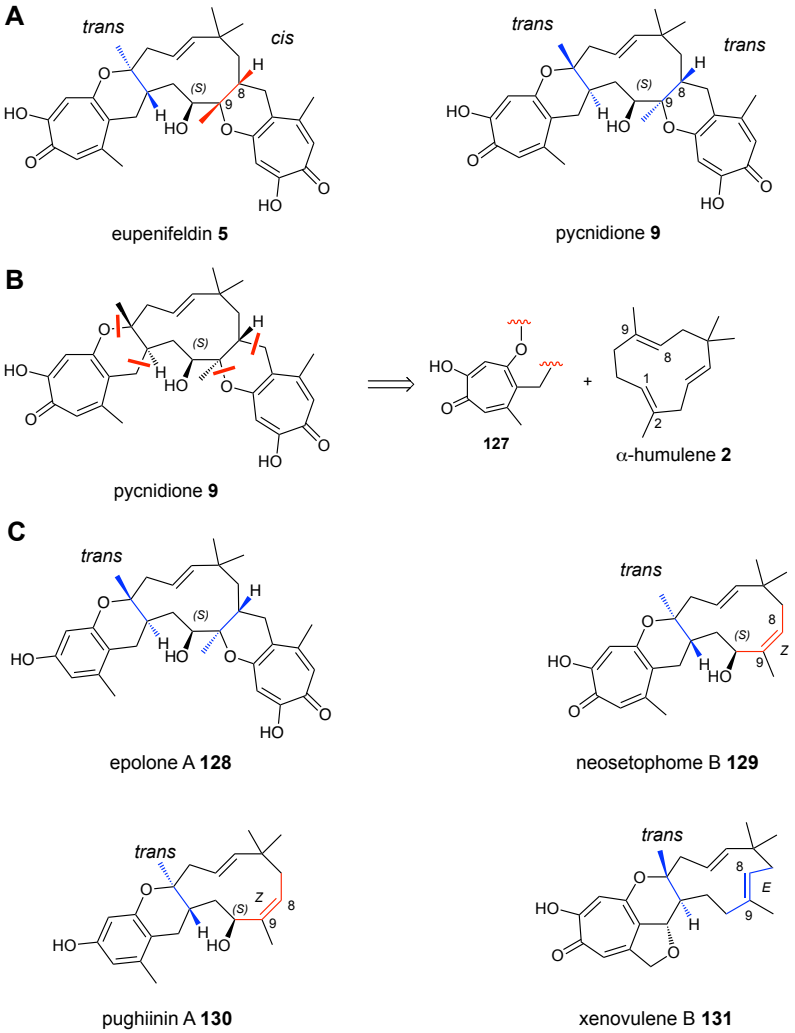



**Scheme 1.14:** Skeletal rearrangements catalysed by the NHI enzyme AndA in the biosynthesis of anditomin **85**; red = stereoselective dehydrogenation of **90**; blue = radical isomerization of **121**.[74]

## 1.4.2. Tropolone Sesquiterpenoids

Tropolone sesquiterpenoids (TS) are a subclass of meroterpenoid natural products that are exclusively produced by fungi. The first reported TS natural product, eupenifeldin **5**, was isolated in 1993 from culture broths of *Eupenicillium brefeldianum* ATCC 74184 (Scheme 1.15 A).[78] The pentacyclic structure and stereochemistry of **5** was deduced from detailed NMR and single-crystal X-ray analysis. It displays partial symmetry as two identically substituted tropolone nuclei are fused to a central 11-membered macrocycle *via* bridging dihydropyran rings.[78] Bistropolone **5** was also isolated from other fungal sources including *Kionochaeta ramifera* BCC 7585, *Phoma* sp. and *Neosetophoma* sp. MSX50044.[79–81] Original characterization of **5** by researchers from BRISTOL-MYERS SQUIBB revealed potent antitumor activity.[78] Eupenifeldin **5** was found to be active both in human colorectal carcinoma cell lines (HCT-116; $IC_{50}$ = 0.005 µg·mL$^{-1}$) and in a

murine leukemia model, leading to a prolonged life-span.[78] Recently OBERLIES *et al*. reported that the use of eupenifeldin-coated buttresses in a murine model of resection delayed tumor recurrence and improved the overall survival rate.[82]




**Scheme 1.15** Overview of the tropolone sesquiterpenoid family of natural products: **A,** chemical structures of the representative TS natural products eupenifeldin **5** and pycnidione **9**; **B,** retrobiosynthetic proposal for pycnidione **9** as suggested by HARRIS *et al*.;[83] **C,** structures of related TS natural products that differ in hydroxylation, ring-size and stereochemistry.

Almost in parallel to the isolation of eupenifeldin **5**, HARRIS *et al*. reported the isolation of pycnidione **9** from an unidentified *Phoma* sp.[83] Pycnidione **9** displays the identical pentacyclic ring structure as observed for **5**, but differs in the stereochemistry (Scheme 1.15 A). In pycnidione **9**, the eastern dihydropyran-tropolone moiety is *trans*-fused to the central 11-membered macrocycle (at C-8/C-9), whereas in eupenifeldin **5** the eastern dihydropyran-tropolone moiety is *cis*-fused. The difference in stereochemistry at the C-8/C-9 carbons (*cis* or *trans* fusion) is observed in all members of this class of natural products and allows the assignment of the compounds to two subclasses: pycnidione-type TS (hitherto also referred to as xenovulene-type

TS) display *trans*-fusion at the C-8/C-9 position and eupenifeldin-type TS display *cis*-fusion at the C-8/C-9 position.

Promising bioactivity was also reported for pycnidione **9**. The bistropolone **9** was originally identified as a potent antiarthritic agent and inhibits stromelysin with an $IC_{50}$ = 31 µM.[83] Investigations into its bioactivity by CHI-LI CHUNG *et al*. revealed additional potent antitumor activity: **9** inhibits the proliferation of A549 cells with a growth inhibition value $GI_{50}$ = 9.3 nM.[84] HARRIS *et al*. also made the first retrobiosynthetic proposal and suggested that **9** is derived from the terpene α-humulene **2** and two tropolone nuclei similar to **127** (Scheme 1.15 B).

Since the discovery of **5** and **9** a series of related tropolone sesquiterpenoids have been isolated from a range of fungal microorganisms that illustrate the structural diversity of this class of natural products. In 1995 HOUCK *et al*. isolated epolone A **128** from the unidentified ascomycete OS-F69284.[85]

Epolone A **128** was concomitantly produced with pycnidione **9** (Scheme 1.15 C). In **128** the western tropolone nucleus is replaced by a tetrasubstituted benzene ring of unknown biosynthetic origin. Epolone A **128** induces the production of erythropoietin gene expression in human cells by a factor of five (**128**-concentration: 1 µM).[85]

Neosetophome B **129** was isolated from *Neosetophoma* sp. MSX50044.[81] Contrary to the previous discussed TS natural products, **129** is a monotropolone in which only one tropolone nucleus is linked to humulene at the western face. Notably, in neosetophome B **129** (and all other monotropolones from the eupenifeldin class of TS) the C-8/C-9 alkene is *Z*-configured. The monotropolone **129** is concomitantly produced with eupenifeldin **5** (Scheme 1.15 C) and is a micromolar antitumor agent towards a panel of human cancer cell lines, including breast, ovarian and lung cancer.[81]

With pughiinin A **130** PITTAYAKHANJONWUT *et al*. reported another monosubstituted TS natural product that contained the previously described tetrasubstituted benzene moiety in place of the tropolone nucleus (Scheme 1.15 C). Pughiinin A **130** is produced alongside pycnidione **9** in *Kionochaeta pughii* BCC 3878 and is a potent antiplasmodial agent ($IC_{50}$ = 2.4 µg·mL$^{-1}$).[86] Finally, Xenovulene B **131** was isolated from the submerged fermentation of *Acremonium strictum* and displays an unprecedented tetrahydrofuran linked to the tropolone/dihydropyran rings (Scheme 1.15 C).[87] Contrary to neosetophome B **129**, in xenovulene B **131** the C-8/C-9 alkene is *E*-configured. The *E*-configuration is mirrored in all other monotropolones from the

xenovulene-type TS. The reason for the difference in stereochemistry at this position remains elusive to date.

In total, more than 20 TS and related natural products have been reported across the last 30 years. The chosen examples serve to illustrate conserved structural features and optional modifications. All TS share the conserved dihydropyran/humulene core but differ in whether they constitute mono- or disubstituted derivatives of humulene. Hydroxylation at C-10 is optional, but all isolated hydroxylated metabolites display the (10*S*)-configuration, suggesting a conserved enantioselective enzymatic reaction responsible for hydroxylation. While the tropolone motif is most frequently observed, replacement with a highly substituted benzene is also possible (*e.g.* **130**).

The 11-membered macrocycle α-humulene **2** itself is a common component of plant essential oils.[88] The terpene **2** and its oxidized congeners are largely responsible for the hoppy flavour of beer and its name originates in **2** being isolated from *Humulus lupulus* (also referred to as Common Hop).[89] Humulene **2** has been rarely isolated from fungi, but examples include *Fusarium fujikuroi*, *Colletotrichum acutatum* and *Stereum hirsutum*.[90–92] Consideration of the isolated tropolone sesquiterpenoids, but also synthetic chemistry using α-humulene **2** as substrate, reveals a preferred regioselectivity in α-humulene functionalisation: Despite displaying two triply substituted alkenes (C-1/C-2 and C-8/C-9) the C-1/C-2 alkene seems to be more reactive as all monotropolone derivatives isolated so far show linkage to the dihydropyran at this position and chemical epoxidation occurs preferably at this position as well.[89] HERMANS *et al*. used density-functional theory calculations to identify distinct humulene conformations: A strong hyper-conjugative $\sigma_{C\alpha\text{-}C\beta}$-$\pi_{C=C}$ orbital overlap was found to discriminate the C-1/C-2 double bond over the second triple substituted alkene, making it more reactive.[89]

## 1.5. Investigation of Natural Product Biosynthesis

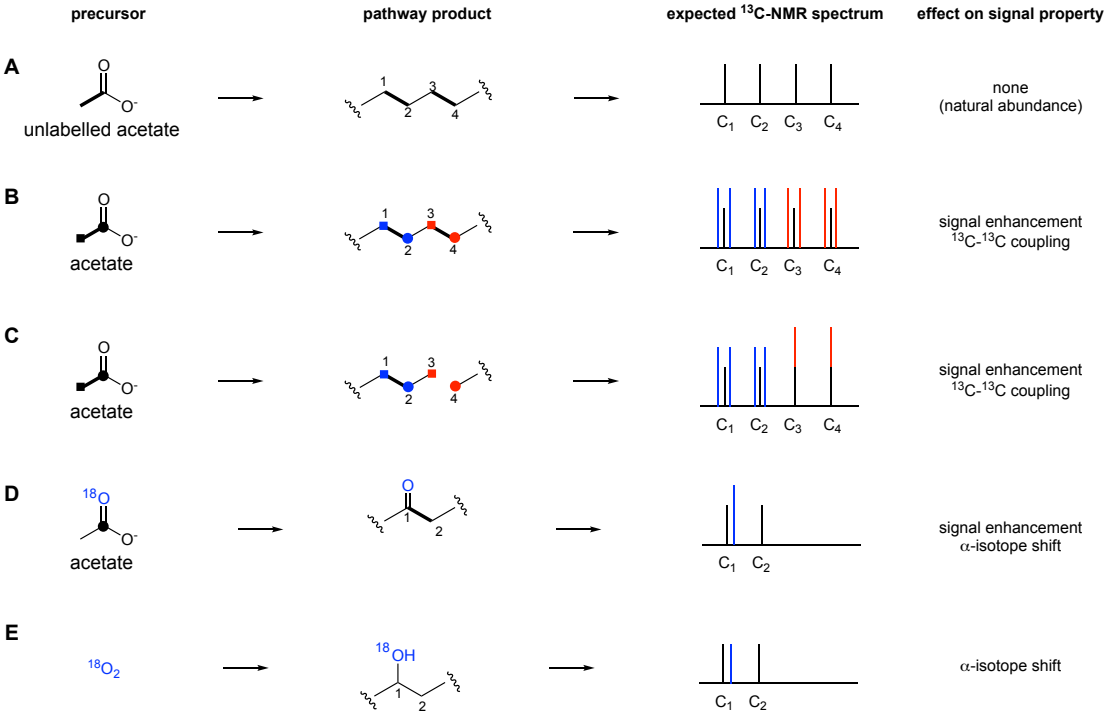
The biosynthesis of natural products has been investigated for decades and numerous genetic, biological and chemical techniques and methods have been used to elucidate entire pathways. The natural product chemist´s toolbox includes technologies such as genetic inactivation of key biosynthetic enzymes (knockout experiments), gene silencing, isotopic labelling and heterologous expression. In this study the two latter technologies were successfully deployed to investigate (and engineer) aspects of tropolone sesquiterpenoid biosynthesis.

### 1.5.1. Isotopic Labelling

The use of stable isotopes significantly spurred our understanding of metabolic pathways and it has contributed enormously to validate the biosynthetic origin of entire natural products or individual functional moieties.[93] Inspiring work by SCHOENHEIMER and RITTENBERG in 1935 laid the foundation for this success: Deuterium-labelled fatty acids and steroids were used to trace their fate in a living organism and could thus identify key intermediates in cholesterol metabolism.[94]

Isotopes of elements occupy identical positions in the periodic table of elements - but they differ in the number of neutrons in the nucleus, thus allowing for analytical differentiation.[95] In early work radioactive isotopes such as $^{14}C$, $^{32}P$ and $^{35}S$ were used.[96] Radioactive isotopes offer the great advantage that they can be detected in minute quantities – however, information on the position of the label in the labelled molecule often requires time-consuming chemical degradation procedures.[96] In the course of rapidly evolving NMR technologies in the 1970s a shift towards non-radioactive isotopes such as $^{13}C$, $^{15}N$, $^{18}O$ and $^{2}H$ occurred.[95] By comparison of $^{13}C$-NMR spectra from labelled and unlabeled compounds differences in spin-spin couplings or isotopically shifted signals would give information about label incorporation and position of the label in the molecule.[97]

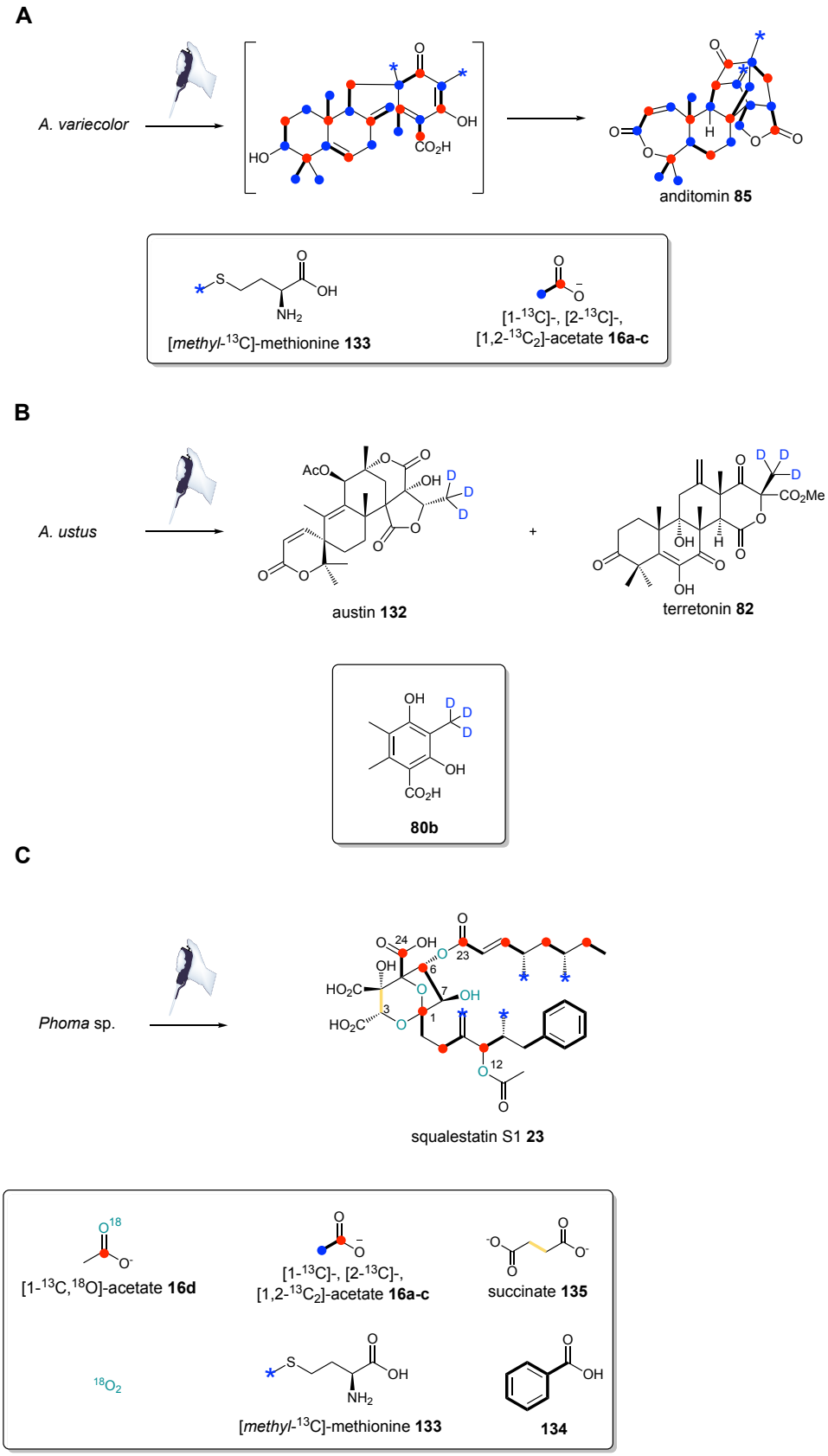
For example, incorporation of an *intact* [1,2-$^{13}C_2$]-acetate into a polyketide leads to $^{13}C$-$^{13}C$ spin coupling of the corresponding carbon pair in the polyketide. The resulting carbon NMR signal shows distinct $^{13}C$-$^{13}C$ coupling satellites flanking the natural abundance signals (Scheme 1.16 A and B).[97] In case the carbon-carbon bond within the acetate unit is disrupted during polyketide formation,$^{13}C$-$^{13}C$ coupling is lost and solely enhanced singlets for the corresponding carbons will be observed (Scheme 1.16 C).[97] The observed coupling constants (30 - 80 Hz) are dependent on the hybridization of the carbon atoms, with $sp^2$-$sp^2$ bonds displaying larger couplings (~ 60 Hz) than $sp^2$-$sp^3$ (~ 45 Hz) or $sp^3$-$sp^3$ (~ 35 Hz) bonds.[97]




**Scheme 1.16** The use of isotopic labelling in natural product biosynthesis: Effects of label incorporation on $^{13}$C-NMR spectra; modified from SIMPSON 1998.[97]

Oxygen-substitutions in fungal polyketides are often derived from acetate C-1 oxygens during polyketide biosynthesis itself or post-translationally inserted by *e.g.* cytochrome P450 enzymes or flavin-dependent monooxygenases. Differentiation between these mechanisms can be achieved through incorporation experiments with atmospheric [$^{18}$O]-labelled oxygen or with doubly labelled [$^{18}$O, 1-$^{13}$C]-acetate. Incorporation of $^{18}$O can be monitored by $^{13}$C-NMR spectroscopy, using the carbon atom as a "reporter" nucleus: Incorporation of $^{18}$O leads to an upfield α-isotope shift (~ $\Delta\delta$ = 0.05 ppm) of the $^{13}$C chemical resonance (Scheme 1.16 D + E).[97] In the case of double-labelled [$^{18}$O, 1-$^{13}$C]-acetate, incorporation of $^{13}$C-labelled C-1 also leads to an enhanced carbon resonance at this position, given that the level of incorporation exceeds 1% and making the experiment more sensitive (Scheme 1.16 D).

The previously outlined biosynthesis of the meroterpenoids anditomin **85**, andilesin C **96**, austin **132** and terretonin **82** was also first investigated using isotope feeding studies. In early studies T. J. SIMPSON studied the incorporation of [1-$^{13}$C]-, [2-$^{13}$C]- and [1,2,-$^{13}$C$_2$]-acetates **16a-c** and [*methyl*-$^{13}$C]-methionine **133** into anditomin **85** (Scheme 1.17 A).[98] The interpretation of the resulting labelling pattern confirmed the meroterpenoid biosynthetic origin: Labelling of the $C_{25}$ metabolite suggested **85** to be derived from the direct prenylation of 3,5-dimethylorsellinic acid **80** and subsequent cyclization (Scheme 1.17 A).[99] Indeed, more sophisticated labelling studies in the related austin **132**/terretonin **82** pathway in *Aspergillus ustus* confirmed the polyketide origin.




**Scheme 1.17** Use of isotopic labelling in the elucidation of fungal meroterpenoid biosynthetic pathways: **A,** isotopic labelling of anditomin **85** in *E. variecolor*; **B,** isotopic labelling of austin **132** and terretonin **82** in *A. ustus*; **C,** labelling of squalestatin S1 **23** in *Phoma* sp.

A synthetic trideuteriomethyl-analogue **80b** of 3,5-dimethylorsellinic acid **80** was incorporated both into austin **132** and terretonin **82** and successfully validated the tetraketide origin of these metabolites (Scheme 1.17 B).[99]

A series of labelling experiments also elucidated the biosynthetic origin of the squalene inhibitor squalestatin S1 **23**. Incorporation studies with [1-$^{13}$C]-, [2-$^{13}$C]- and [1,2-$^{13}C_2$]-acetate **16a-c** revealed the presence of a tetraketide and a hexaketide chain that forms the carbon backbone.[100] The hexaketide chain in **23** is primed with an unusual benzoic acid **134**-derived starter unit, as confirmed by the incorporation of [*carboxy*-$^{13}$C]- and [*aromatic*-$^{13}C_6$]-benzoic acid.[100] Further complementary studies using [*methyl*-$^{13}$C]-methionine **133** and [2,3-$^{13}C_2$]-succinate **135** identified the biosynthetic origin of the remaining carbon atoms. Finally, incorporation of $^{18}O_2$ and [1-$^{13}$C,$^{18}$O]-acetate **16d** in separate experiments elucidated the origin of all oxygen atoms within **23** and allowed to differentiate between polyketide origin (carbonyl oxygens on carbon C-23 and C-24) and origin from atmospheric oxygen (oxygens on carbon C-1, C-3, C-6, C-7 and C-12; Scheme 1.17 C).[100]

The selected examples serve to illustrate the versatile utility of isotopic labelling experiments. In many cases labelling experiments pave the way for following detailed genetic and biochemical investigations and they facilitate the proposal of biosynthetic pathways and the identification of candidate BGC.

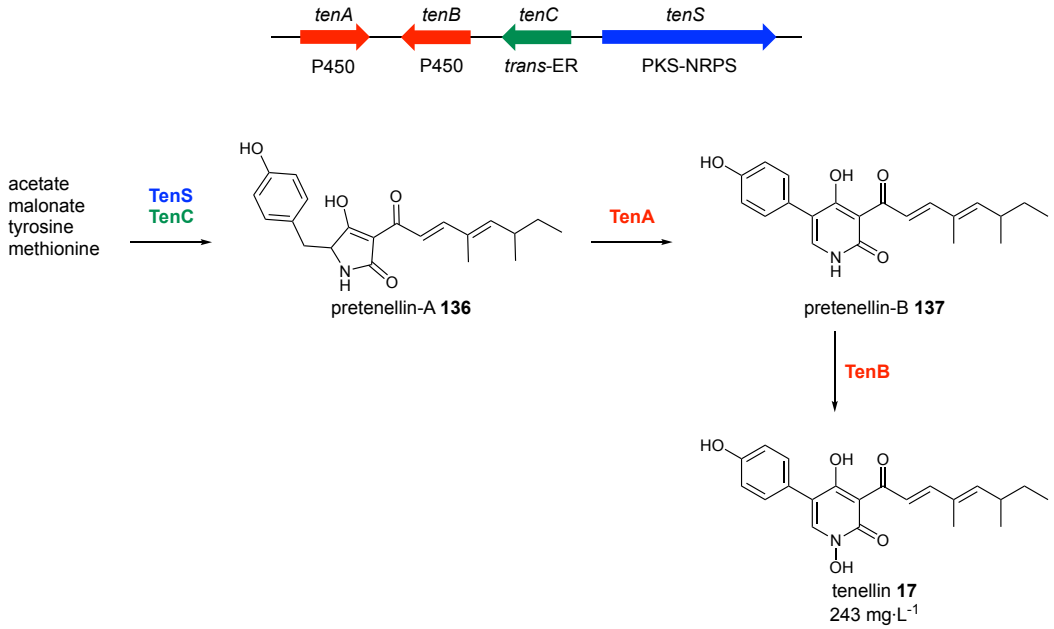
Despite impressive advances in the area of genomic and biochemical methodology stable isotope labelling remains a significant tool to address biosynthetic questions. TRENTI and COX recently successfully deployed feeding studies to revise the structure of the fungal phytotoxin phyllostictine A and B.[101] Likewise, the research group of PROF. JEROEN DICKSCHAT makes excellent use of labelled isoprenoids to elucidate the complicated mechanisms of carbocationic rearrangements during terpene biosynthesis, as most recently demonstrated for the reaction mechanism of catenul-14-en-6-ol synthase.[102]

### 1.5.2. Heterologous Expression

In recent years heterologous expression of partial and complete biosynthetic pathways has become an increasingly important tool to study natural product biosynthesis. For this purpose, individual genes or sets of genes are transferred from a donor strain to a genetically tractable host organism. Bacteria are the most convenient expression hosts, as genetic manipulation and cultivation techniques are often already established. However, due to the inability of bacteria to process eukaryotic introns, codon bias, reported difficulties in correct folding of fungal polypeptides and lack of post-translational phosphopantetheinylation-machinery, this is often not

successful for fungal pathways.[103] Instead, yeast and fungal hosts have been successfully deployed for fungal pathway reconstitution.[104]

The first complete fungal BGC that was reconstituted in a fungal host was the tenellin BGC from *Beauveria bassiana*. As expression host *Aspergillus oryzae* was successfully used, a filamentous fungus with GRAS status ("generally recognized as safe") that has been used for thousands of years in the japanese fermentation industry.[105] For example, transformation of *A. oryzae* with *tenS* (encoding a hybrid PKS-synthase/NRPS-synthetase), *tenA* and *tenB* (encoding two cytochrome P450s) and *tenC* (encoding a *trans*-acting ER-domain) afforded tenellin **17** at impressive 243 mg·L$^{-1}$ (Scheme 1.18).[103]

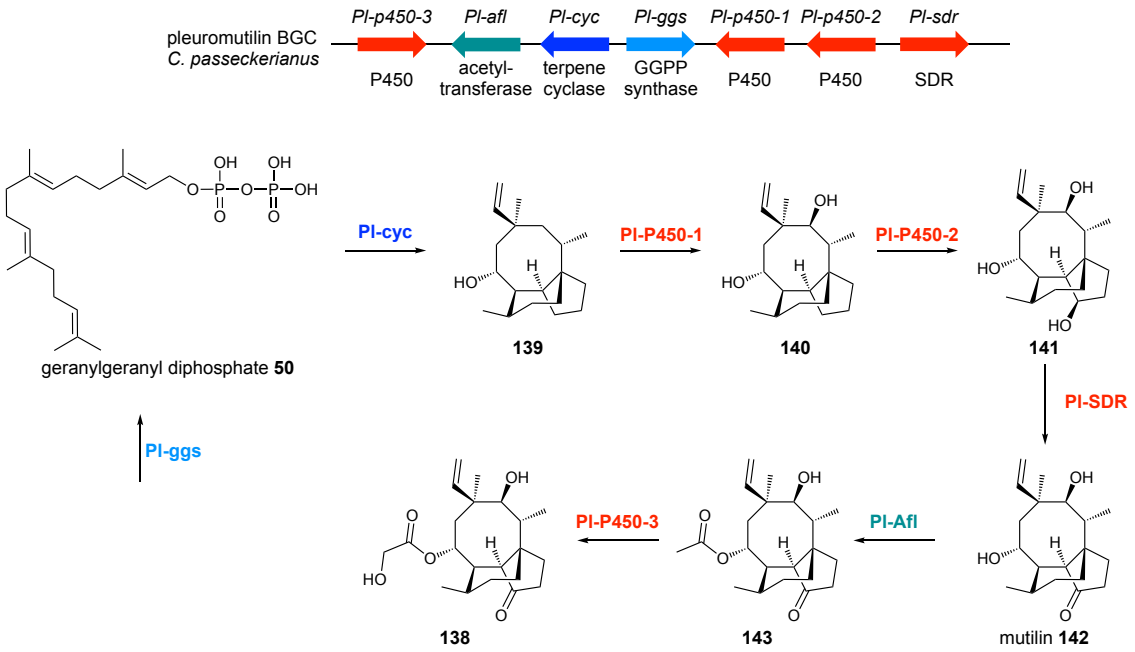



**Scheme 1.18:** Reconstitution of the tenellin **17** biosynthetic pathway in *A. oryzae*; P450 = cytochrome P450; ER = enoylreductase; PKS-NRPS = polyketide synthase/non-ribosomal peptide synthetase.

Successive advances in genetic tools to manipulate fungal genomes (*e.g.* transformation protocols, CRISPR-Cas, *etc*.), improved host systems (not limited to *A. oryzae*) and technologies for multiple gene expression nowadays allow heterologous expression to be used routinously in a standardized fashion.[104] Heterologous expression is deployed for a number of different purposes, *e.g.*: **a),** elucidation of biosynthetic pathways; **b)**, titre improvement; **c),** gene cluster activation; and **d),** combinatorial biosynthesis.

Many natural products are produced in their native ecological context at very low concentrations or at highly specific timepoints that complicate their isolation. One such example is the avirulence conferring enzyme 1 (ACE1) metabolite produced by *Pyricularia oryzae*. *P. oryzae* is a major fungal pathogen of rice plants, leading to large annual losses in crops. Resistance in certain rice

plants has been linked to avirulent strains of *P. oryzae* that express the *ACE1* BGC. However, as this BGC is only expressed during a very short period upon fungal appressorium penetration into the plant tissue, the isolation and characterization of the *ACE1* metabolite has not yet been possible.[106,107] In some cases heterologous expression of biosynthetic gene clusters can help to overcome these issues, by activating silent clusters or significantly increasing the production titres. For example, heterologous expression of the entire seven gene pleuromutilin **138** biosynthetic gene cluster in *A. oryzae* increased production by impressive 2106 %![108] Subsequently, co-expression of individual (sets of) genes in *A. oryzae* was successfully used to delineate the complete biosynthetic pathway: Pleuromutilin **138**-formation proceeds through the initial cyclization of geranylgeranyl diphosphate **50**, that affords the 14-hydroxytricyclic diterpene **139**. Stepwise inclusion of genes encoding *Pl*-P450-1 and *Pl*-P450-2 resulted in formation of diol **140** and triol **141**, respectively. Successive incorporation of genes encoding the short-chain dehydrogenase *Pl*-SDR, acetyl transferase *Pl*-Afl and *Pl*-P450-3 led to full reconstitution of pleuromutilin biosynthesis (Scheme 1.19).[109]




**Scheme 1.19:** Proposed biosynthetic pathway leading to formation of **138** in *A. oryzae*.

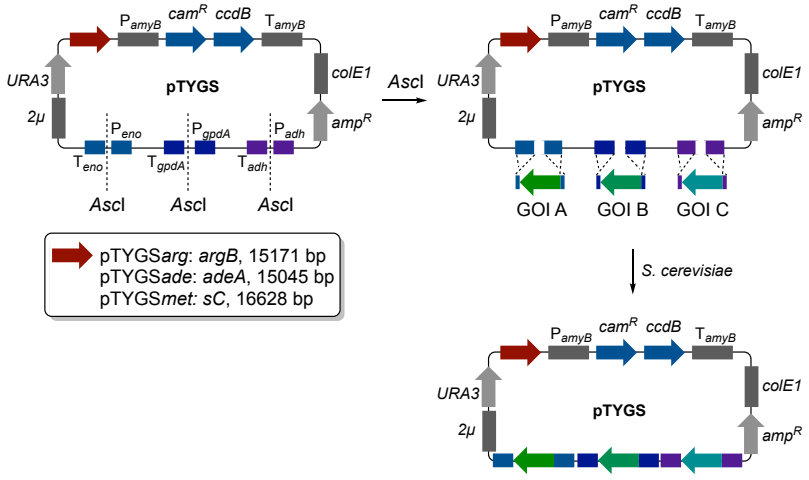
In respect to pathway elucidation, heterologous expression is often used when more `classical´ methods (*e.g.* genetic knockout experiments, gene silencing, *etc.*) have failed. Basidiomycetes for example are often found to be genetically intransigent and can be resistant to means of transformation and genetic manipulation.[110] This genetic intractability has impeded investigations into many intriguing and important natural products for decades. The psychoactive metabolite psilocybin, for instance, is produced by a number of basidiomycete fungi and was first characterized in 1959.[111] In the absence of genetic tools to manipulate the host organism, however, its biosynthetic pathway was only elucidated 60 years later. In 2017 HOFFMEISTER *et*

*al*. expressed all pathway specific genes in *E. coli* and reconstituted the pathway *in vitro*.[111] Subsequent transfer of the entire pathway into *Saccharomyces cerevisiae* and strain optimization afforded a psilocybin producing strain that affords psilocybin at titres around 627 ± 140 mg·L$^{-1}$.[112] Such methods are now regularly used for the total mycosynthesis of fungal natural products.[113]

The majority of fungal BGC are transcriptionally silent or active at a very low level under the typically deployed culture conditions in the laboratory.[114] Fungal cultivation in the lab and cluster activation thus represents one of the major bottlenecks in the discovery of new fungal metabolites. Heterologous expression of silent clusters in genetically tractable and well-studied host organisms has become a very promising tool for the rapid screening of large sets of fungal BGC. HILLENMEYER *et al.* recently reported a synthetic biology platform, HEx (Heterologous EXpression), that represents a scalable approach to rapidly assemble and screen cryptic fungal BGC in *S. cerevisiae*.[115] In this study the authors reconstituted 41 fungal BGC under strong auto-inducible promotors. In the case of 22/41 clusters, expression resulted in the formation of detectable amounts (by liquid chromatography mass spectrometry analysis) of non-yeast secondary metabolites.[115] More elaborate approaches expand this "one strain fits all" strategy to a multispecies expression system: MORTENSEN *et al*., for instance, recently introduced the DIVERSIFY synthetic biology platform that exploits a common synthetic gene integration site for the transformation of a single gene cassette into four different *Aspergilli* species.[116]

In this study a convenient and modular vector-based multigene expression system for the quadruple auxotrophic strain *A. oryzae* NSAR1 was used.[117] This strain is deficient in the production of arginine (Δ*argB*) and adenine (Δ*adeA*$^-$) and defective in nitrate reduction (Δ*niaD*$^-$) and sulfate assimilation (Δ*sC*$^-$).[117,118] In addition to the auxotrophic markers two dominant selectable markers can be used. These are the *bar* resistance gene used with the herbicide glufosinate ammonium and the *ble* resistance gene used with phleomycin.[119,120] A series of fungal expression plasmids (pTYGS) complementing the auxotrophic/selectable markers (*argB, adeA, sC, niaD, ble, bar*) allow for reconstitution of up to 24 genes in the host:[105] Each vector comprises four cloning sites equipped with a fungal promotor/terminator pair (Scheme 1.20). One of these pairs is the malto-oligosaccaride inducible $P_{amyB}$-promotor/$T_{amyB}$ terminator pair (repressed by glucose) that flanks a rapidly accessible Gateway® destination site, ideally suited to quickly incorporate large megasynthase genes *via in vitro* recombinatorial cloning.[105] The other three pairs are strong constitutive promoters/terminators ($P/T_{adh}$, $P/T_{gpdA}$, $P/T_{eno}$) that contain an *Asc*I-restriction site directly downstream of the promoter region. *Asc*I-digestion enables the uptake of up to three tailoring genes *via in vivo* yeast homologous recombination. Tailoring genes for homologous recombination are prepared as PCR amplified fragments that contain at least 30 bp

homologous overlap with the respective promoter/terminator site. Additional features of the pTYGS plasmid series include a copy of the *ura3* gene (enabling selection in *ura3* mutant strains of yeast), a *2μ* origin of replication (plasmid replication in yeast), *colE1* origin of replication (plasmid replication in *E. coli*) as well as $cam^R$ and $amp^R$ genes encoding resistance towards ampicillin and chloramphenicol.[105]




**Scheme 1.20** Vector-based multigene expression system used in this study.

## 1.6. Overall Aims

The main focus of this thesis is on the understanding and engineering of tropolone sesquiterpenoid biosynthesis. For this, initial experiments will be directed to elucidate the complete biosynthetic pathway towards xenovulene A **1**. Heterologous expression experiments in the fungal host *A. oryzae* will be performed to identify a minimal gene set required for **1**-formation. Key biosynthetic transformations will then be reconstituted *in vitro* using recombinantly produced enzymes. A special focus will be on the identification and characterization of enzymes involved in formation of α-humulene **2** and a putative core hetero DIELS-ALDER reaction, that connects the terpene and polyketide moieties. If applicable, identified new classes of enzymes will be characterized structurally in cooperation with the Helmholtz Centre for Infection research in Braunschweig. Site-directed mutagenesis should be used to identify and validate structure-function relationships.

The biosynthetic information obtained from the xenovulene pathway will then be transferred to investigate the biosynthesis of the related TS eupenifeldin **5** and noreupenifeldin B **6** in *Phaeosphaeriaceae* sp. CF-150626. Feeding studies using stable $^{13}$C-labelled precursors will be deployed to investigate the origin of the tetrasubstituted benzene ring in **6** and whole genome sequencing should identify candidate BGC responsible for **5**- and **6**-formation.

This basic research should then pave the way to systematically and rationally engineer TS biosynthetic pathways in *A. oryzae* and to produce, identify and characterize new-to-nature TS natural products, using heterologous expression in *A. oryzae*.

# 2. Biosynthetic Studies of Xenovulene-type Tropolone Sesquiterpenoids

Part of the presented work has been published in *Nat. Commun.* **2018**, *9* and in *Angew. Chem. Int. Ed.* **2020**, *59*, 23870-23878 and in Schor, Raissa: Biosynthesis of Xenovulenes in *Acremonium strictum*. Unv. Diss. Gottfried Wilhelm Leibniz Universität Hannover, **2018**.[121–123]

## 2.1. Introduction

Xenovulene A **1** was first reported in 1995 as a metabolite isolated from submerged cultures of *Acremonium strictum* IMI 501407.[1] It was identified in the course of a pharmaceutical screening program for inhibitors of flunitrazepam **144**-binding to the GABA benzodiazepine receptor ($IC_{50}$ of **1** = 40 nM).[124] Subsequent isolation and extensive NMR analysis determined the structure of **1**. Xenovulene A **1** contains an unusual furocyclopentenone linked to α-humulene **2** *via* a bridging dihydropyran.[124] **1** was concomitantly produced with a series of related metabolites (**131**, **145**-**149**) that all share a conserved 5,6,11-tricyclic core but in which the furoyclopentenone was replaced by (highly) oxygenated benzene or tropolone rings (Figure 2.1 A).[87,123]

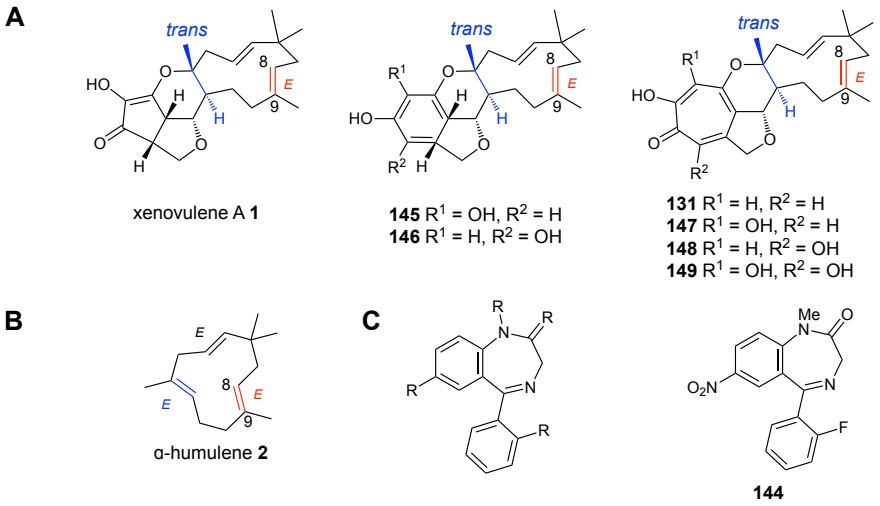


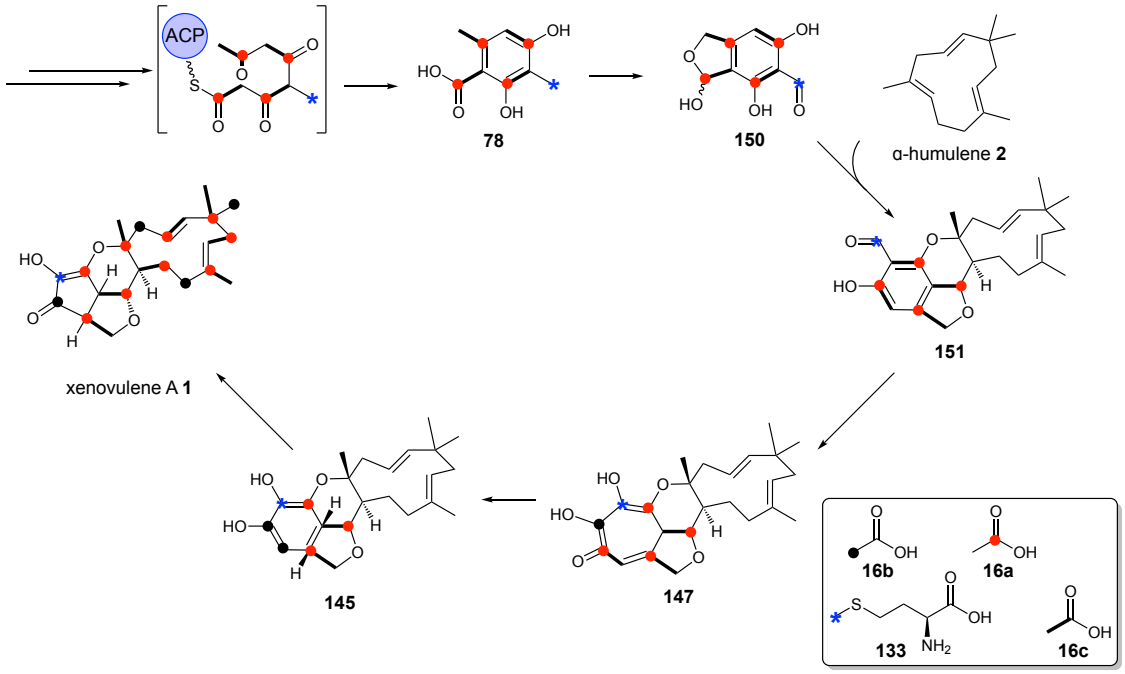

**Figure 2.1** Chemical structures of: **A,** xenovulenes isolated from submerged cultures of *Acremonium strictum*; **B,** structure of all-*trans* α-humulene **2**; **C,** 1,4-benzodiazepine core structure (left) and structure of flunitrazepam **144** (right).

Furthermore, the xenovulenes are all derived from the terpene α-humulene **2** (Figure 2.1 B). The all-*trans* configuration of **2** is mirrored in the *E*-configuration of the C-8/C-9 alkene in the xenovulenes and in the *trans*-fusion at the dihydropyran/humulene ring junction (Figure 2.1 A).

[1] The taxonomy of this strain has not yet been properly determined; preliminary investigations indicate that the strain constitutes a novel species, belonging to the *Acremonium/Sarocladium* family. Alternatively referred to as *Sarocladium schorii*.[123]

The intriguing bioactivity of **1** established the xenovulenes as interesting pharmaceutical lead structures.[125] Particularly the low structural resemblance to the classical benzodiazepines (Figure 2.1 C) is noteworthy. The latter are very potent tranquilising and anticonvulsant agents but are known to create a strong dependence in patients upon prolonged intake.[126] Less-addictive alternatives are needed, making **1** an interesting target molecule.[124]

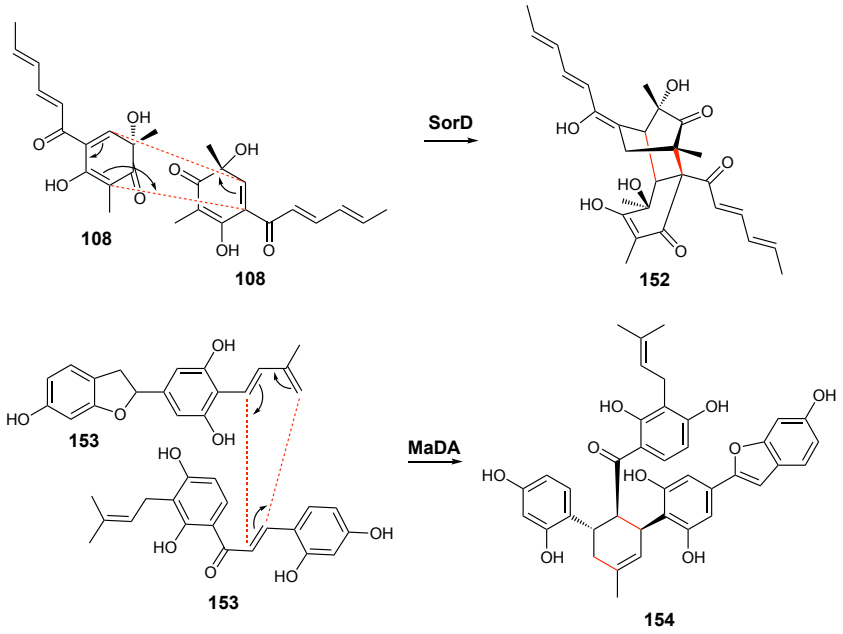
Already two years after the first reports of xenovulene A **1**, the first biosynthetic route towards **1** was proposed by SIMPSON *et al.* The proposed pathway was based on elaborate feeding studies with [1-$^{13}$C]-, [2-$^{13}$C]-, [1,2-$^{13}C_2$]-labelled acetates (**16a-c**) and [*methyl*-$^{13}$C]-methionine **133** (Scheme 2.1).[87] While the labelling of the 11-membered macrocycle corresponded to the expected pattern for the terpene α-humulene **2**, the proposed tetraketide moiety of xenovulene A **1** only contained two intact acetate units. A third acetate unit was interrupted by an inserted methyl-group, derived from [*methyl*-$^{13}$C]-methionine **133**. This latter finding was in agreement with earlier labelling experiments on tropolones by RONALD BENTLEY.[127]




**Scheme 2.1:** Proposed biosynthesis of xenovulene A **1** based on isotopic labelling by SIMPSON *et al.*[87]

Based on the co-isolation of related xenovulenes with benzene and tropolone rings, SIMPSON *et al.* proposed a complex ring-expansion/ring-contraction/ring-contraction sequence.[87] The first enzyme-free intermediate, 3-methylorsellinic acid **78**, was proposed to be transformed to the corresponding lactol **150**. Subsequently **150** would undergo an intermolecular DIELS-ALDER reaction with α-humulene **2** to give the first tropolone sesquiterpenoid **151**. Ring-expansion to **147** and two successive ring-contractions would result in the final pathway product xenovulene A **1** (Scheme 2.1).[87]

From a biochemist´s perspective the proposed pathway was intriguing. While the isotopic feeding studies confirmed the xenovulenes as meroterpenoids, the proposed pathway giving rise to **1** is very unusual for fungal meroterpenoid formation. For the majority of meroterpenoids the cyclization of the prenyl chain occurs *after* fusion with the non-terpenoid moiety (*e.g.* during the formation of anditomin **85**).[64] In xenovulene biosynthesis, contrasting this common route, the cyclization appears to occur *prior* to the fusion with the non-terpenoid part.[59] The proposed *inter*molecular hetero DIELS-ALDER reaction, that drives fusion of terpenoid and non-terpenoid moiety, is also unusual. While enzymes have been successfully demonstrated to catalyse *intra*molecular DIELS-ALDER reactions (*e.g.* IccD during ilicicolin H formation;[128] PyiF during pyrichalasin H formation;[129] MPS during macrophomate formation[130]), *inter*molecular DIELS-ALDER reactions are rarely observed in nature, and *hetero* DIELS-ALDER reactions are also rare.

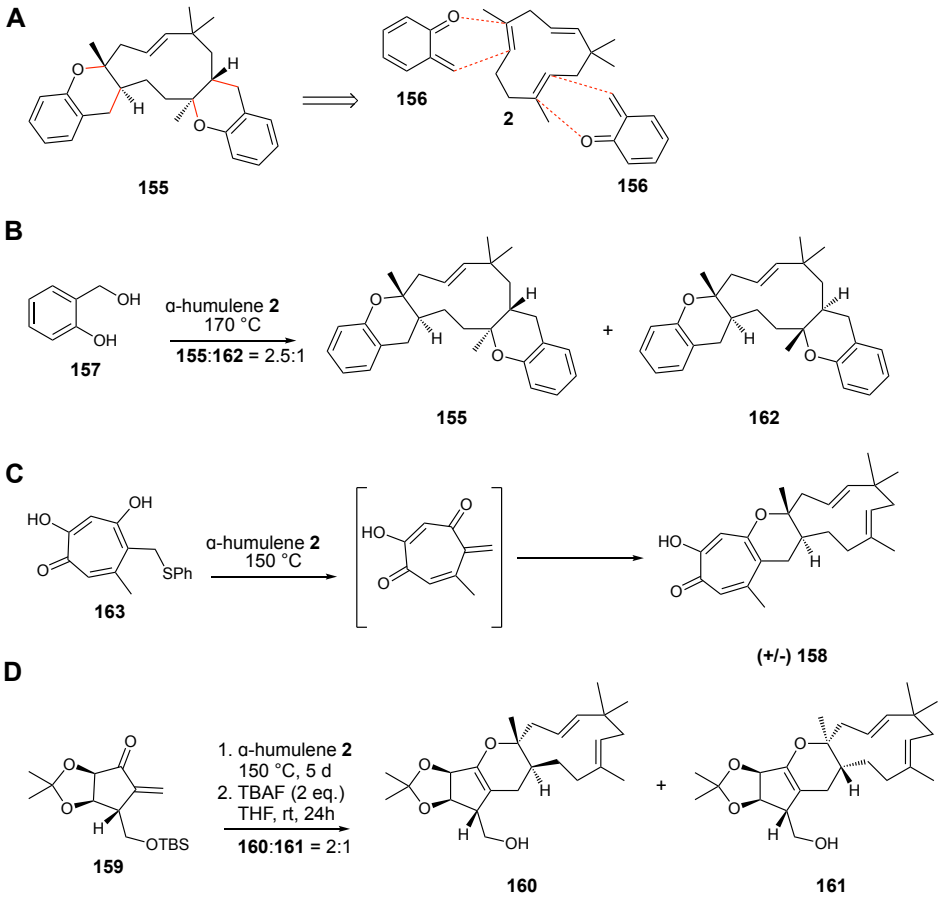



**Scheme 2.2:** Intermolecular DIELS-ALDER reaction in the biosynthesis of bisorbicilinol **152** and chalcomoracin **154**.

Enzymes catalysing *inter*molecular DIELS-ALDER reactions have only been reported for the sorbicillinoid pathway (SorD) in *Trichoderma reesei*/*Penicillium chrysogenum* and the chalcomoracin pathway in *Morus alba* (MaDA). Both examples were identified only within the last two years (Scheme 2.2).[70,131,132]

Fusion of the terpene and non-terpenoid moiety through an *inter*molecular DIELS-ALDER reaction in tropolone sesquiterpenoid formation was first proposed by ZWANENBURG *et al.* for the plant natural product lucidene **155.**[133] Lucidene was isolated from the rootbark of *Uvaria lucida* spp. *lucida* and the authors proposed **155** to be derived from α-humulene **2** and two molecules of *o*-benzoquino methide **156** (Scheme 2.3 A).[133] Subsequent biomimetic synthetic studies by

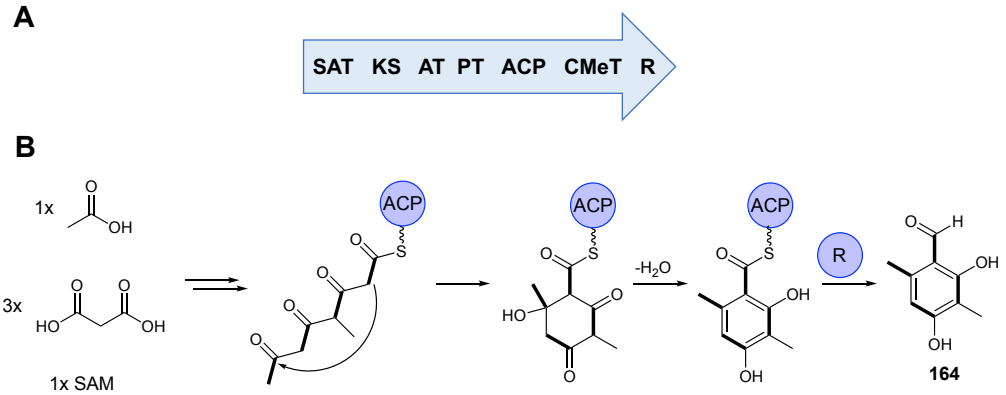
BALDWIN *et al.* supported this hypothesis. Thermolysis of the hydroxybenzyl alcohol **157** at 170 °C afforded the corresponding *o*-quinomethide **156** and *in situ* reaction with α-humulene **2** afforded lucidene **155** as the major product (Scheme 2.3 B).[134] Similar chemistry was successfully deployed to synthesize a deoxy analogue **158** of epolone B (Scheme 2.3 C).[135]




**Scheme 2.3** Proposed hetero DIELS-ALDER reaction in the biosynthesis of tropolone sesquiterpenoids: **A,** proposed biosynthesis of lucidene **155**; **B,** biomimetic synthesis of lucidene **155** by BALDWIN *et al.*; **C,** biomimetic synthesis of a deoxy analogue **158** of epolone B by BALDWIN *et al.*; **D,** biomimetic synthesis of the xenovulene core by KIRSCHNING *et al.*[133–136]

Most recently, KIRSCHNING *et al.* applied the same principle in the total synthesis of the xenovulene core scaffold. Hetero DIELS-ALDER reaction between humulene **2** and methyl vinyl ketone **159** at 150 °C for 5 d afforded the expected cycloadducts **160** and **161** (Scheme 2.3 D).[136] As the biomimetic chemical reaction only occurred at elevated temperatures (>> 100 °C), nature is likely to deploy enzyme catalysis to achieve this transformation. However, despite significant advances in synthetic chemistry a universal access to all members of this class of natural products remains to be found. Biosynthesis might thus be a viable alternative approach to assess different TS natural products.

Investigations into the biosynthesis of the xenovulenes now span a time period of over 20 years and showcase the extensive advancements in the field of natural product biosynthesis. Evolving new technologies and methodologies spurred the understanding of xenovulene formation in *A. strictum*. Ten years after the first biosynthetic proposal a partial xenovulene A biosynthetic gene cluster was identified. The BGC was clustered around a non-reducing polyketide synthase gene (*asPKS1*).[137] Heterologous expression of *asPKS1* in the auxotrophic fungal host *A. oryzae* NSAR1 identified 3-methylorcinaldehyde **164** (instead of the acid **78**) as the sole product of AsPKS1. The isolation of **164** agreed with the identification of a reductive release domain within the *asPKS1* gene sequence (Scheme 2.4).[137] Intriguingly, 3-methylorcinaldehyde **164** also represented a known precursor of the tropolone stipitatic acid **165** and thus linked tropolone and xenovulene biosynthesis for the first time.[138]

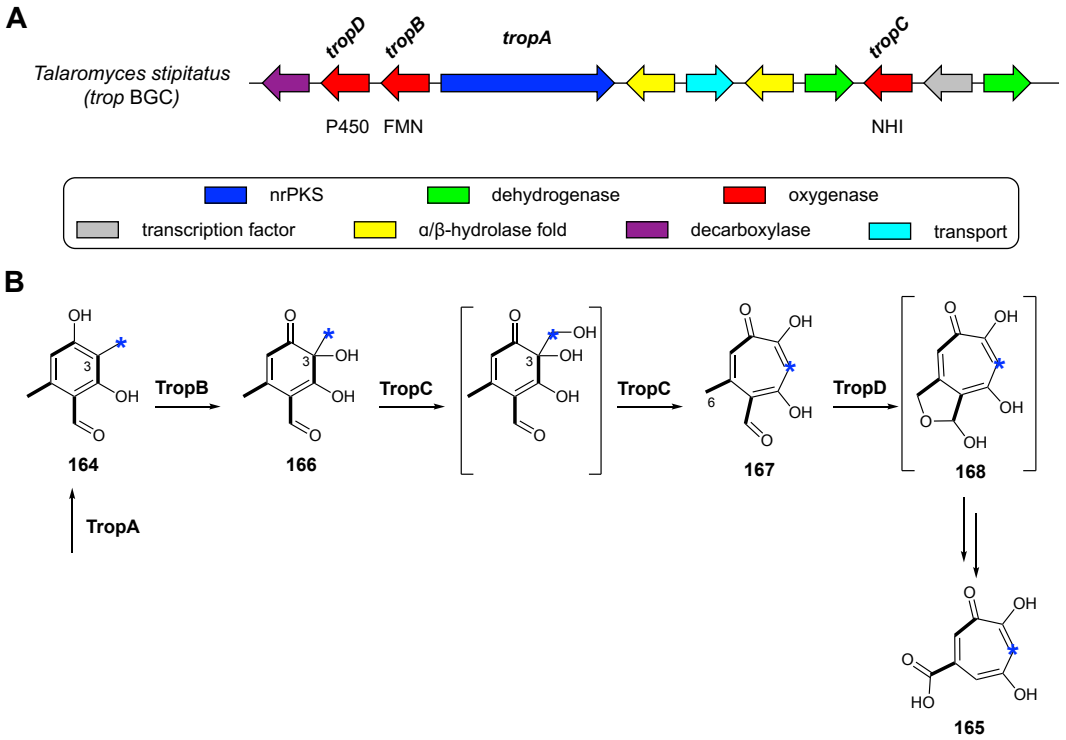



**Scheme 2.4** Analysis of AsPKS1 from *Acremonium strictum*: **A,** domain architecture of AsPKS1; SAT = starter unit acyltransferase; KS = ketosynthase; AT = acyltransferase; PT = product template domain; ACP = acyl carrier protein; *C*-MeT = *C*-methyltransferase; R = reductive release domain; **B,** proposed biosynthesis of 3-methylorcinaldehyde **164**; incorporated acetate units highlighted in bold.

Natural products containing the tropolone motif are widely found in plants, bacteria and fungi.[139] Plants deploy terpene/alkaloid pathways to assess the tropolone core structure and bacteria either use the shikimate-pathway or type II-PKS.[139] 3-methylorcinaldehyde **164** being the precursor of fungal tropolones suggested that fungi use a third, non-reducing PKS-based, route towards tropolones.[139] Fungal tropolone biosynthesis was subsequently unveiled in molecular detail during studies of the formation of stipitatic acid **165** in *Talaromyces stipitatus*.[140] Through an elaborate combination of gene knockouts and heterologous expression, three core genes were identified which form the tropolone nucleus (Scheme 2.5).

TropA, a non-reducing polyketide synthase (38.7 % sequence identity with AsPKS1 on amino acid sequence level) was demonstrated to catalyse formation of 3-methylorcinaldehyde **164**. An FAD-dependent monooxygenase, TropB, then oxidatively dearomatizes **164** to afford the 3-hydroxylated enone **166**. TropC, a non-heme Fe(II)-dependent oxygenase, catalyses a key oxidative ring-expansion *via* hydroxylation of the C-3 methyl group and subsequent radical

rearrangement to yield stipitaldehyde **167**.[140,141] Additional data obtained from knockout experiments proposed the cytochrome P450, TropD, to act on **167** subsequently. Most likely TropD oxidizes at the C-6 methyl-group, followed by the immediate formation of the hemiacetal **168**. However, **168** was not directly observed.[140]




**Scheme 2.5** Proposed biosynthesis of stipitatic acid **165**: **A,** schematic representation of the *trop* BGC from *T. stipitatus*; **B,** proposed pathway towards stipitatic acid **165**; compounds in brackets not directly observed; P450 = cytochrome P450; FMN = FAD-dependent monooxygenase; NHI = non-heme iron dioxygenase; labelling pattern adopted from SIMPSON *et al.*[87]

In 2013 RAISSA SCHOR (COX research group, Hannover) performed whole genome sequencing of the xenovulene A **1**-producing strain *Acremonium strictum* and identified a ~49 kb BGC centered around the previously described core *asPKS1* gene (Figure 2.2).[123]

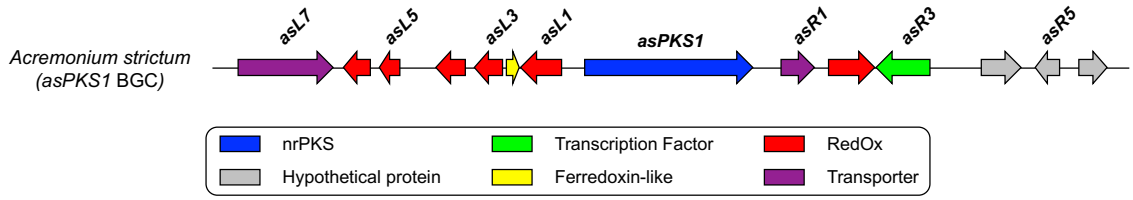



**Figure 2.2:** Schematic overview of the *asPKS1* biosynthetic gene cluster; genes are scaled in respect to actual size and distance from each other.

Since the early biosynthetic steps of xenovulene A **1** and stipitatic acid **165** biosynthesis were assumed to be identical, a global sequence alignment was performed between the *trop* and *asPKS1* biosynthetic gene clusters. The reanalyzed complete *asPKS1* BGC indeed harbored a

complete set of genes (*asPKS1, asL1, asL3, asR2*) required for tropolone formation, homologous to *tropA-D* in *T. stipitatus* (Figure 2.3).[123] Transcriptomic analysis under producing and non-producing conditions linked the cluster to the production of **1** and established the cluster boundaries.[123] Genetic inactivation of the central non-reducing PKS gene, *asPKS1*, abolished xenovulene A **1** biosynthesis and confirmed the biosynthetic gene cluster.[123] Notably, the cluster did not comprise an obvious class I terpene cyclase (required for α-humulene **2** formation) nor a gene showing similarity to any known DIELS-ALDERase. This suggested the presence of unprecedented enzymatic potential within the *asPKS1* BGC. Indeed, with AsR4, AsR5 and AsR6 the cluster harbored three enzymes that bioinformatics analysis using blastp[142] identified as hypothetical proteins, not showing any similarity to enzymes of known function.[123]

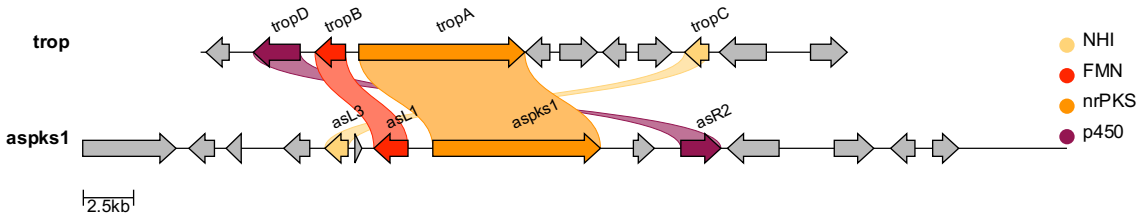
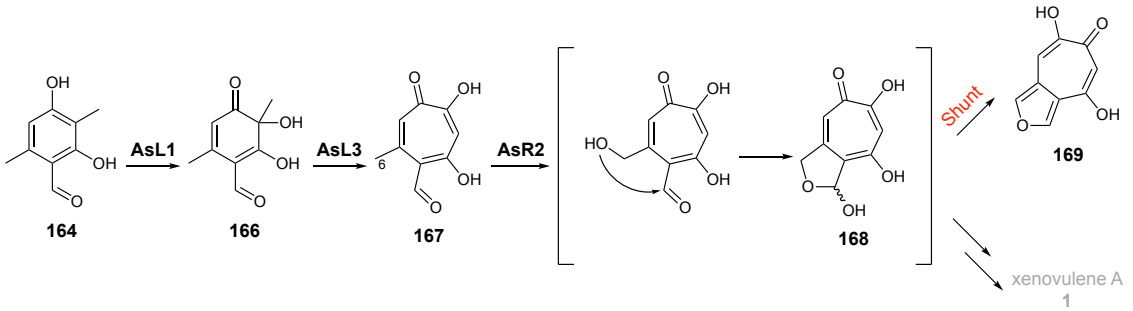



**Figure 2.3:** Global sequence alignment of the *trop* and *asPKS1* BGC; NHI = non-heme iron dioxygenase; FMN = FAD-dependent monooxygenase; nrPKS = non-reducing PKS; P450 = cytochrome P450; alignment and visualisation performed with Clinker & Clustalmaps.js.[143]

With genetic engineering of *Acremonium strictum* being a complicated endeavour (only one genetic knockout experiment [*asPKS1*] was successful),[123] RAISSA SCHOR investigated the biosynthesis of **1** through a series of heterologous expression experiments in the quadruple auxotrophic strain of *A. oryzae* NSAR1.[117] Stepwise inclusion of the *tropA-D* homologues *asPKS1, asL1, asL3* and *asR2* in the modular expression system confirmed that the early biosynthetic steps were identical as previously observed for tropolone formation in *T. stipitatus* (Scheme 2.6).[123] AsPKS1 generates the non-reduced polyketide 3-methylorcinaldehyde **164** as the first enzyme-free intermediate. The aldehyde **164** is oxidatively dearomatized through the FAD-dependent monooxygenase AsL1, yielding **166**. Enone **166** then undergoes oxidative ring-expansion to give the key intermediate stipitaldehyde **167**.[123]




**Scheme 2.6:** Early biosynthetic steps in xenovulene A **1** formation as identified through heterologous expression in *A. oryzae*.[123]

As for the formation of stipitatic acid **165** the cytochrome P450 AsR2 probably catalyzes formation of the hemiacetal **168** through initial hydroxylation at the C-6 methyl group. However, instead of the hemiacetal **168**, stipitafuran **169** was isolated, probably as the result of shunting in *A. oryzae*.[123]

Further heterologous expression of the `full´ *asPKS1* biosynthetic gene cluster (11 genes, solely omitting transporters and transcription factors) in *Aspergillus oryzae* NSAR1 was sufficient to establish production of **1**.[123] Stepwise omission of individual genes from the `full´ cluster (`knockout by expression´), thus expressing different combinations of 10 genes, allowed confirmation of various genes (*asL2, asL5, asR4*) as non-essential: Their omission did not affect **1** production (Table 2.1).[123]

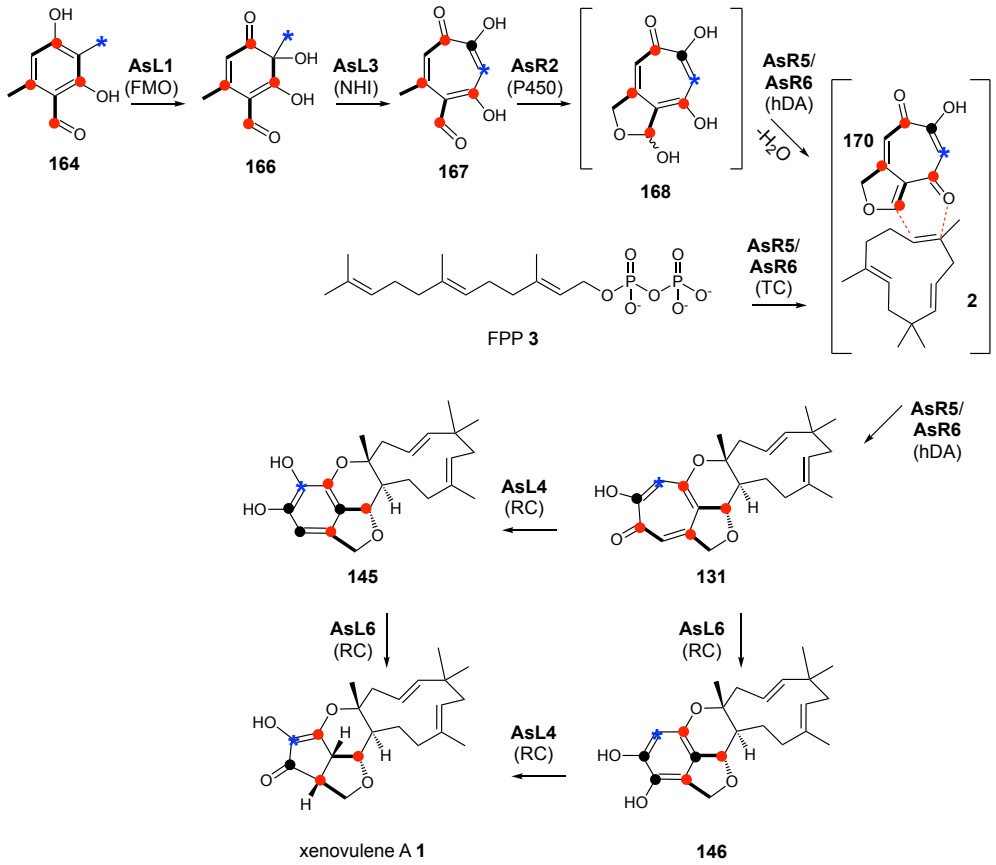
**Table 2.1** `Knockout by expression´ strategy perused by RAISSA SCHOR; all transformed gene sets constitute an individual expression experiment in *A. oryzae* NSAR1;[123] red. = reduced levels of xenovulene A **1** production.

| Exp. ID | Transformed genes | | | | | | | | | | | Prod. of |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***L6*** | ***L5*** | ***L4*** | ***L3*** | ***L2*** | ***L1*** | ***PKS*** | ***R2*** | ***R4*** | ***R5*** | ***R6*** | **1** |
| `full´ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Δ*asL2* | ✓ | ✓ | ✓ | ✓ | x | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Δ*asL4* | ✓ | ✓ | x | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | red. **1** |
| Δ*asL5* | ✓ | x | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Δ*asL6* | x | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | red. **1** |
| Δ*asR4* | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | x | ✓ | ✓ | ✓ |
| Δ*asR5* | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | x | ✓ | x |
| Δ*asR6* | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | x | x |
| Δ*asL4+asL5+asL6* | x | x | x | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | x |

Notably, omission of either *asR5* or *asR6* abolished production of xenovulene A **1**, albeit the early tropolone intermediates were still produced.[123] This suggested these two enzymes to be involved in α-humulene **2** formation and subsequent hetero DIELS-ALDER reaction. Omission of either one of the FAD-dependent oxidoreductases AsL4 or AsL6 drastically reduced the titers of **1** and gave rise to two new compounds, **145** and **146** (Scheme 2.7).[123] The two metabolites contained a benzene ring instead of a tropolone or furocyclopentenone, implying AsL4/AsL6 to be involved in the two ring-contractions occurring during xenovulene A **1** biosynthesis. As both experiments yielded **1**, albeit in lower quantities, AsL4 and AsL6 seem to have partial capability to replace each other.[123] Intriguingly, however, omission of *asL4*, *asL5* and *asL6* in Exp. Δ*asL4+asL5+asL6* (Table 2.1) did not afford the intact tropolone nucleus containing xenovulene B **131**, as would have been expected upon omission of the ring-contracting enzymes.[123]

Combined, these results indicated a possible route towards **1** that was in agreement with the labelling pattern observed by SIMPSON *et al.*, the xenovulenes isolated from wild-type *Acremonium strictum* and the results of the heterologous expression experiments (Scheme 2.7).

Accordingly, the well-known tropolone pathway would generate the tropolone nucleus of **167**. The AsR2-derived hemiacetal **168**, despite being unobserved, is the likely substrate of the hetero DIELS-ALDERase (either AsR5 or AsR6): Dehydration would yield the reactive *o*-quinomethide **170** that could undergo the intermolecular hetero DIELS-ALDER reaction with α-humulene **2**, generated by either AsR5 or AsR6 from FPP **3**, to form the first meroterpenoid, xenovulene B **131**. Subsequently, **131** undergoes two successive ring-contractions that can follow two different routes, depending on the presence/absence of AsL4/AsL6.[123]




**Scheme 2.7** Proposed biosynthesis of xenovulene A **1** based on the `knock-out by expression´ strategy perused by RAISSA SCHOR;[123] FMO = FAD-dependent monooxygenase; NHI = non-heme iron dioxygenase; P450 = cytochrome P450; hDA = hetero DIELS-ALDERase; TC = terpene cyclase; RC = ring-contraction enzyme.

## 2.2. Project Aims

Despite the successful reconstitution of the xenovulene A **1** biosynthesis in *A. oryzae* NSAR1, the key biosynthetic steps (hetero DIELS-ALDER reaction; α-humulene **2** formation) await further elucidation. While the results indicate that *at least* eight genes are required for the production of **1** (*asPKS1, asL1, asL3, asL4, asL6, asR2, asR5, asR6*), this awaits further verification. In initial experiments this hypothesis should be probed through heterologous expression of the respective eight core genes alone. If production of **1** is observed, a further reduction of the expressed set of genes should be attempted. A possible candidate for omission from the set of expressed genes is

the cytochrome P450 encoding gene *asR2*. Neither for tropolone production nor for xenovulene A **1** production could its exact role be experimentally verified. Additionally, omission of the putative ring-contracting enzyme encoding genes *asL4* and *asL6* should be attempted. It is surprising that in previous work by RAISSA SCHOR omission of *asL4*+*asL6*+*asL5* (Table 2.1) did not afford xenovulene B **131**. This may point to the ring-contracting enzymes acting prior to the occuring DIELS-ALDER reaction and this point should be investigated in detail.

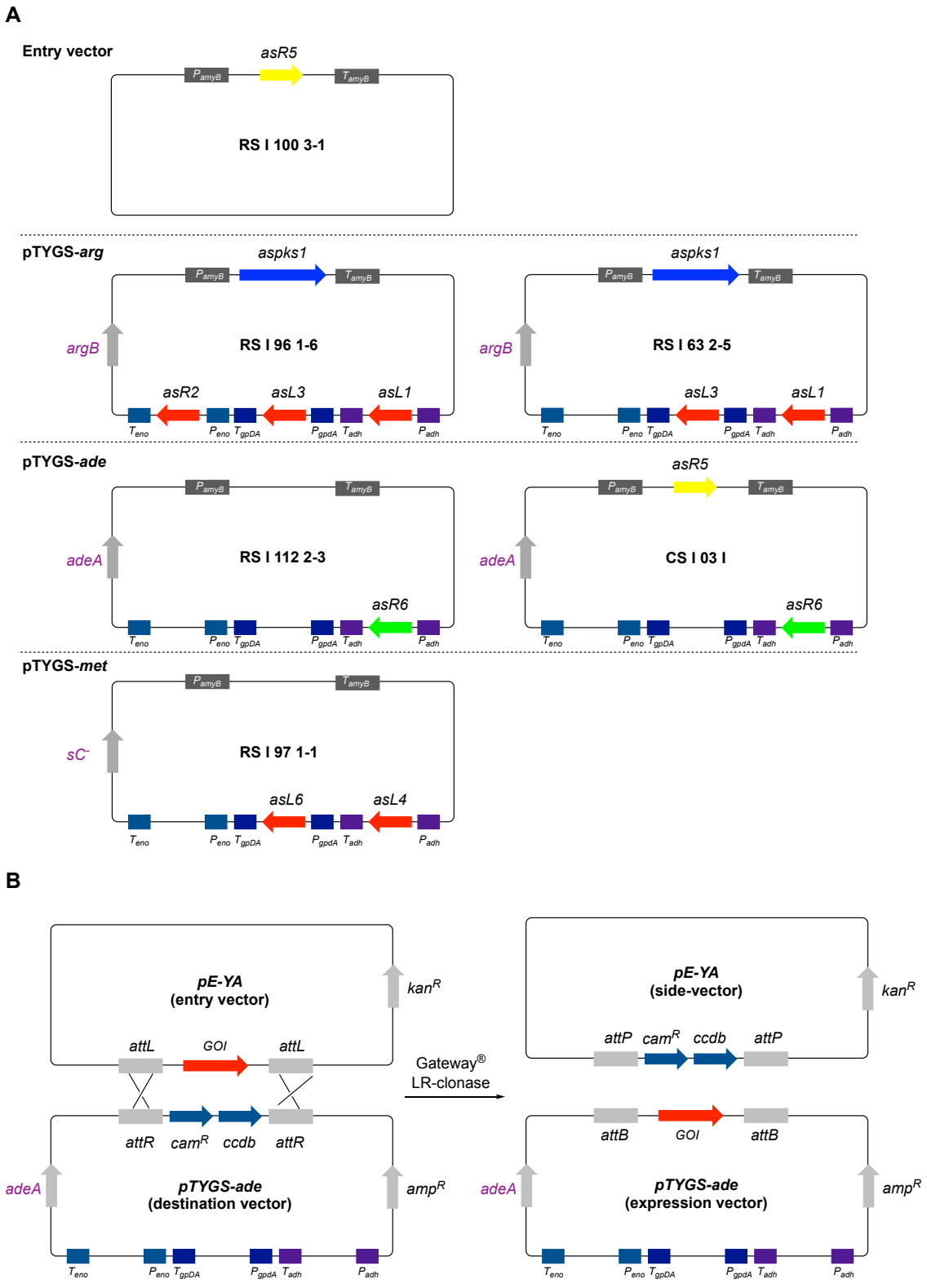
Most intriguingly, *in silico* analysis of the *asPKS1* BGC does not identify any candidate gene within the cluster that resembles a known DIELS-ALDERase or class I terpene cyclase. Initial results by RAISSA SCHOR suggest that the hypothetical enzymes AsR5 and AsR6 may fulfil these functions. Accordingly, both may constitute the first members of novel families of DIELS-ALDERases and cryptic terpene cyclases, respectively. Identification and analysis of these enzymes should contribute fundamentally to our understanding of tropolone sesquiterpenoid biosynthesis. Suitable *in vitro* assays should be established to investigate the function of enzymes through production of recombinant proteins in *E. coli* and incubation with suitable substrates.

## 2.3. Results – Heterologous Expression of Genes from the *asPKS1* BGC

### 2.3.1. Construction of Expression Plasmids

To investigate the biosynthesis of xenovulenes through heterologous expression, initially suitable expression plasmids were either provided by RAISSA SCHOR or generated in this study (Scheme 2.8 A).

For the heterologous expression of genes from the *asPKS1* BGC in *A. oryzae* NSAR1, the modular expression system established by LAZARUS and co-workers was used.[103] Biosynthetic genes were cloned into one of the three expression plasmids *pTYGS-arg/ade/met*, each complementing one of the auxotrophies of *A. oryzae* NSAR1. A new *pTYGS-ade* vector carrying the gene cassettes for the two hypothetical proteins *asR5* and *asR6* was generated in this study using Gateway® technology.




**Scheme 2.8: A,** overview of expression plasmids used in this study; plasmids named with the prefix `RS´ were kindly provided by RAISSA SCHOR;[123] **B,** Principle of Gateway® cloning technology; *amp*$^R$ = carbenicillin resistance gene; *kan*$^R$ = kanamycin resistance gene; *GOI* = gene of interest; P = promoter T = terminator.[144]

Gateway® cloning (also referred to as recombinatorial cloning)[108] is routinely used to shuffle intact DNA sequences from an entry vector (here: *pE-YA* with the gene of interest [*goi*] flanked by *attL* recombination sites) into a destination vector (here: *pTYGS-ade* with a *ccdB*+*cam*$^R$ gene

cassette, flanked by *attR* recombination sites; the *ccdb*+*cam*$^R$ gene cassette encodes for the CcdB killer protein and chloramphenicol resistance). An enzyme mixture comprising the bacteriophage λ integrase, excisionase and *E. coli* integration host factor facilitates recombination of the *attL*- and *attR*-flanked sequence *in vitro*, resulting in the formation of two new vectors: The expression vector now comprising the gene of interest flanked by *attB*-sites and a side-vector comprising the *ccdb*+*cam*$^R$ gene cassette flanked by *attP*-sites.[144] Transformation into *E. coli* TOP10 supplemented with carbenicillin$^{50}$ allows selection for the desired expression vector as the expression vector and side-vector carry different resistance genes (*carb* vs. *kan*). Additionally, the side-vector now carries the *ccdb* gene which encodes the CcdB killer protein that is highly toxic to *E. coli* TOP10 (Scheme 2.8 B).[144] The obtained plasmid was confirmed by DNA sequencing (*Eurofins*, Ebersberg) for the correct integration of the target gene.

### 2.3.2. Transformation of *A. oryzae* NSAR1

*A. oryzae* NSAR1 was transformed with up to three plasmids carrying a maximum of eight biosynthetic genes using a well-established $CaCl_2$/PEG-mediated protocol.[103] Obtained transformants (typically ≥ 5) were selected on appropriate agar, depending on the presence of auxotrophic markers on the transformed *pTYGS*-plasmids. For example, transformants obtained from the transformation of *pTYGS-arg*, *pTYGS-ade* and *pTYGS-met* were selected on medium deficient in arginine, adenine and methionine as each vector complements one of the three auxotrophic markers (note: methionine complements sulfate reductase auxotrophy). Three different gene combinations were transformed into *A. oryzae* NSAR1 (Table 2.2).

**Table 2.2** Overview of transformation experiments performed in this study; each Exp. ID represents an individual expression experiment in *A. oryzae* NSAR1.

| Exp. ID | Transformed plasmids with pTYGS-arg/ade/met backbone | Transformed genes | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | ***PKS*** | ***L1*** | ***L3*** | ***L4*** | ***L6*** | ***R2*** | ***R5*** | ***R6*** |
| #1 | RSI96 1-6, CSI03 I, RSI97 1-1 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| #2 | RSI62 2-5, CSI03 I, RSI97 1-1 | ✓ | ✓ | ✓ | ✓ | ✓ | x | ✓ | ✓ |
| #3 | RSI96 1-6, CSI03 I, *pTYGS-met* | ✓ | ✓ | ✓ | x | x | ✓ | ✓ | ✓ |

After three rounds of selection gDNA of all transformants was extracted using a commercially available kit (*Sigma Aldrich*). Gene-specific primers for each transformed biosynthetic gene were used in a standard PCR reaction to confirm that the integration of the corresponding gene cassette had indeed occurred (Figure 2.4). Transformants, which gDNA tested positive for the presence of all transformed genes, were cultivated in DPY medium at 28 °C and 110 rpm for 5 d prior to chemical analysis.

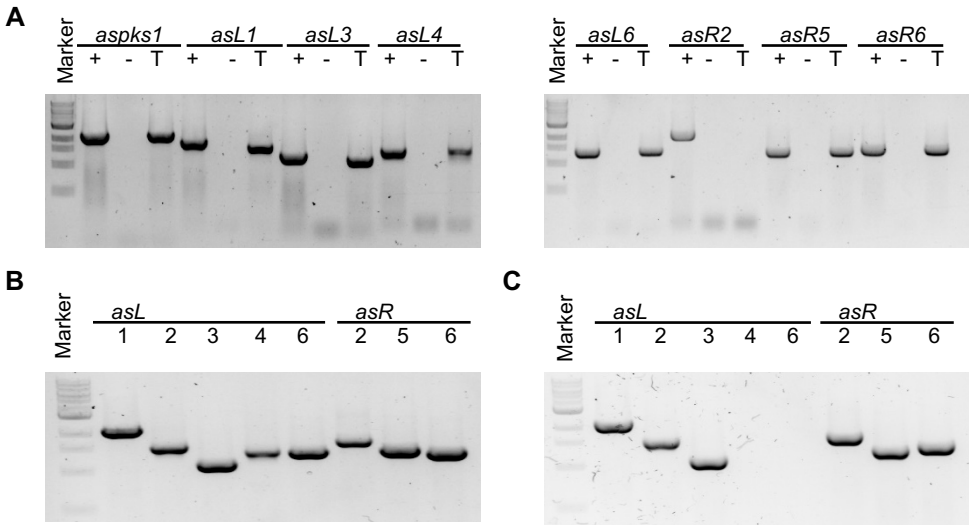



**Figure 2.4** PCR analysis of representative *A. oryzae* NSAR1 transformants: **A,** Exp. ID #2. + = positive control with *A. strictum* gDNA; - = negative control with water; T = gDNA of transformant; **B,** Exp. ID #1; **C,** Exp. ID #3. Oligonucleotides used: *asPKS1* (83+322); *asL1* (421+423); *asL3* (424+426); *asL4* (645+706); *asL6* (650+709); *asR2* (427+428); *asR5* (761+762); *asR6* (759+760).

### 2.3.3. #Exp.1 - The Minimal *asPKS1* BGC

Transformation of *A. oryzae* NSAR1 with eight genes (*asPKS1* + *asL1* + *asL3* + *asL4* + *asL6* + *asR2* + *asR5* + *asR6*; Exp. #1, Table 2.2) yielded six positive transformants carrying all eight biosynthetic genes. When cultivated under inducing conditions the fungal cultures turned yellow after about 2 d, indicating that the early tropolone producing steps were active in the heterologous host (Figure 2.5).

*A. oryzae* NSAR1 WT

*A. oryzae* Exp. #1
+*aspks1*+*asL1*+*asL3*+*asR2*
+*asR5*+*asR6*+*asL4*+*asL6*

**Figure 2.5:** Comparison of fungal cultures of *A. oryzae* NSAR1 WT (left) and *A. oryzae* NSAR1 transformed with eight genes (ID #Exp. 1; right); cultures were cultivated in DPY medium, 28 °C and 110 rpm for 4d.

After 5 d the crude EtOAc extract (dissolved in acetonitrile:water [9:1] to a final concentration of 10 mg·mL$^{-1}$) of each transformant was analysed by LCMS (Figure 2.6).

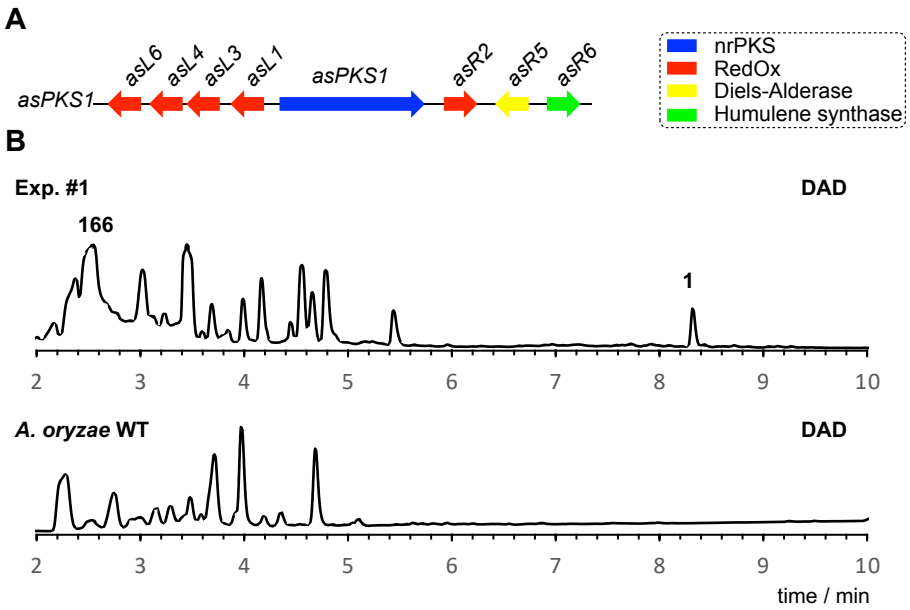



**Figure 2.6** Overview heterologous expression of the minimal *asPKS1* BGC: **A,** co-expression of eight genes from the *asPKS1* BGC; **B,** LCMS DAD chromatograms of organic extracts of *A. oryzae* expression strain and wild-type control (arbitrary units).

Upon comparison with a control strain (*A. oryzae* NSAR1 transformed with "empty" vectors *pTYGS-arg/ade/met*) a new metabolite, eluting at $t_R$ = 8.3 min, and corresponding to the nominal mass of xenovulene A **1** (358) was identified only in the representative transformants of Exp. #1. The peak had a characteristic UV absorption of $\lambda = 276$ nm that corresponded to data previously reported for **1**, and, together with the retention time being identical as observed previously, this was sufficient to confirm production of **1** (Figure 2.7).[123]

Additionally, the transformants produced a series of polar compounds eluting at $t_R$ = 2-5 min. These compounds were only assessed by low-resolution mass spectrometry and UV absorption spectra and were partially identified as intermediates/shunts of the early tropolone producing steps. For example, the pre-tropolone intermediate generated by the FAD-dependent oxygenase AsL1, enone **166**, was identified based on retention time ($t_R$ = 2.5 min), low-resolution MS (*m/z* $ES^+$ = 183) and UV absorption ($\lambda = 301$ nm; Figure 2.7).[123] However, as these polar compounds were already characterized in detail by RAISSA SCHOR they were not further analysed/annotated in the scope of this thesis.

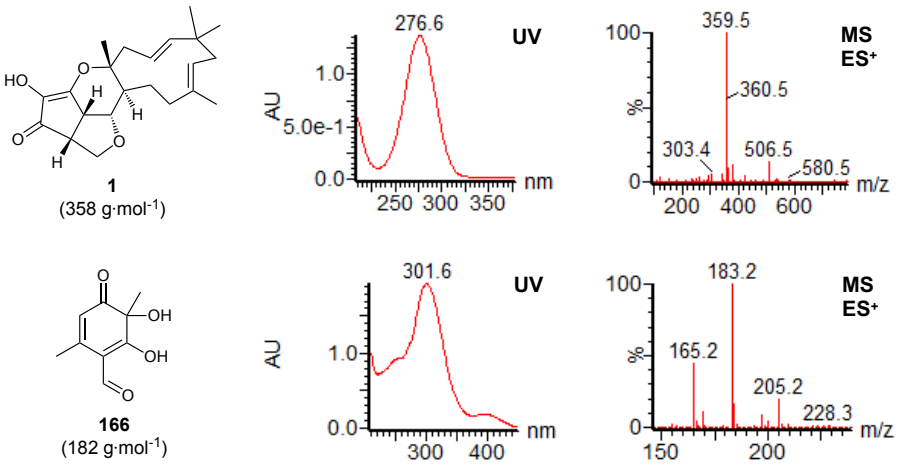
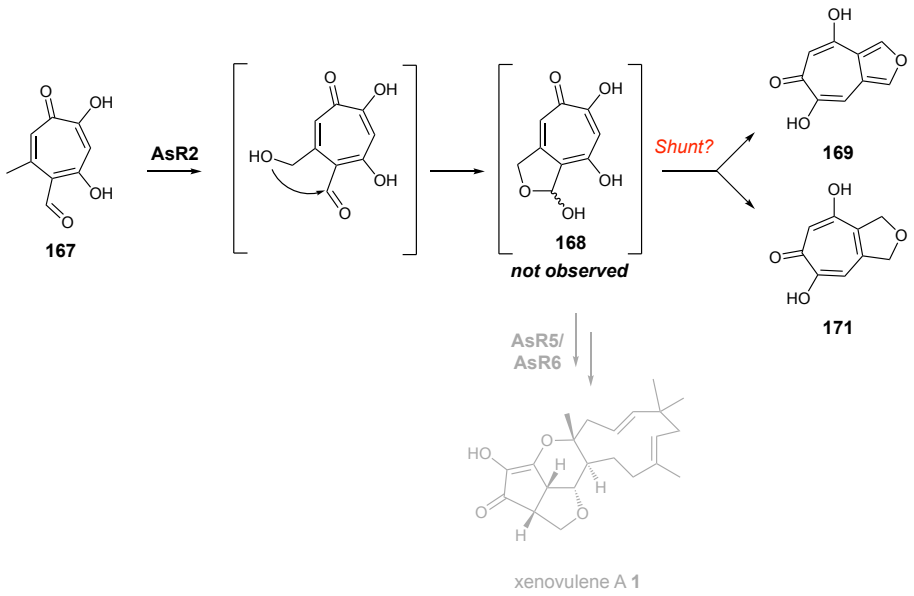



**Figure 2.7:** Chemical characterisation of xenovulene A **1** ($t_R$ = 8.3 min) by UV and mass spectrum obtained from $ES^+$ and chemical characterisation of **166** ($t_R$ = 2.5 min) by UV and mass spectrum obtained from $ES^+$.

Production of **1** through heterologous expression of eight biosynthetic genes in *A. oryzae* supports previous results obtained through heterologous expression by RAISSA SCHOR, who proposed these genes to be essential for **1** generation.[123] However, it is also possible that fewer than the eight expressed genes are required for **1**-formation. The role of *asPKS1*, *asL1* and *asL3* in tropolone formation is well-understood and therefore these must be essential genes. The same accounts for the hypothetical protein encoding genes *asR5* and *asR6*, omission of which had previously been proven to be detrimental to **1**-production.[123] More careful consideration has to be paid to the cytochrome P450 encoding gene *asR2*, and the FAD-dependent monooxygenase encoding genes *asL4* and *asL6*, whose role is not completely understood. These genes will be omitted from the expression system in follow-up experiments.

### 2.3.4. #Exp.2 - Omission of Cytochrome P450 Encoding Gene *asR2*

Based on the genetic inactivation of the cytochrome P450 encoding gene *tropD* (homologous to *asR2*) in the fungus *T. stipitatus*, AsR2 was proposed to act after the core proteins required for fungal stipitaldehyde **167** formation (AsPKS1, AsL1, AsL3).[140] However, co-expression of *asPKS1*, *asL1*, *asL3* and *asR2* in *A. oryzae* did not lead to production of the expected hemiacetal **168** but instead to the production of stipitafuran **169** and cordytropolone **171** (Scheme 2.9).[123] While these are likely to be (more stable) shunts in *A. oryzae*, the exact function of *asR2* in xenovulene biosynthesis needs to be validated.




**Scheme 2.9:** Proposed biosynthetic role of AsR2 in the biosynthesis of xenovulene A **1**.

Transformation of *A. oryzae* NSAR1 with seven genes (*asPKS1* + *asL2* + *asL3* + *asL4* + *asL6* + *asR5* + *asR6*; Exp. ID #2, Table 2.2), omitting *asR2*, resulted in three transformants showing intact integration of all transformed genes. Analysis of the crude extracts of these transformants by LCMS showed that the early tropolone producing steps were intact, as confirmed by observation of the previously reported set of tropolone intermediates and shunts (*e.g.*, 3-methylorcinaldehyde **164** eluting at $t_R$ = 6.0 min; enone **166** eluting at $t_R$ = 2.6 min; Figure 2.8). Trace amounts of xenovulene A **1**, eluting at the same retention time ($t_R$ = 8.3 min) as previously, were identified only in extracted ion chromatograms (EIC) when searching for *m/z* = 359 in $ES^+$ mode (Figure 2.8 C). Further evidence for trace production of **1** was obtained from HRMS ($[M]H^+$ calculated $C_{22}H_{31}O_4$ 359.2222, found 359.2223), which confirmed the expected molecular formula of $C_{22}H_{30}O_4$ for xenovulene A **1**.

Despite trace production of **1**, the cytochrome P450 AsR2 proves vital in the formation of **1**. Residual traces of **1** are likely the effect of housekeeping cytochromes P450 from *A. oryzae* NSAR1 that are partially capable of complementing AsR2 function (the genome of *A. oryzae* houses more than > 150 annotated cytochrome P450s).[145]

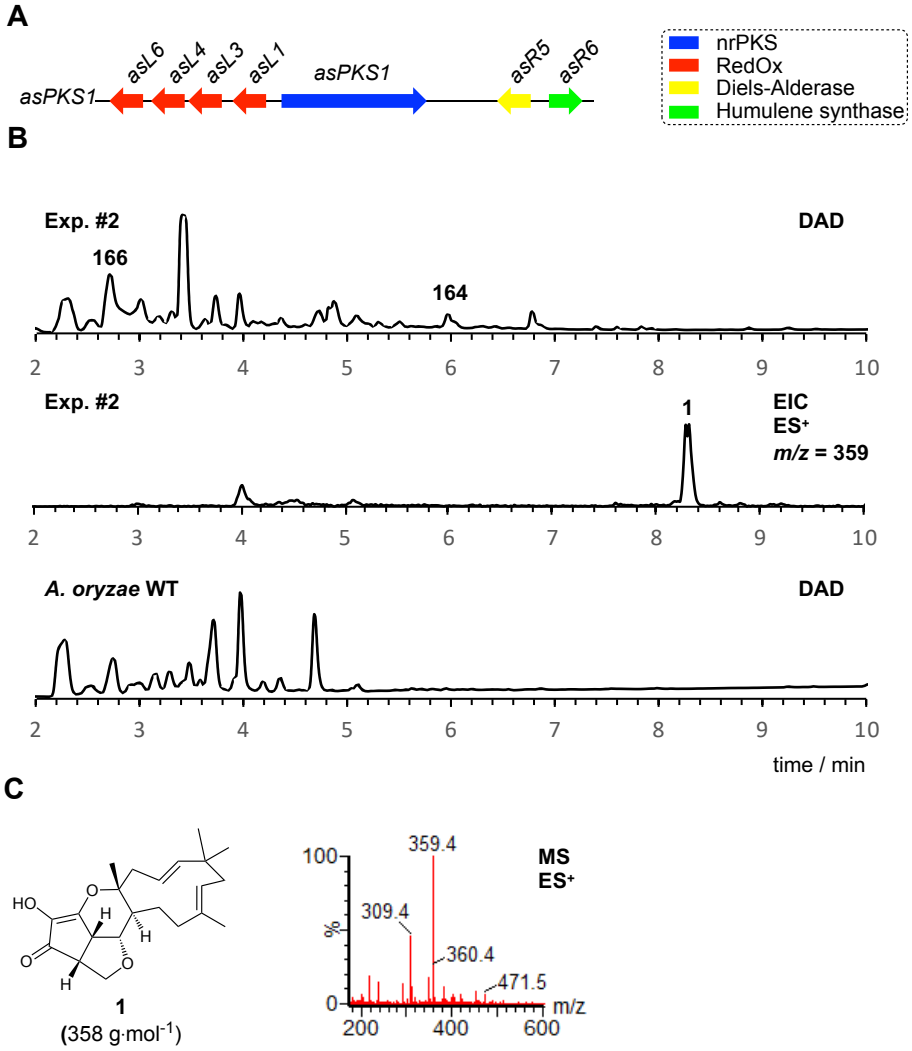



**Figure 2.8** Overview heterologous expression of the minimal *asPKS1* BGC omitting *asR2*: **A,** co-expression of seven genes from the *asPKS1* BGC; **B,** LCMS DAD/EIC chromatograms of organic extracts of *A. oryzae* expression strain and wild-type control (arbitrary units); **C,** chemical characterisation of xenovulene A **1** by mass spectrum obtained from $ES^+$.

### 2.3.5. #Exp. 3 - Omission of the FMO Encoding Genes *asL4* and *asL6*

Previous experiments by RAISSA SCHOR were directed to analyse the role of the two FAD-dependent monooxygenases AsL4 and AsL6. Omission of either one of them resulted in drastically reduced titres for xenovulene A **1** and gave rise to two new xenovulenes (**145** and **146**), harbouring benzofuran rings instead of the furocyclopentenone observed in xenovulene A **1**.[123] The latter finding implied both proteins to be involved in the two ring-contractions occurring during the xenovulene biosynthetic pathway. However, while omission of either *asL4* or *asL6* afforded novel phenolic xenovulenes, the proposed tropolone nucleus-containing precursor xenovulene B **131** was not observed. Furthermore, in a different experiment, omission of *asL4*, *asL6* and the non-essential gene *asL5* also did not afford the expected xenovulene B **131**.[123] Absence of **131**, however, suggests that the ring-contractions might occur prior to the proposed

hetero DIELS-ALDER reaction. Alternatively, due to technical difficulties at the time when RAISSA SCHOR performed her analysis, xenovulene B **131** might not have been detected.

To confirm that both genes are essential for the production of xenovulene A **1** and act after formation of xenovulene B **131** *in vivo*, in this study both genes were omitted from heterologous expression of the minimal *asPKS1* BGC.

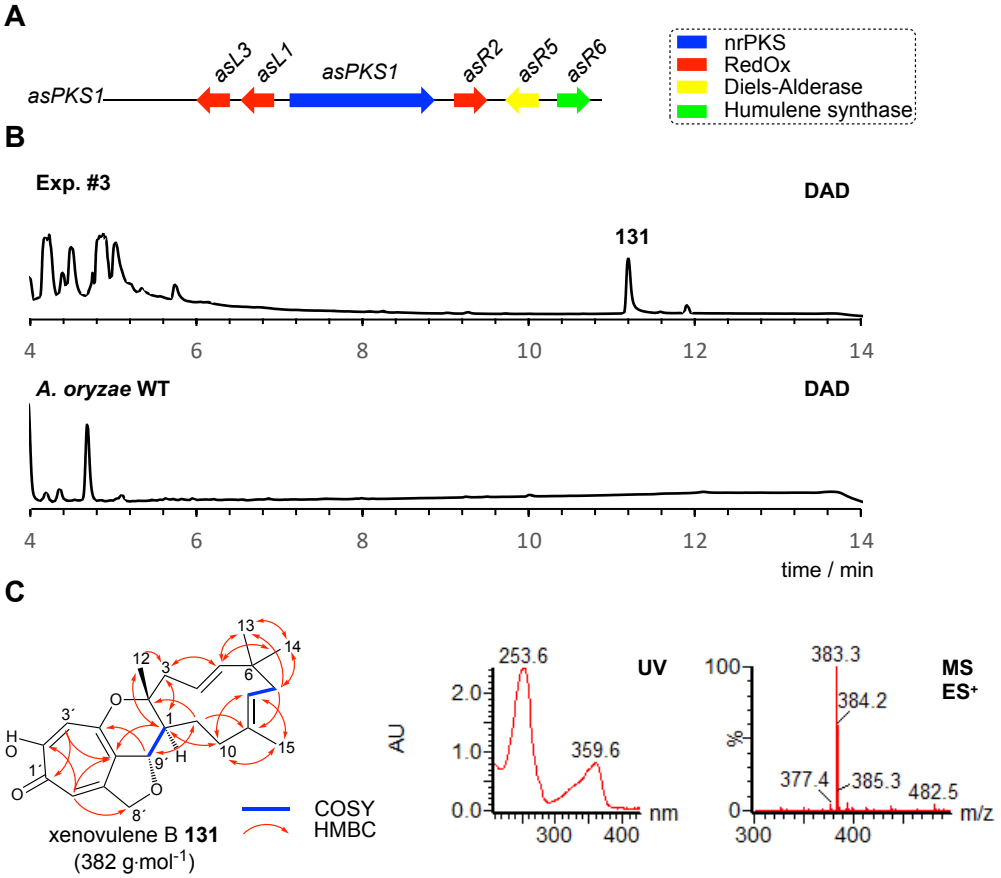



**Figure 2.9** Overview heterologous expression of the minimal *asPKS1* BGC omitting *asL4* and *asL6*: **A,** co-expression of six genes from the *asPKS1* BGC; **B,** LCMS DAD chromatograms of organic extracts of *A. oryzae* expression strain and wild-type control (arbitrary units); **C,** chemical characterisation of xenovulene B **131** by NMR spectroscopy (key COSY and HMBC correlations highlighted), UV and mass spectrum obtained from $ES^+$.

Transformation of *A. oryzae* NSAR1 with six genes (*asPKS1* + *asL1* + *asL3* + *asR2* + *asR5* + *asR6*) yielded five transformants that were shown to have integrated all gene cassettes successfully. Crude extracts were analysed by LCMS (Figure 2.9 A + B). Comparison of the data to a wild-type control identified in addition to the typical tropolone intermediates/shunts a single non-polar new compound **131** with a retention time $t_R = 11.1$ min. The nominal mass (382), retention time and UV absorption ($\lambda = 253$ nm, 360 nm) were consistent with literature data reported for xenovulene B **131**, which was previously isolated from *Acremonium strictum* wild-type.[123] Large-scale fermentation (0.5 L) and subsequent mass-directed purification to homogeneity afforded 5.1 mg. HRMS analysis ($[M]H^+$ $C_{24}H_{31}O_4$ calculated 383.2222, found

383.2223) supported **131** to be xenovulene B and full NMR characterisation confirmed the structure (Figure 2.9 C and Table 2.3). $^{1}$H and $^{13}$C-NMR data was identical to data reported previously for **131**.[123]

**Table 2.3** Chemical shifts of **131** in $CDCl_3$ (500 MHz).

| Pos. | $\delta_H$ / ppm (mult, *J* in Hz) | $\delta_C$ / ppm | C-type | HMBC (H to C) | COSY |
|---|---|---|---|---|---|
| 1 | 1.94 (dd, 10.6, 8.1) | 40.5 | CH | 2, 3, 10, 12, 5′, 9′ | 9′ |
| 2 | - | 86.8 | C | - | - |
| 3a | 2.32 (m) | 43.2 | $CH_2$ | 4, 5 | - |
| 3b | 2.63 (dt, 14.7, 2.2) | | | 1, 2, 4, 5 | - |
| 4 | 4.94 (m) | 119.2 | CH | - | - |
| 5 | 5.17 (dd, 15.8, 1.8) | 144.0 | CH | 3, 6, 13, 14 | - |
| 6 | - | 38.5 | C | - | - |
| 7a | 1.75 (dd, 12.9, 4.8) | 41.7 | $CH_2$ | 5, 6, 8, 9, 13, 14 | 8 |
| 7b | 2.18 (d, 12.5) | | | 6, 8, 9, 13, 14 | 8 |
| 8 | 5.06 (ddd, 11.7, 4.5, 1.6) | 123.8 | CH | 7, 10, 15 | - |
| 9 | - | 136.3 | C | - | - |
| 10a | 2.12 (m) | 38.0 | $CH_2$ | 1, 11, 15 | - |
| 10b | 2.26 (dd, 11.9, 4.9) | | | 1, 8, 9, 11, 15 | - |
| 11a | 1.37 (m) | 28.3 | $CH_2$ | 1, 2, 9, 10, 9′ | - |
| 11b | 1.65 (m) | | | 1, 9′ | - |
| 12 | 1.26 (s) | 22.7 | $CH_3$ | 1, 2, 3, 4, | - |
| 13 | 1.04 (s) | 30.0 | $CH_3$ | 5, 6, 7, 14 | - |
| 14 | 0.97 (s) | 24.3 | $CH_3$ | 5, 6, 7, 13 | - |
| 15 | 1.61 (s) | 17.1 | $CH_3$ | 8, 9, 10 | - |
| 1′ | - | 173.5 | C | - | - |
| 2′ | - | 166.1 | C | - | - |
| 3′ | 6.96 (s) | 111.0 | CH | 1′, 2′, 4′, 5′ | - |
| 4′ | - | 158.9 | C | - | - |
| 5′ | - | 121.3 | C | - | - |
| 6′ | - | 152.1 | C | - | - |
| 7′ | 6.91 (s) | 113.1 | CH | 1′, 2′, 4′, 5′, 6′, 8′ | - |
| 8′a | 4.90 (m) | 75.2 | CH | - | - |
| 8′b | 4.95 (m) | | | - | - |
| 9′ | 4.80 (dd, 10.6, 2.1) | 84.2 | CH | 1, 11, 4′, 5′, 6′ | - |

## 2.4. *In Vitro* Analysis of Key Biosynthetic Enzymes

### 2.4.1. Identification of Hetero DIELS-ALDERase and Terpene Cyclase

Based on the preliminary analysis of the xenovulene A biosynthetic pathway, AsR5 and AsR6 are responsible for humulene formation and catalysis of the subsequent DIELS-ALDER reaction.[123] Unfortunately, while omission of either gene from the expression system abolished xenovulene A **1** production, no new metabolites were observed.[123] The chosen approach is thus unsuited to further differentiate between the function of both genes. Simultaneously, bioinformatic analysis (using blastp,[142] InterPro[146] and NCBI´s conserved domains tool)[147] of both proteins did not allow determination of a possible role based on sequence similarities to proteins of known function: Neither enzyme has any similarity to known biosynthetic enzymes. More sophisticated analysis using the Phyre2 protein fold recognition server[148] did not reveal any further information on AsR6. However, Phyre2 suggested AsR5 to be a $Ca^{2+}$-dependent phosphotriesterase (Appendix section 8.1). Although class I terpene cyclases typically coordinate

$Mg^{2+}$, the presence of a divalent cation binding site might hint at a novel metal binding motif in AsR5.[39,148] Structural predictions using the fully automated protein structure homology modelling server (SWISS-MODEL)[149] proposed a six-bladed propeller structure for AsR5 though (based on the structure of a putative lactonase; template ID 2P4O.1; Figure 2.10). Initial searches with the AsR6 protein sequence did not return any structural hits. However, in the course of this study (probably) improved underlying algorithms finally suggested a fragmentary model based on the structure of the terpene cyclase pentalenene synthase (template ID 1HM4.1.A; Figure 2.10).

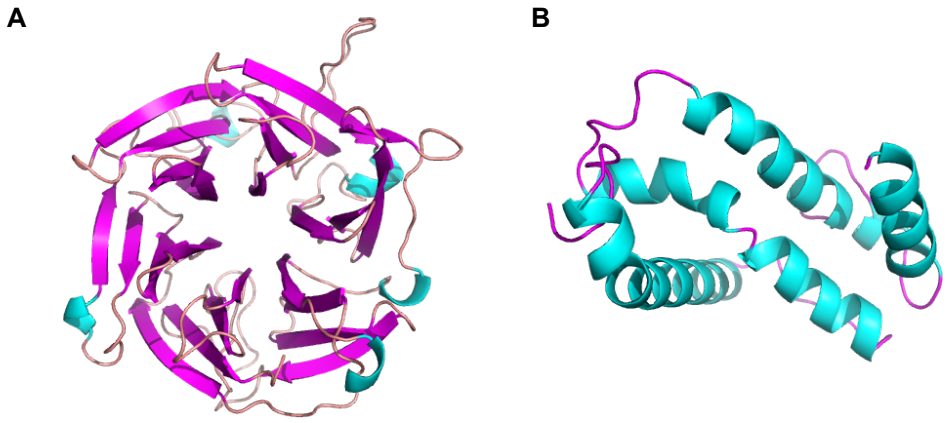



**Figure 2.10** Structural prediction of: **A,** AsR5; and **B,** AsR6; models were created using SWISS Model.[149]

Notably, both models were of poor quality, based on the low sequence identity of query and template (AsR5: 21.68%; AsR6: 18.80%) and the calculated QMEAN values (AsR5: -5.11; AsR6: -4.97). QMEAN values of approximately 0 qualify as good alignments, whereas QMEAN values < -4 indicate poor quality models.

To shed light on the biosynthetic role of *asR5* and *asR6* both were heterologously expressed in *E. coli*. For this, synthetic gene cassettes codon optimized for high protein expression in *E. coli* were obtained from *Thermo Fischer Scientific*. Expression of *asR6* in *E. coli* BL21 afforded recombinant *N*-$His_6$-tagged AsR6 in good yields (for details of expression and purification see Section 4.3). *N*-$His_6$-tagged AsR5 proved insoluble upon expression in *E. coli* BL21 (data not shown). To obtain soluble recombinant AsR5, the codon-optimized *asR5* gene cassette was cloned downstream of a thioredoxin tag (TrxA), which is routinely used to enhance the solubility of fusion proteins.[150] Heterologous expression of *asR5* on pET32a (plasmid CS I 01 J) in *E. coli* Arctic Express cells, co-expressing the chaperonins Cpn60 and Cpn10 that confer high refolding activity between 4 °C and 12 °C,[151,152] finally afforded soluble *N*-TrxA-$His_6$-tagged AsR5.[2]

[2] Heterologous expression of *asR5* was performed by RAISSA SCHOR. Purification and *in vitro* assays by CARSTEN SCHOTTE. Plasmid CS I 01 J cloned in joint work.

Purification was achieved by $Ni^{2+}$ affinity chromatography and after cleavage of the *N*-TrxA-$His_6$-Tags with enterokinase (*New England Biolabs*), the obtained recombinant AsR5 (Figure 2.11) was used to determine the bioactivity.

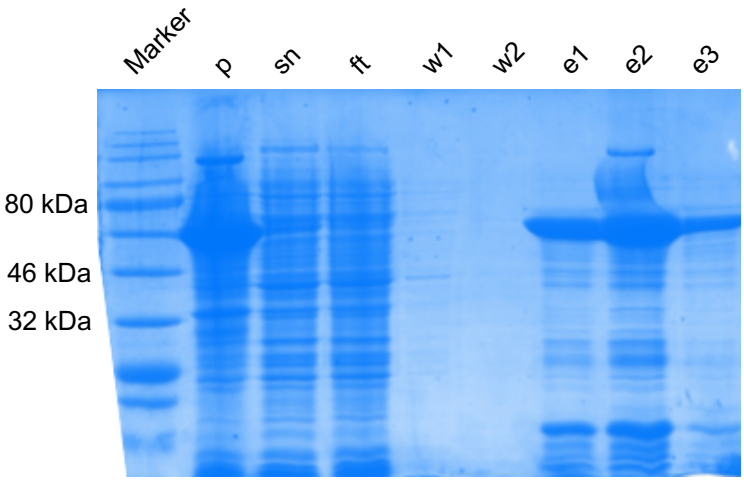
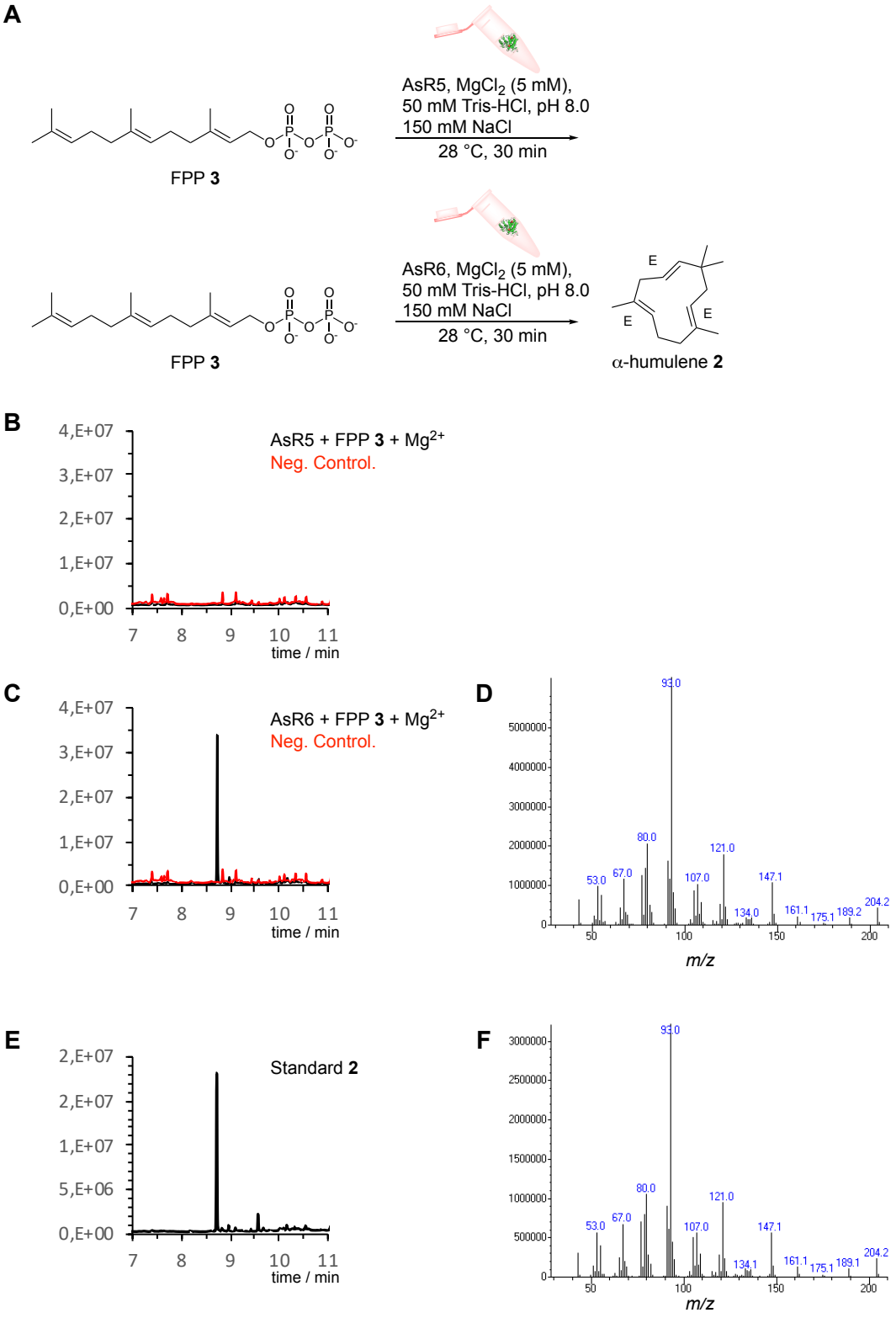



**Figure 2.11** Heterologous expression of *N-TrxA-His$_6$-asR5*: SDS-PAGE analysis of the purification process; Marker = *NEB* Color Prestained Protein Standard, Broad Range (11 - 245 kDa); p = pellet; sn = total lysate; ft = flow-through; w1 = wash step 1; w2 = wash step 2; e1-e3 = elution with 1 ml 500 mM imidazole.

Weak homology between AsR6 and pentalenene synthase suggested that it could be the required terpene cyclase. Initial experiments therefore focussed on testing its activity with FPP **3**. With FPP **3** in hand, probing the capability of humulene formation by either AsR5 or AsR6 was straightforward. Following standard protocols for monitoring terpene cyclase activity,[153] the activity was assessed on a 500 µl scale. Co-incubation of either *N*-$His_6$-tagged AsR6 or AsR5 with FPP **3** in the presence of $Mg^{2+}$ was followed by *n*-hexane extraction and subsequent analysis by gas-chromatography coupled to mass spectrometry (GCMS).[3] A major new peak, identified as α-humulene **2** by comparison of the retention time and mass fragmentation pattern to a standard of **2** (*Sigma Aldrich*), was only observed in assays with AsR6 (Scheme 2.10). No other new peaks were observed upon incubation of FPP **3** with AsR6, demonstrating α-humulene **2** to be the sole product. The data thus conclusively demonstrates AsR6 to be a novel type of terpene cyclase responsible for humulene **2** formation during xenovulene A **1** biosynthesis (for detailed analysis of AsR6 see Section 4). With AsR6 responsible for the formation of humulene **2**, this leaves AsR5 as the remaining candidate enzyme to catalyse the unique hetero DIELS-ALDER reaction during **1** formation. However, in absence of the native substrate **168** the *in vitro* activity of AsR5 was not further investigated.

[3] The assistance of CLARA OBERHAUSER and DR. GERALD DRÄGER to establish suitable GCMS methods is gratefully acknowledged.




**Scheme 2.10** Characterisation of AsR5 and AsR6: **A,** *in vitro* assays performed with AsR5 and AsR6**; B,** GCMS TIC chromatogram of *n*-pentane extract of assay with AsR5 and **3**; **C,** GCMS TIC chromatogram of *n*-pentane extract of assay with AsR6 and **3**; **D,** EI mass spectrum of **2** produced by AsR6; **E,** GCMS TIC chromatogram of standard **2**; **F,** EI mass spectrum of standard **2**.

### 2.4.2. Analysis of Putative Ring-Contraction Enzymes AsL4 and AsL6

The FAD-dependent monooxygenases AsL4 and AsL6 were shown to catalyse the observed ring-contractions during xenovulene A **1** biosynthesis. Omission of the two corresponding gene cassettes from expression in *A. oryzae* NSAR1 resulted in loss of **1** production and only production of xenovulene B **131** was observed. Subsequently, substantial efforts were made in the course of this study to investigate the two enzymes and the proposed ring-contractions *in vitro*. However, *asL4* and *asL6* proved challenging targets for expression as *N*-$His_6$-tagged enzymes in *E. coli* as all attempts resulted in the preparation of insoluble protein (data not shown). Further attempts to co-express the gene cassettes with *C*-terminal $His_6$-Tags (instead of the more commonly used *N*-terminal $His_6$-Tag) proved likewise ineffective. Attempts to co-express *asL4* and *asL6* with chaperones (*e.g.* using *E. coli* Arctic Express[151,152] and *E. coli* TaKaRa[154]) similarly did not lead to the production of soluble enzyme. A detailed *in silico* analysis using the *phobius web server*[155] revealed putative transmembrane spanning domains, that could interfere with the preparation of soluble enzyme (Figure 2.12). As a consequence, *N*- and/or *C*-terminally truncated constructs of *asL4* (encoding for $AsL4_{29\text{-}376}$ and $AsL4_{1\text{-}376}$ respectively) were used for expression experiments, but similarly did not lead to active enzymes. In a final attempt *asL4* and *asL6* were expressed in *S. cerevisiae* W303B but no protein production was observed (data not shown). In the absence of soluble preparations of AsL4/AsL6 the mechanism of ring-contraction was not further investigated.

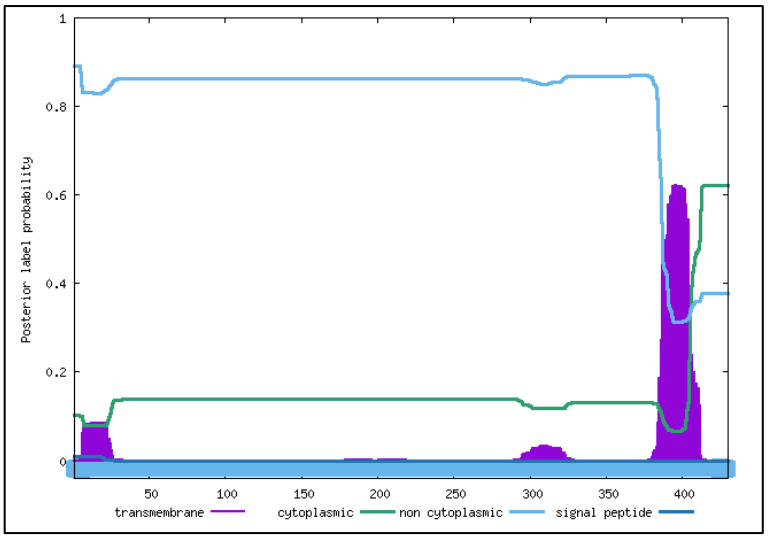
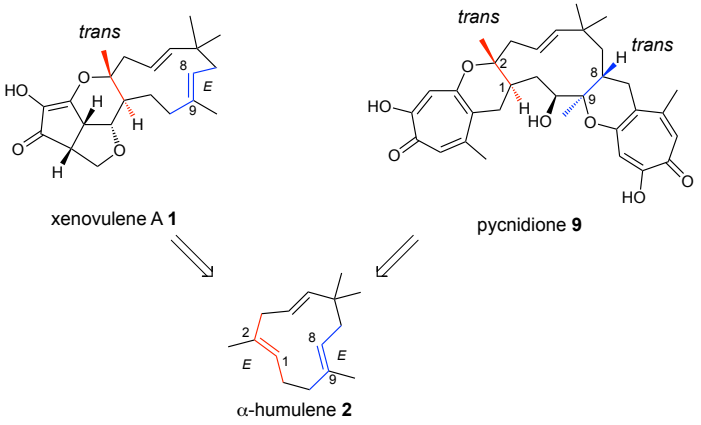



**Figure 2.12** Phobius analysis of the AsL4 amino acid sequence for the presence of signal peptides and transmembrane spanning domains.[155]

## 2.5. Identification of Homologous Terpene Cyclases

As outlined in section 1.4.2 the xenovulenes belong to a wider family of tropolone sesquiterpenoid natural products. In the course of this study fellow PhD student LEI LI (Cox research group) obtained a fungal producing organism (*Leptobacillium* sp. CF-236968) of the bistropolone pycnidione **9** from the Fundación MEDINA. Pycnidione **9** was also proposed to be derived from α-humulene **2**.[83] The all-*trans* olefin configuration of **2** is mirrored in the *trans* C-1/C-2 and C-8/C-9 ring junctions in pycnidione **9** (Scheme 2.11).




**Scheme 2.11** Chemical structure and proposed biosynthetic origin of pycnidione **9**.

A putative biosynthetic gene cluster for the production of pycnidione **9** was identified by LEI LI (COX group) in the genome sequence of *Leptobacillium* sp. CF-236968 (*pyc* BGC). Subsequent global sequence alignment of the *pyc* and *asPKS1* BGC using Clinker & Clustermap.js[143] showed both clusters to share a high degree of homology. Core genes required for tropolone formation (*pycPKS/asPKS1, pycL1/asL1, pycL3/asL3*), DIELS-ALDER reaction (*pycR1*/*asR5*) and humulene formation (*pycR6/asR6*) are conserved in both clusters (Figure 2.13). Attempts to link the *pyc* BGC to the production of pycnidione **9** were unsuccessful, as *Leptobacillium* sp. proved resistant to all commonly deployed antibiotics (testing performed by LEI LI).

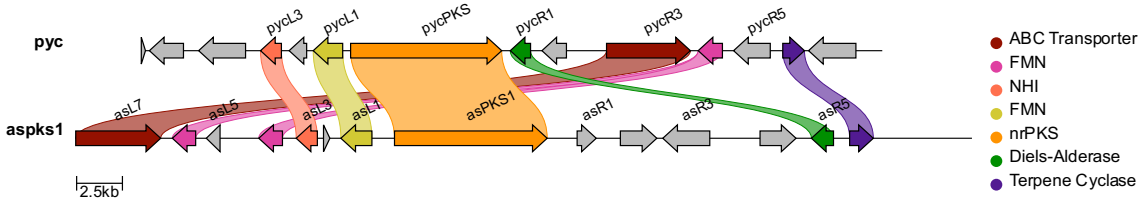
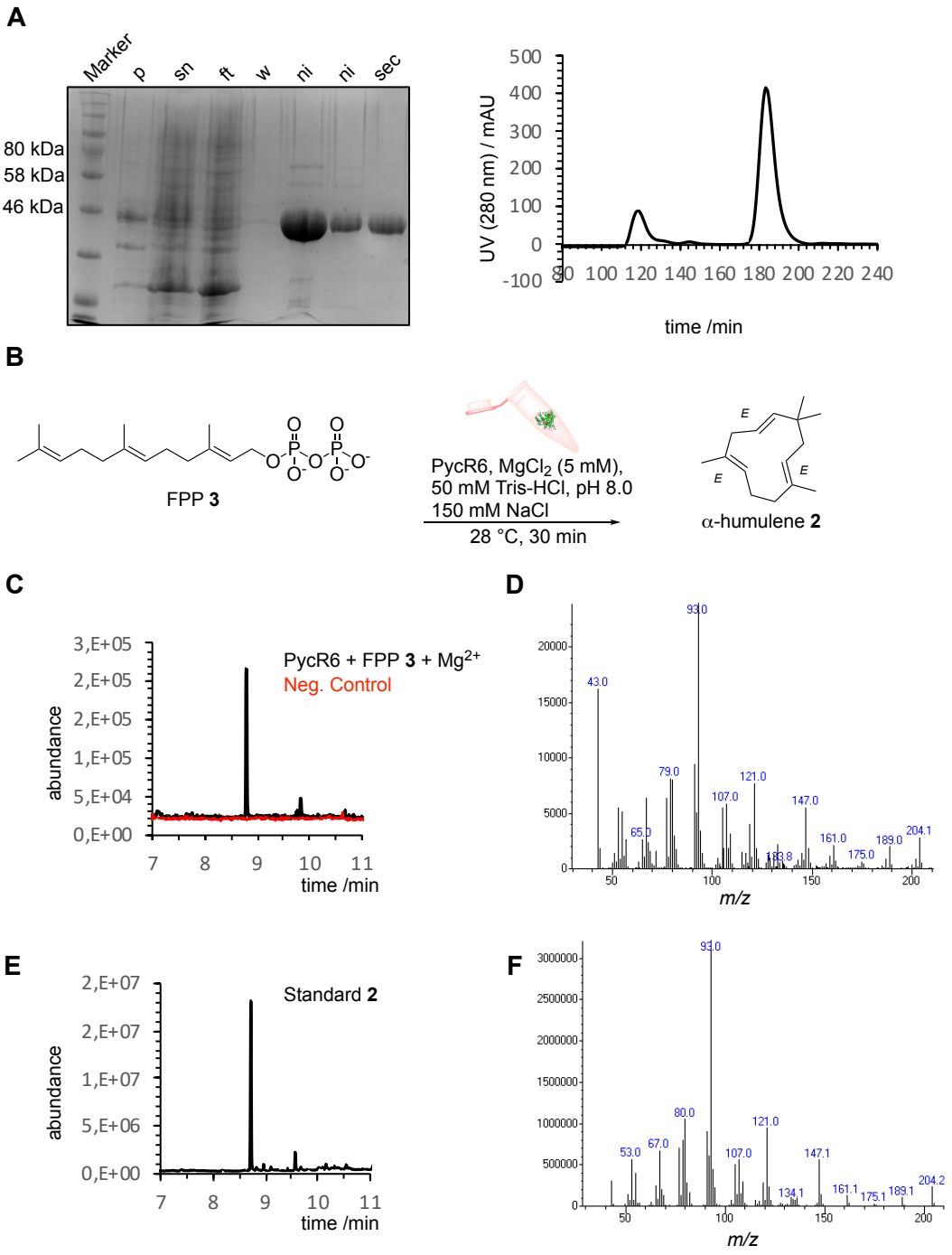



**Figure 2.13** Cluster comparison between the *asPKS1* BGC and the *pyc* BGC: Global sequence alignment and visualisation were achieved with Clinker & Clustermaps.js;[143] FMN = FAD-dependent monooxygenase; NHI = non-heme iron dioxygenase; nrPKS = non-reducing polyketide synthase.




**Scheme 2.12** Purification and *in vitro* activity of PycR6: **A,** SDS PAGE of PycR6 purification (left); Marker (*NEB* Color Prestained Protein Standard, Broad Range (11 - 245 kDa); p = pellet; sn = total lysate; ft = flowthrough; w = wash fraction; ni = eluate $Ni^{2+}$ affinity chromatography; sec = eluate size-exclusion chromatography; UV-chromatogram ($\lambda = 280$ nm) of size-exclusion chromatography (right); **B,** Performed *in vitro* assays with PycR6; **C,** GCMS TIC chromatogram of n-hexane extract of *in vitro* assay with PycR6 and FPP **3**; **D,** EI mass spectrum of **2** produced by PycR6; **E,** GCMS TIC chromatogram of standard **2**; **F,** EI mass spectrum of standard **2**.

We reasoned that PycR6 should produce α-humulene **2**. In the absence of successful gene knockouts we decided to link the cluster to production of tropolone sesquiterpenoids through heterologous expression of *pycR6* and subsequent characterisation of its potential to produce humulene **2** *in vitro*. To validate our hypothesis the codon-optimized *pycR6* gene fused to an *N*-

terminal $His_6$-Tag was expressed in *E. coli* BL21 (identical conditions as reported for *asR6*). *N*-$His_6$-PycR6 was isolated from 1 L expression cultures by $Ni^{2+}$ affinity chromatography and further purified using size-exclusion chromatography (Scheme 2.12 A).

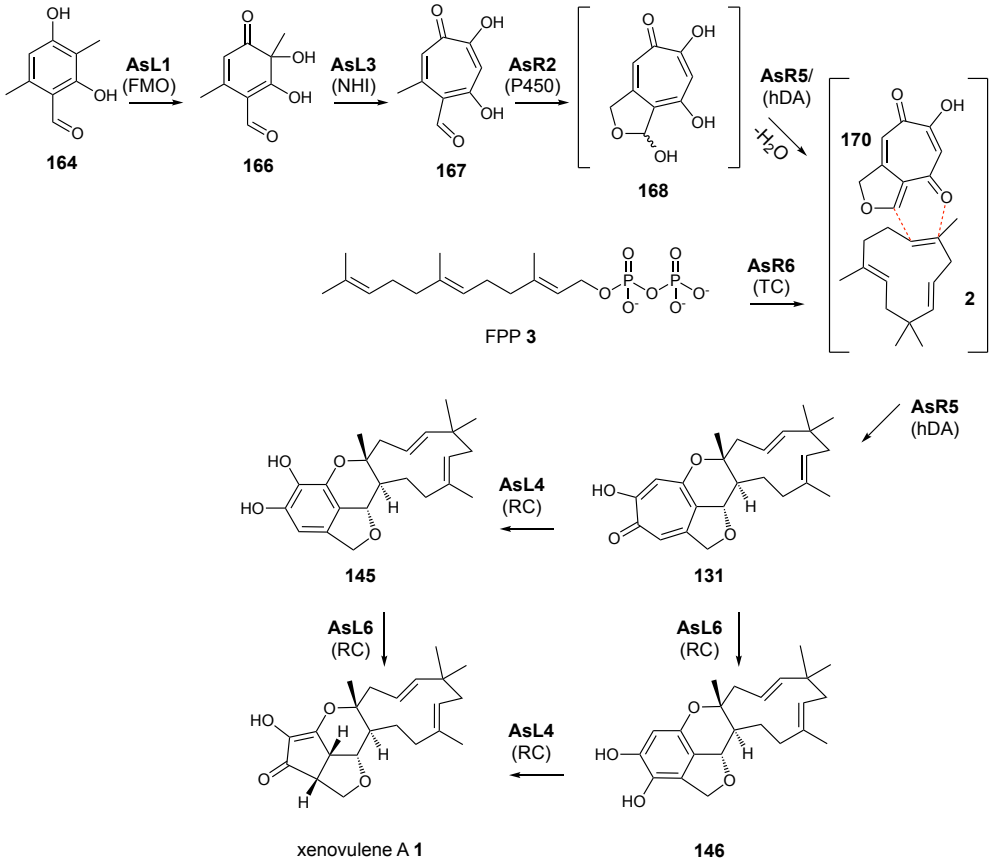
Correct expression and purification of PycR6 was assessed and confirmed by mass spectrometry (DR. JENNIFER SENKLER, BRAUN research group, Leibniz Universität Hannover; Appendix 8.1). The *in vitro* activity of pure PycR6 was assessed as done previously with AsR6 (Scheme 2.12 B). Incubation with FPP **3** in the presence of $Mg^{2+}$ afforded α-humulene **2**, as identified by GCMS analysis and comparison of retention time and mass fragmentation pattern to a commercial standard (*Sigma Aldrich*, Scheme 2.12 C-F). Notably, PycR6 only gives rise to a single terpene product and no other terpenes were observed upon incubation with FPP **3**. Formation of humulene **2** by PycR6 links the identified *pyc* BGC successfully to the production of pycnidione **9**, as no other homologue of AsR6 was identified within the genome of *Leptobacillium* sp.

## 2.6. Conclusion, Discussion and Outlook

In the course of this study the biosynthetic studies on xenovulene A **1** were complemented and key genes/enzymes were linked to specific biosynthetic functions. Initially, completion of the `knock-out by expression´ strategy deployed by RAISSA SCHOR confirmed the proposed – but previously not validated – minimal *asPKS1* gene cluster. Eight dedicated enzymes (AsPKS1 + AsL1 + AsL3 + AsR2 + AsR5 + AsR6 + AsL4 + AsL6) are required to biosynthesize **1** in *A. oryzae* NSAR1. A further reduction of this minimal gene set was probed (*e.g.* omission of cytochrome P450 encoding gene *asR2*) but not successful in the production of **1**. The conducted *in vivo* experiments were unsuited to differentiate between the function of *asR5* and *asR6* but heterologous expression of both genes in *E. coli* and purification of recombinant proteins allowed their function to be assessed *in vitro*. Conversion of FPP **3** to α-humulene **2** by AsR6 – but not AsR5 –identified AsR6 as the dedicated humulene synthase encoded in the *asPKS1* BGC. By inference, this leaves AsR5 as the sole remaining candidate to catalyse the hetero DIELS-ALDER reaction involved in fusion of the terpene and polyketide moiety. AsR6 does not show sequence homology to other terpene cyclases/DIELS-ALDERases. Given the abundance and structural diversity of tropolone sesquiterpenoids in nature, AsR6 likely represents the first member of a new familiy of class I terpene cyclase. AsR5 could be the first member of a family of hetero DIELS-ALDERases but further studies are required to validate this proposed function. Omission of the FAD-dependent monooxygenases AsL4 and AsL6 afforded xenovulene B **131**, herein observed for the first time upon heterologous expression.

The exact timing of the ring-contraction in regard to the central DIELS-ALDER reaction was previously not known. However, combining the results from this study and from previous studies

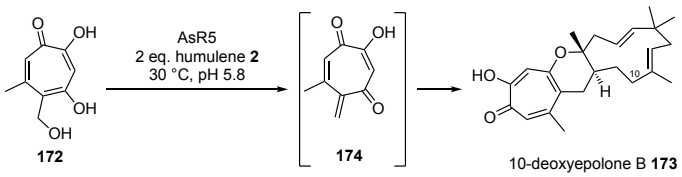
suggest the ring-contractions to occur *after* the DIELS-ALDER reaction. This conclusion is supported by the fact that omission of *asL4* and *asL6* from the expression experiments afforded xenovulene B **131**; and by the fact that no ring contracted tropolones were observed in expression experiments lacking the DIELS-ALDERase *asR5* by Raissa Schor.[123] Taken together the data support the biosynthetic proposal outlined in Scheme 2.13.




**Scheme 2.13** Proposed biosynthesis of xenovulene A **1**: FMO = FAD-dependent monooxygenase; NHI = non-heme iron dioxygenase; P450 = cytochrome P450; hDA = hetero DIELS-ALDERase; TC = terpene cyclase; RC = ring-contraction enzyme.

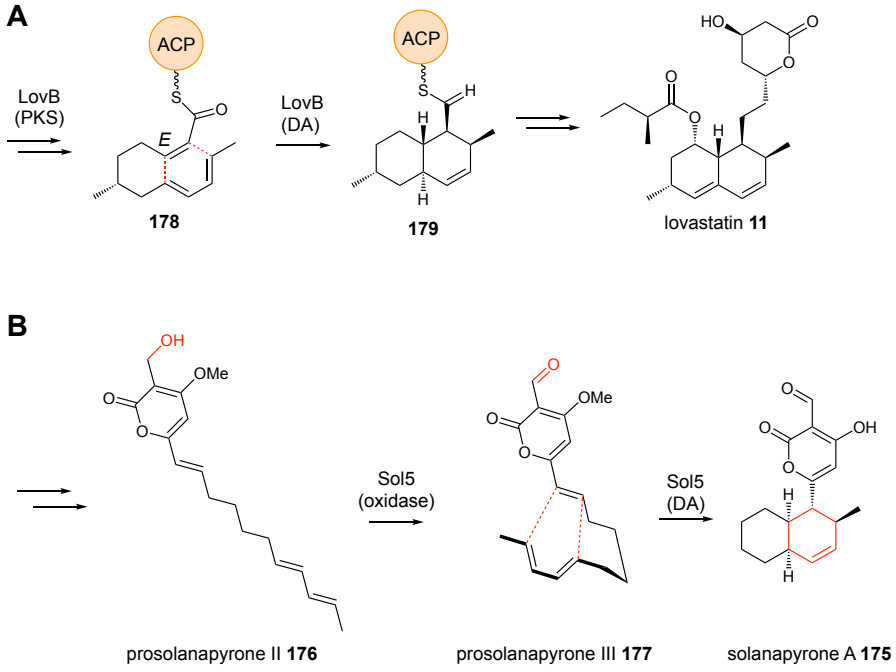
With AsR5 the first fungal *inter*molecular DIELS-ALDERase was putatively identified in the course of this study. However, in absence of the likely native substrate, hemiacetal **168**, the function of AsR5 was not further investigated and a biochemical characterisation is still pending. In the course of this study the doctoral thesis of TYLER JOSEPH DOYON (NARAYAN Lab, University of Michigan, 2020) became publicly available. DOYON demonstrated that AsR5 accepts the non-native substrate stipitol **172**, a proposed key intermediate in the biosynthesis of *e.g.* eupenifeldin **5** (see Section 3). In the presence of α-humulene **2**, stipitol **172** was converted to 10-deoxyepolone B **173**, in 95 % yield (Scheme 2.14). Intriguingly, the activity was pH-dependent, with best conversion at pH 5.8 and 0 % conversion at pH values > 6.8.[156] An acidic environment likely facilitates elimination of water during formation of the *o*-quinomethide **174**, while a basic environment might inhibit the dehydration of **172**. Conversion of stipitol **172** by

AsR5, a non-native substrate, suggests a broader substrate tolerance of the DIELS-ALDERase and in future work the substrate promiscuity of AsR5 should be systematically assessed.




**Scheme 2.14:** *In vitro* activity of AsR5 as reported by TYLER JOSEPH DOYON.[156]

Intriguingly, bioinformatic analysis of AsR5 did not identify homologues of known biosynthetic function. In many examples of recently identified DIELS-ALDERases, the DA-enzyme activity co-occurs with additional biosynthetic capacities (*e.g.*, oxidase or polyketide activity).[157] This additional biosynthetic catalytic function typically precedes the DIELS-ALDER activity, thus priming the subsequent DA-reaction by generating highly reactive biosynthetic intermediates *in situ*, stopping too reactive intermediates from forming undesired side products/shunts.[157,158]




**Scheme 2.15** DA reactions in the biosynthesis of: **A,** lovastatin **11**; and **B,** solanapyrone A **175**.

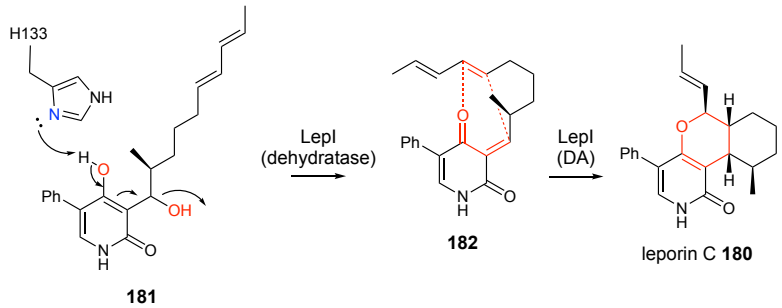
For example, during the biosynthesis of lovastatin **11**, *trans*-decalin formation is catalysed by the polyketide synthase LovB (Scheme 2.15 A).[159] Likewise, the unrelated enzyme Sol5, an annotated oxidase, catalyses *cis*-decalin formation during the biosynthesis of solanapyrone A **175** (Scheme 2.15 B).[160,161]

While for LovB the catalytic domain responsible for the DIELS-ALDER activity has not yet been identified, the mechanism of Sol5 has been investigated in greater detail.[157] Sol5 initially

catalyses oxidation of prosolanapyrone II **176** to afford the aldehyde prosolanapyrone III **177**, before catalysing the *exo*-DA reaction to give solanapyrone A **175**.[161] Both LovB and Sol5 have distinct non-DA functions and it is easy to imagine how a gain-of-function evolution led to the additional DA-activity that exploits structural preorganisation of the substrate on the enzyme scaffold (Scheme 2.15). Apparently, AsR5 also fulfils a dual function as dehydratase and DIELS-ALDERase, but the dehydratase activity evades bioinformatic detection.

Other recent examples of DIELS-ALDERases act as standalone enzymes that, unlike LovB or Sol5, do not catalyse any other biosynthetic transformation but seem to have evolved from enzyme precursors with different functions. For example, the first reported standalone DIELS-ALDERase, SpnF, catalyses a transannular [4+2] cycloaddition to form the cyclohexene core of spinosyn A.[162] SpnF was originally annotated as a SAM-dependent methyltransferase. Similarly, in the biosynthesis of leporin C **180** the key intramolecular DA reaction is catalysed by the annotated SAM-dependent *O*-methyltransferase LepI.[163]

Intriguingly, from a chemical perspective the *intra*molecular biosynthetic mechanism of LepI resembles the proposed *inter*molecular mechanism of AsR5 (Scheme 2.16): Initially, key 2-pyridone alcohol **181** is dehydrated to yield the reactive quinomethide **182**. The subsequent hetero DIELS-ALDER reaction, also catalysed by LepI, establishes the dihydropyran core of leporin C **180**.[163] Further structural characterisation of LepI and sequence alignments with other methyltransferases demonstrated how nature evolved an ancestral methyltransferase into a dual dehydratase/DA-enzyme, where the cofactor SAM is no longer directly involved in catalysis: Disruption of the conserved His-Glu catalytic dyad in LepI prevents the methyltransferase activity, whereas a unique key histidine H133 initiates dehydration *via* deprotonation of the 4-OH group in **181** (Scheme 2.16).[158]



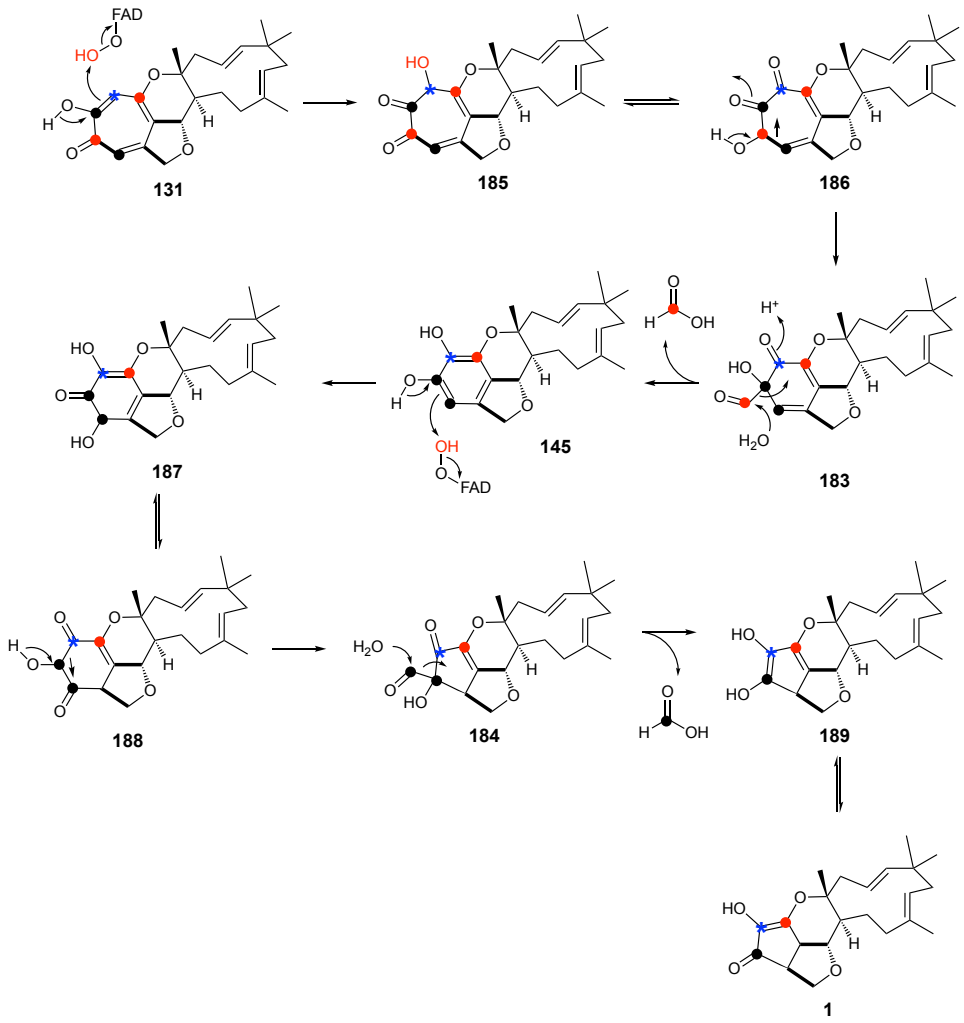

**Scheme 2.16:** Hetero DIELS-ALDER reaction in the biosynthesis of leporin C **180**.

Identification of AsR6 as a humulene synthase unveils a new type of terpene cyclase. Notably, AsR6 does not bear significant sequence homology to either terpene cyclases or to known

humulene synthases. Most importantly, the highly conserved $Mg^{2+}$-binding motifs (DDxx(x)D/E and NSE), typically encountered in class I terpene cyclases, are not readily identified in AsR6.

With ZSS1 a dedicated α-humulene synthase has been linked to the formation of α-humulene **2** during the biosynthesis of the putative anticancer agent zerumbone. Zerumbone is produced in the shampoo ginger *Zingiber zerumbet* Smith but shows only 2.8% sequence identity to AsR6 (as determined with EMBOSS Needle[164]).[88,165] A more detailed analysis of AsR6 is therefore required in order to understand the mechanism of humulene formation and to determine whether the lack of homology on primary sequence level is mirrored on the tertiary sequence level (see Section 4).

Substantial efforts have been directed to investigate the roles of AsL4 and AsL6, the ring-contraction enzymes encoded in the *asPKS1* BGC. Previously, SIMPSON *et al*. proposed a mechanism for an oxidative ring-contraction that agreed with the observed labelling pattern of xenovulene A **1** and B **131** and the wild-type metabolites isolated from *A. strictum* (Scheme 2.17).[87] Two subsequent hydroxylation events (catalysed by AsL4/AsL6) trigger the two rounds of ring-contraction: initial hydroxylation of xenovulene B **131** leads to formation of aldehyde **183** and hydrolytic elimination of formic acid subsequently affords the 5,6-ring system in **145**. The subsequent second round of hydroxylation leads to formation of aldehyde **184** and, again, elimination of formic acid affords the ring-contracted furocyclopentenone of xenovulene A **1**.[166] The initiation of the two ring-contractions *via* hydroxylation is in agreement with Phyre2 analysis of AsL4 and AsL6, that identifies both enzymes as FAD-dependent aromatic hydroxylases (Appendix section 8.1).

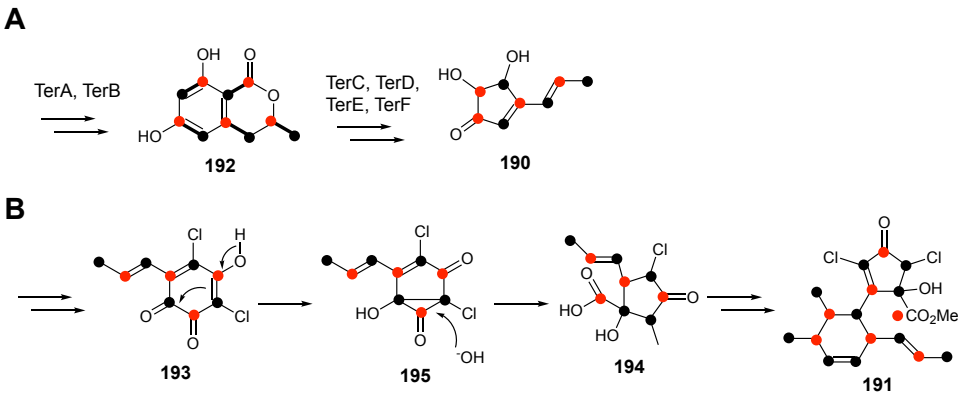



**Scheme 2.17:** Proposed mechanism of oxidative ring-contraction during the biosynthesis of xenovulene A **1**.[87]

Interestingly, AsL4 and AsL6 catalyse chemically equivalent reactions despite sharing only 33.4% sequence identity on amino acid sequence level (determined with EMBOSS Needle[164]). However, AsL4 and AsL6 appear to have distinct regioselective preferences, as the phenolic reaction products of AsL4 and AsL6 differ in the position of one of the two attached hydroxyl groups (Scheme 2.13).[123] It is unfortunate, that in absence of soluble preparations of AsL4 and AsL6, the mechanism could not be investigated in detail.

While oxidative ring-contractions have been proposed for the biosynthesis of other natural products, remarkably few enzymes have been linked to the catalysis of these reactions. For example, isotopic labelling of the fungal metabolites terrein **190** and roussoellatide **191** suggests a ring-contraction to be responsible for generation of the five-membered rings.[167,168] Terrein **190** biosynthesis was studied intensively by BROCK *et al*. and as a key intermediate 6-hydroxymellein **192** is formed by the interplay of a non-reducing polyketide synthase (TerA) and a DH-KR multidomain protein (TerB; Scheme 2.18 A).[167] Four additional enzymes of unknown function are required to furnish terrein **190** (TerC-F) and intriguingly *terC* and *terD* encode FAD-

dependent monooxygenases, although there is only weak homology to AsL4 and AsL6.[123,167] In the biosynthesis of the related metabolite roussoellatide **191** a key favorskii-like ring-contraction was proposed to transform the orthoquinone **193** into the cyclopentenone **194** (Scheme 2.18 B).[168]



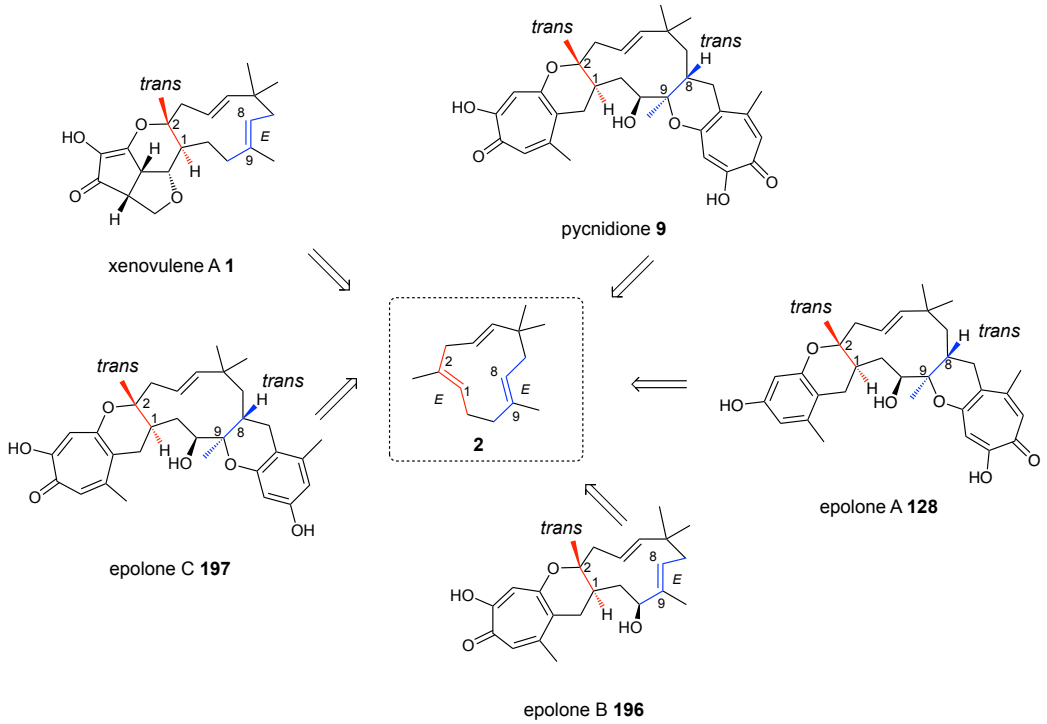

**Scheme 2.18:** Proposed oxidative ring-contractions in the biosynthesis of terrein **190** and roussoellatide **191**.

Other examples of oxidative ring-contractions include the pathways leading to viridicatin (catalysed by cytochrome P450 AsqJ),[77] gibberellic acid (catalysed by cytochrome P450 CYP88A),[169] griseorhodin A (catalysed by FAD-dependent oxygenase GrhO6)[170] and spirotryprostatin (catalysed by cytochrome P450 FqzG).[171]

Finally, initial attempts were conducted to apply the insights gained from xenovulene A biosynthesis to the biosynthesis of related tropolone meroterpenoids. The key structural feature of a tropolone connected to α-humulene **2** *via* a dihydropyran-ring is encountered in numerous literature known metabolites, including pycnidione **9**, epolone A **128** and epolone B **196**.[83,85] In these compounds the all-*trans* double bond configuration of α-humulene **2** is mirrored in *trans*-standing substituents at the western humulene-dihydropyran ring junction and in either a C-8=C-9 *E*-configured olefin or *trans*-standing substituents at the eastern humulene-dihydropyran ring-junction (Scheme 2.19). Notably, the xenovulenes are also related to sterhirsutin A and B as well as hyperjapone A. However, in these compounds the humulene-macrocycle is connected to a non-tropolone moiety and their racemic nature suggests that the core hetero DIELS-ALDER reaction occurs non-enzymatically.[172,173]

In parallel work by LEI LI a biosynthetic gene cluster likely to be responsible for pycnidione **9** and epolone C **197** formation in the fungus *Leptobacillium* sp. CF-236968 was identified (the *pyc* BGC). Recombinant production of PycR6, homologous to the new terpene cyclase AsR6, in this study and subsequent *in vitro* characterisation identified PycR6 to be a second α-humulene synthase. It is reasonable to assume that all fungal tropolone sesquiterpenoids are derived from

similar enzymology and AsR6 and PycR6 constitute the first members of a wider family of new terpene cyclases.



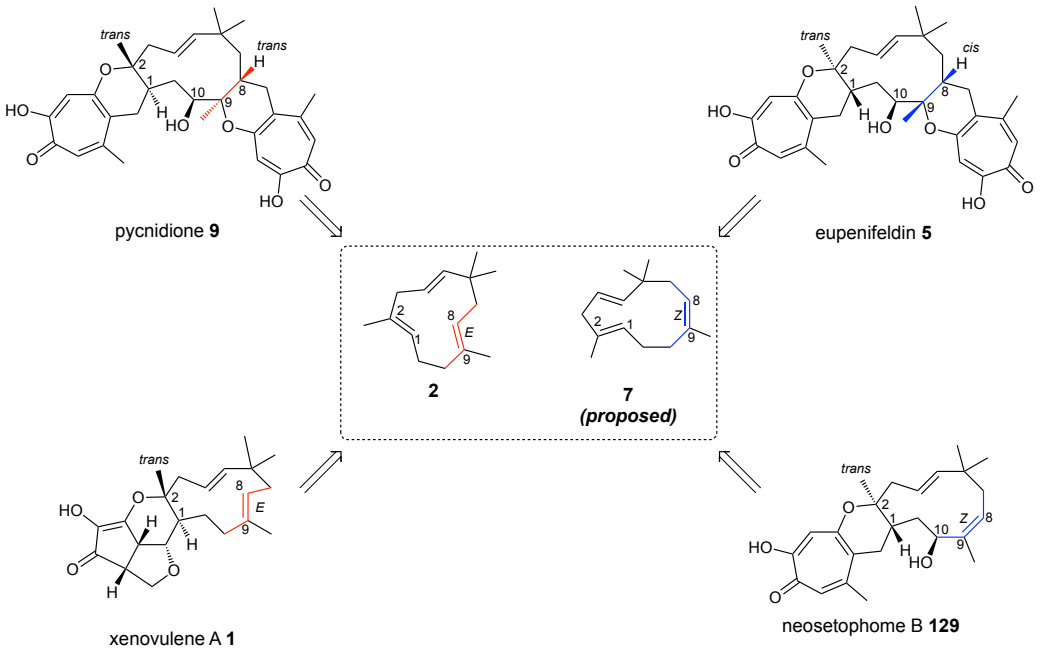

**Scheme 2.19:** Proposed biosynthetic origin of tropolone sesquiterpenoids with an α-humulene backbone; epolone C **197** concomitantly produced with pycnidione **9** in *Leptobacillium* sp. 236968.

# 3. Biosynthetic Studies of Eupenifeldin-type Tropolone Sesquiterpenoids

Part of the presented work has been published in *Angew. Chem. Int. Ed.* **2020**, *59*, 23870-23878.[122]

## 3.1. Introduction

Eupenifeldin **5** has been isolated from a range of fungal sources and is one of the most frequently encountered tropolone sesquiterpenoids found in nature.[78,79,81,174] Structurally, **5** is a bistropolone with two identically substituted tropolone nuclei fused *via* bridging dihydropyran rings to the core humulene-derived macrocycle (Scheme 3.1).[78] In the various producing organisms eupenifeldin **5** is usually produced alongside a range of mono- and bistropolones that vary in the degree of hydroxylation at the C-10 position, and the nature of the tropolone nucleus that is sometimes exchanged for tetrasubstituted benzene rings (*vide infra*).

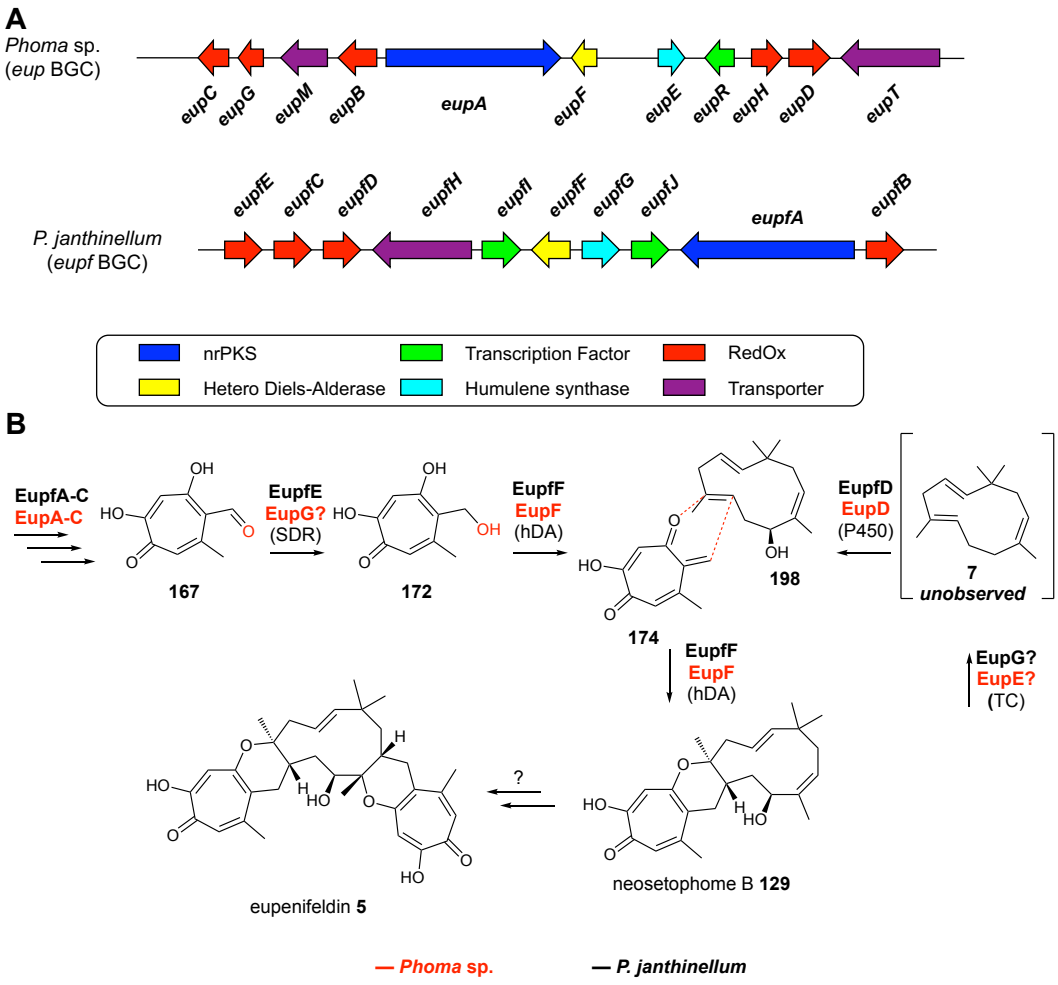



**Scheme 3.1** Proposed biosynthetic origin of eupenifeldin **5** and neosetophome B **129** as opposed to the biosynthetic origin of xenovulene A **1** and pycnidione **9**.

Eupenifeldin-type tropolone sesquiterpenoids differ from the xenovulene-type natural products in respect to the stereochemistry at the C-8/C-9-carbons. Disubstituted compounds such as eupenifeldin **5** display *cis*-fused substituents at the respective humulene/dihydropyran ring junction. Monosubstituted compounds such as neosetophome B **129** contain a *Z*-configured C-8/C-9 alkene instead of the *E*-olefin observed in the xenovulenes.[81] Contrary to the xenovulenes,

the eupenifeldin-type tropolone sesquiterpenoids are thus not derived from all-*trans* α-humulene **2**, but must be derived from an hypothetical (1*E*,4*E*,8*Z*)-humulene **7** precursor if the DIELS-ALDER reaction is stereospecific (Scheme 3.1).[175]

In the course of this study the biosynthetic pathway towards neosetophome B **129** and eupenifeldin **5** was partially validated. In 2019 CHE *et al.* reported the first biosynthetic gene cluster, the *eup* BGC, responsible for formation of eupenifeldin **5** in *Phoma* sp.[80] Through intensive genetic inactivation studies a reasonable biosynthesis of eupenifeldin **5** was proposed. The cluster contained genes homologous to *asR5* and *asR6*, the DIELS-ALDERase and terpene cyclase from the xenovulene A **1** biosynthetic pathway. Later in 2019 HU *et al*. reported a second BGC involved in eupenifeldin **5** and neosetophome B **129** formation, the *eupf* BGC from *Penicillium janthinellum* (Scheme 3.2 A).[175]

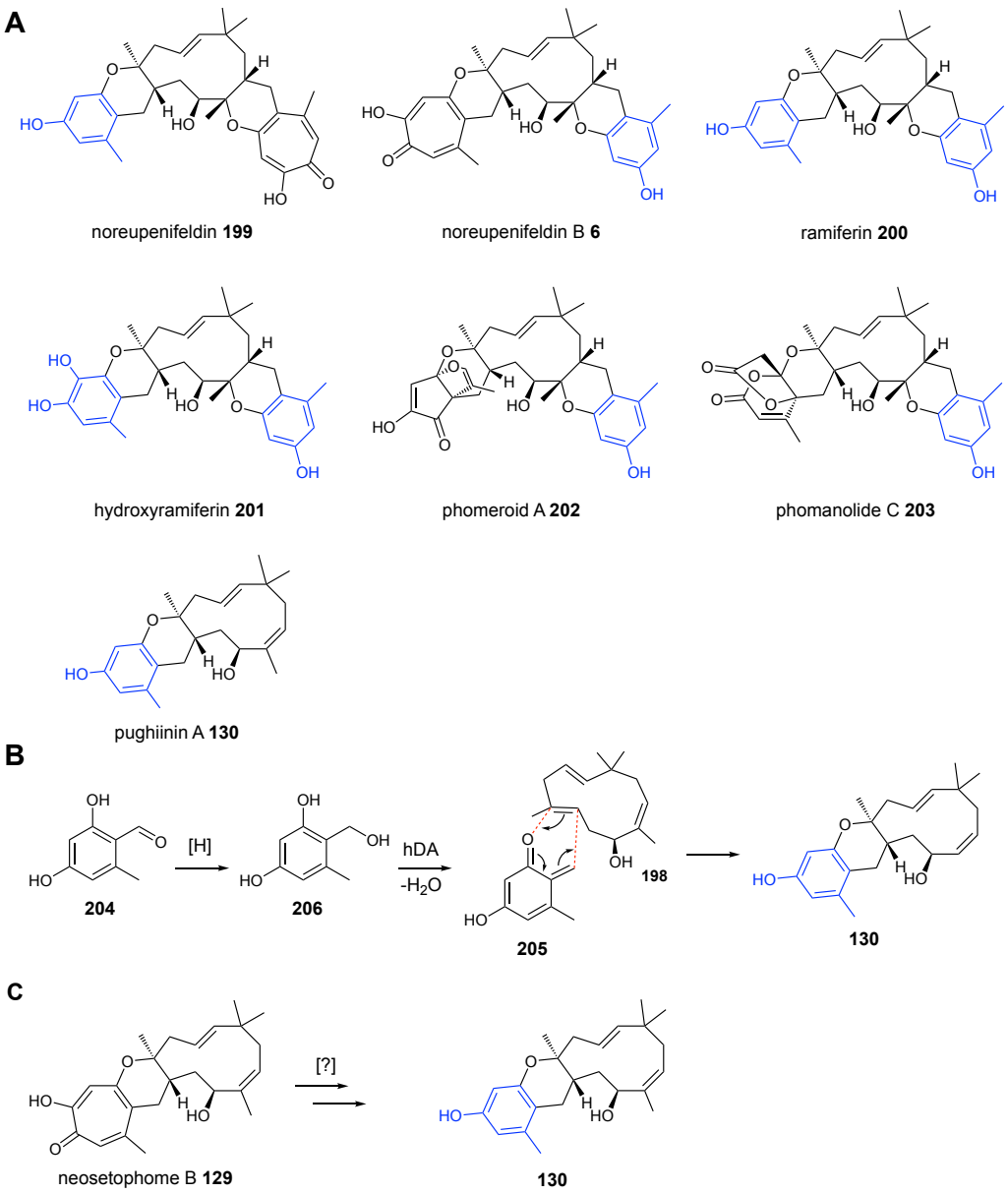



**Scheme 3.2** Biosynthesis of eupenifeldin **5**: **A,** reported BGCs responsible for eupenifeldin **5** biosynthesis; the *eup* BGC is scaled in respect to actual gene size and orientation; schematic representation of the *eupf* BGC as details of the cluster have not been published at the time this thesis was submitted; **B,** proposed biosynthesis of **5**; SDR = short-chain dehydrogenase; hDA = hetero DIELS-ALDERase; P450 = cytochrome P450; TC = terpene cyclase; question marks behind enzymes indicate that the specific function has not been validated *in vitro* or *in vivo*.[80,175]

Together, the results clearly demonstrate that biosynthesis of **5** and **129** proceeds *via* initial tropolone formation: Three genes in *P. janthinellum* (*eupfA-C*), homologous to *tropA-C* (and *asPKS1, asL1, asL3,* respectively), are responsible for the formation of stipitaldehyde **167** (Scheme 3.2 B). A short-chain dehydrogenase (SDR), EupfE, reduces **167** to the corresponding alcohol, stipitol **172**. The DIELS-ALDERase EupfF then catalyzes dehydration of the hydroxymethyl **172** to produce the *o*-quinomethide **174**; EupfF additionally catalyzes the hetero DIELS-ALDER reaction with humulenol **198** that stereoselectively yields neosetophome B **129** (Scheme 3.2 B). Humulenol **198** was shown to be the product of the terpene cyclase EupfG and the cytochrome P450 EupD *in vivo*. Notably, the non-hydroxylated precursor 8*Z*-humulene **7** was not observed and only postulated as the correct terpene product of EupfG.[80,175] Furthermore, only the monotropolone neosetophome B **129** was observed in the conducted expression experiments and in *in vitro* assays; bistropolones have not yet been successfully biosynthesized. The authors thus proposed that additional enzymes have to be involved in the biosynthesis of the bistropolone eupenifeldin **5**.[175]

Eupenifeldin **5** is usually produced with a range of related metabolites, in which one or both tropolone rings are exchanged for a phenol (Scheme 3.3 A). Prominent examples are the regioisomers noreupenifeldin **199**[174] and noreupenifeldin B **6**,[81] ramiferin **200**[79] and its hydroxylated derivative hydroxyramiferin **201**[81] as well as phomeroid A **202**,[176] phomanolide C **203**[177] and pughiinin A **130**.[86] All phenol rings feature at least one hydroxy- and one methyl group-substitution at the same position and are fused to the bridging dihydropyran ring in identical orientations. Tropolone sesquiterpenoids with phenols have been reported to possess intriguing bioactivity. For example, epolone A **128**, a stereoisomer of noreupenifeldin **199**, was reported to induce the production of erythropoietin in human cells at a concentration of 1.0 µM.[85] Pughiinin A **130** has potent antiplasmodial activity ($IC_{50}$ = 2.4 µg·mL$^{-1}$)[86] and ramiferin **200** antimalarial ($IC_{50}$ = 6.3 µg·mL$^{-1}$).[79]

The biosynthetic origin of the six-membered ring remains enigmatic. Both CHE *et al*. and ZHANG *et al*. proposed *o*-orsellinaldehyde **204** as a likely biosynthetic precursor (Scheme 3.3 B).[176,177] With **204** being a common fungal metabolite[178–180] it is easy to imagine how a similar enzymatic machinery as observed for tropolone activation (comprising a short-chain dehydrogenase and DIELS-ALDERase) activates **204** through reduction and dehydration. The resulting active quinomethide **205** can then undergo the DIELS-ALDER reaction with humulenol **198** (Scheme 3.3 B). Alternatively, a ring-contraction similar to that already observed during xenovulene biosynthesis might be responsible for the formation of **130** (Scheme 3.3 C).




**Scheme 3.3** Tropolone sesquiterpenoids with non-tropolone nuclei: **A,** representative examples of tropolone sesquiterpenoids with benzene ring(s) (indicated in blue) in place of the tropolone nucleus; **B,** possible biosynthetic formation of the monobenzopyranyl moiety *via* a DIELS-ALDER reaction; **C,** possible biosynthetic formation of the monobenzopyranyl moiety *via* a ring-contraction of a tropolone precursor.

## 3.2. Project Aims

With the xenovulene A **1**-biosynthetic pathway unveiled, the information needed to identify and characterize biosynthetic pathways to related metabolites was readily available. Of particular interest are tropolone sesquiterpenoids with six-membered rings as these display a broad spectrum of biological activity. The biosynthetic origin of these phenol moieties is currently unknown. We speculated that a ring-contraction of a tropolone precursor was the most likely pathway leading to the six-membered ring, given the prevailing evidence for ring-contractions in the xenovulene A **1** pathway. However, the alternative DIELS-ALDER reaction with **204** could not be excluded

and validation for either pathway is required. Furthermore, the proposed nature of the 1*E*,4*E*,8*Z*-humulene backbone of **5** needs further validation.

Initial screening of fungal strains obtained from cooperation partners at the Fundación MEDINA should establish fungal producers of tropolone sesquiterpenoids with six-membered rings. Using isotopic feeding studies the biosynthetic origin of these compounds will be investigated. In parallel, genome sequencing of producing organism(s) should enable the rapid identification of biosynthetic gene clusters responsible for eupenifeldin **5** biosynthesis in the respective producing organism. With a candidate cluster in hand, heterologous expression of key biosynthetic enzymes should shed light on the identity of the terpene product and possible DIELS-ALDER/ring-contraction steps during tropolone sesquiterpenoid formation. Together, the findings will pave the way for subsequent engineering of the whole pathway and/or individual enzymes.

## 3.3. Results – Analysis of *Phaeosphaeriaceae* sp. CF-150626

To investigate the biosynthesis of eupenifeldin **5** and related compounds five fungal strains were obtained from the Fundación MEDINA (Granada, Spain), where they were previously assessed by LCMS/UV and classified as putative producers of tropolone sesquiterpenoids (Olga Genilloud; personal communication). Three fungal strains (strains CF-150626, CF-253093 and CF-251580) were cultivated in the scope of this thesis and analysed for their potential to produce tropolone sesquiterpenoids. Fungus CF-251580 was quickly eliminated as no production of compounds of interest was detected under the chosen culture conditions. Fungus CF-253093 was established as a good producer of eupenifeldin **5** (data not shown). Fungus CF-150626 was previously described as a prolific producer of eupenifeldin **5** and noreupenifeldin **199**. Indeed, preliminary assessment of this strain returned the most promising results in respect to its ability to produce a range of tropolone sesquiterpenoids and it was investigated in detail.

### 3.3.1. Classification of Fungus CF-150626

Fungus CF-150626 was first described in 2008, with no proper taxonomic identification.[174] Since the time of the first report huge sequencing initiatives such as the 1000 fungal genome project have advanced the publicly available repertoire of fungal genomes dramatically.[181] The internal transcribed spacer region (ITS), a conserved DNA region between the 18S and 28S rRNA genes, is widely used as a barcoding marker for the identification and classification of fungal strains.[182] The CF-150626 ITS sequence was amplified from gDNA using oligonucleotides P11 and P12 and sequenced (*Eurofins Genomics*) to afford a 469 bp DNA sequence (Appendix 8.2). BLASTn sequence comparison (limited to DNA sequences from type material)[142] identified the ITS sequence of *Pseudoophiobolus achilleae* sp. MFLU 17-0925[183] and *Muriphaeosphaeria viburni*

sp. CPC 26610[184] as best alignments. The calculated sequence identities on the DNA sequence level were low (below 95%) and did not allow the classification of fungus CF-150626 at a species level. However, as both strains belong to the fungal family of *Phaeosphaeriaceae*, fungus CF-150626 was preliminarily classified as *Phaeosphaeriaceae* sp. CF-150626. A detailed morphology-based analysis may allow for a more precise classification but was not performed in the scope of this study.

### 3.3.2. Analysis of *Phaeosphaeriaceae* sp. CF-150626 Secondary Metabolite Production

For the production of secondary metabolites *Phaeosphaeriaceae* sp. CF-150626 was cultured under the previously reported fermentation conditions.[174] Growth under producing conditions in PM medium for 13 d was followed by EtOAc extraction. The crude extract was dissolved in acetonitrile:water [9:1] and analysed by LCMS (Figure 3.1). Under standard chromatographic conditions, the non-polar tropolone sesquiterpenoids were expected to elute at the end of the 10-90 % acetonitrile gradient. Indeed, three major constituents were identified eluting between 9.0 min and 11.0 min that showed characteristic UV absorption maxima and had nominal masses within the expected *m/z* range of tropolone sesquiterpenoids. Mass-directed purification of each compound followed by full NMR characterization identified three tropolone sesquiterpenoids.

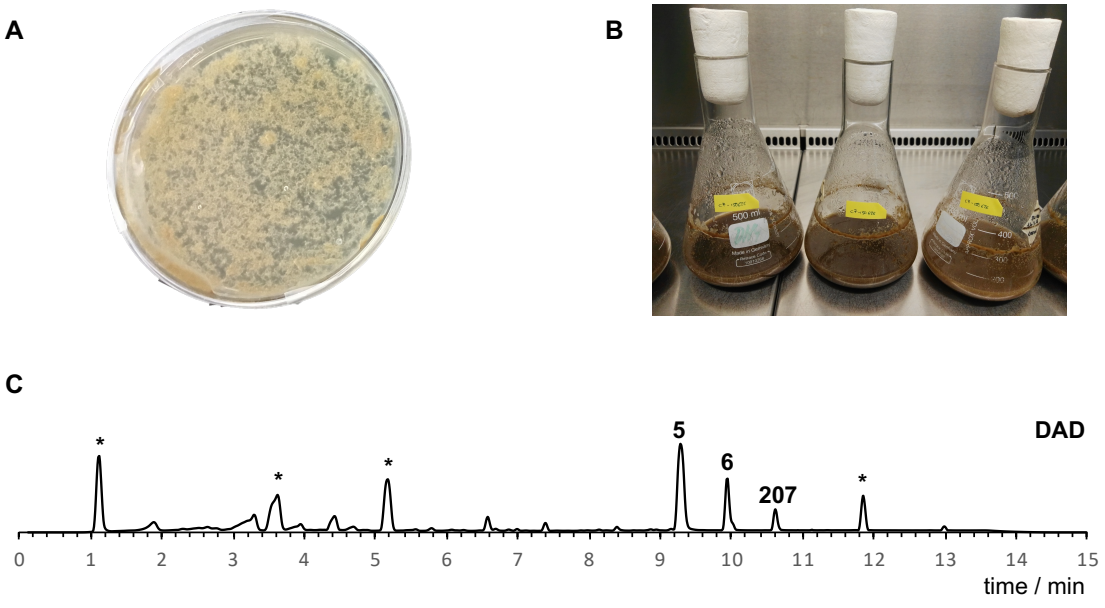



**Figure 3.1** Analysis of *Phaeosphaeriaceae* sp. CF-150626: **A,** 13 d solid culture on PD agar; **B,** 13 d submerged culture in PM medium; **C,** LCMS analysis of crude extract obtained from submerged fermentation of *Phaeosphaeriaceae* sp. CF-150626; displayed is the diode array detector chromatogram of an extract of the culture displayed in B); * = unrelated metabolites.

*Identification of eupenifeldin **5***

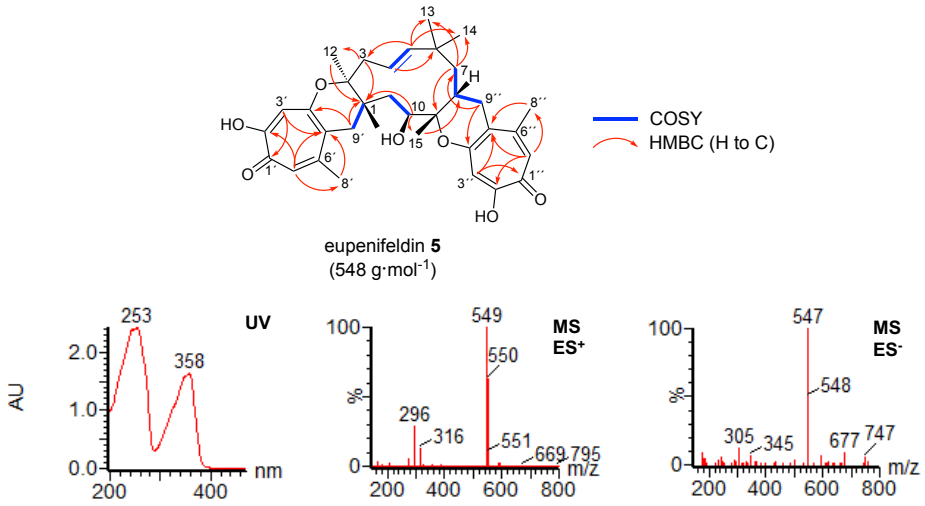



**Figure 3.2:** Chemical characterization of eupenifeldin **5** by NMR spectroscopy (key COSY and HMBC correlations highlighted), UV and mass spectra obtained from $ES^+$ and $ES^-$.

**Table 3.1** Chemical shifts of eupenifeldin **5** in $CDCl_3$ (500 MHz). Referenced to $CDCl_3$.

| Pos. | $\delta_H$ / ppm (mult, *J* in Hz) | $\delta_C$ / ppm | C-type | HMBC (H to C) | NOESY |
|---|---|---|---|---|---|
| **1** | 2.21 (m) | 41.6 | CH | 2, 9′, 10, 12 | |
| **2** | - | 80.5 | C | - | |
| **3a** | 2.73 (ddd, 13.4, 4.3, 1.9) | 46.3 | $CH_2$ | 1, 2, 4, 5 | |
| **3b** | 2.52 (dd, 13.4, 10.7) | | | 1, 2, 4, 5, 12 | |
| **4** | 5.68 (ddd, 15.8, 10,7, 4.3.) | 125.9 | CH | 3, 5, 6 | 3a, 12, 13 |
| **5** | 5.80 (dd, 15.8, 1.9) | 144.1 | CH | 3, 4, 6, 13, 14 | 3b, 8, 14 |
| **6** | - | 35.0 | C | - | |
| **7a** | 1.78 (d, 14.2) overlap | 46.6 | $CH_2$ | 6, 8, 9, 13, 14 | 7b, 10 |
| **7b** | 0.78 (dd, 14.2, 4.3) | | | 5, 6, 8, 13, 14 | |
| **8** | 1.82 (m) | 32.0 | CH | 5′′, 7 | 9′′a, 14, 15 |
| **9** | - | 82.0 | C | - | |
| **10** | 4.22 (d, 10.7) | 70.8 | CH | 1, 9, 11, 15 | 7a, 11b, 12 |
| **11a** | 2.22 (m) | 30.2 | $CH_2$ | - | |
| **11b** | 1.53 (m) | | | 2, 10 | |
| **12** | 1.41 (s) | 19.4 | $CH_3$ | 1, 2, 3 | 10 |
| **13** | 1.07 (s) | 29.7 | $CH_3$ | 5, 6, 7, 14 | 7b |
| **14** | 1.10 (s) | 27.3 | $CH_3$ | 5, 6, 7, 13 | 7b, 8 |
| **15** | 1.14 (s) | 16.1 | $CH_3$ | 8, 9, 10 | 11b, 9′′a |
| **1′** | - | 172.6 | C | - | |
| **2′** | - | 163.2 | C | - | |
| **3′** | 6.94 (s) | 113.5 | CH | 1′, 2′, 4′, 5′ | |
| **4′** | - | 160.3 | C | - | |
| **5′** | - | 122.5 | C | - | |
| **6′** | - | 150.5 | C | - | |
| **7′** | 7.13 (s) | 124.7 | CH | 1′, 2′, 5′, 8′ | |
| **8′** | 2.41 (s) | 27.4 | $CH_3$ | 5′, 6′, 7′ | |
| **9′a** | 3.38 (dd, 18.3, 13.1) | 33.0 | $CH_2$ | 1, 4′, 5′ | 12 |
| **9′b** | 2.41 (m) | | | - | |
| **1′′** | - | 173.3 | C | - | |
| **2′′** | - | 162.9 | C | - | |
| **3′′** | 6.92 (s) | 112.8 | CH | 1′′, 2′′, 4′′, 5′′ | |
| **4′′** | - | 159.5 | C | - | |
| **5′′** | - | 118.7 | C | - | |
| **6′′** | | 151.5 | C | - | |
| **7′′** | 7.16 (s) | 125.6 | CH | 1′′, 2′′, 5′′, 8′′ | |
| **8′′** | 2.37 (s) | 27.5 | $CH_3$ | 5′′, 6′′, 7′′ | |
| **9′′a** | 2.85 (dd, 17.2, 5.3) | 34.4 | $CH_2$ | 7, 8, 4′′, 5′′ | 8, 15 |
| **9′′b** | 2.37 (dd, 17.2) | | | overlap | |

Isolation of the main constituent **5** from the crude extract ($t_R$ = 9.2 min) afforded 17.5 mg from 2 L submerged cultures. The nominal mass (548), characteristic UV absorption maxima ($\lambda$ = 253 nm, 358 nm) and optical rotation ($[\alpha]_D^{28}$ = + 200.0 [$c$ = 0.045 g/100 ml, $CH_2Cl_2$]) of **5** were consistent with previous data published for the bistropolone eupenifeldin.[185] Analysis by HRMS ($[M]H^+$ calculated $C_{33}H_{41}O_7$ 549.2852, found 549.2856) and NMR unambiguously confirmed **5** as eupenifeldin (Table 3.1 and Figure 3.2).

*Identification of noreupenifeldin B* ***6***

A second metabolite **6** eluted at 10.0 min. The nominal mass (520) and UV absorption maxima ($\lambda$ = 210 nm, 253 nm and 360 nm) were in agreement with data previously reported for noreupenifeldin **199**.[174] Purification to homogeneity afforded 5.6 mg from 2 L submerged cultures. HRMS ($[M]H^+$ calculated $C_{32}H_{41}O_6$ 521.2903, found 521.2907) confirmed the expected molecular formula of noreupenifeldin. However, careful analysis of the obtained NMR data revealed small but significant differences in $^1H$ and $^{13}C$ NMR. Full structure elucidation confirmed that **6** was not noreupenifeldin but a regioisomer, in which the position of the benzene ring system was switched (Figure 3.3; Table 3.2). Comparison to literature NMR data identified **6** as noreupenifeldin B, previously reported by OBERLIES *et al*.[81] The optical rotation determined for **6** agreed with the identification as noreupenifeldin B ($[\alpha]_D^{28}$ = + 85.7 [$c$ = 0.047 g/100 ml, $CHCl_3$]). It is intriguing that *Phaeosphaeriaceae* sp. CF-150626 now produces noreupenifeldin B **6**, when initially it was reported to produce noreupenifeldin **199**.[174] The reason for the switch in regioselectivity was not further investigated.

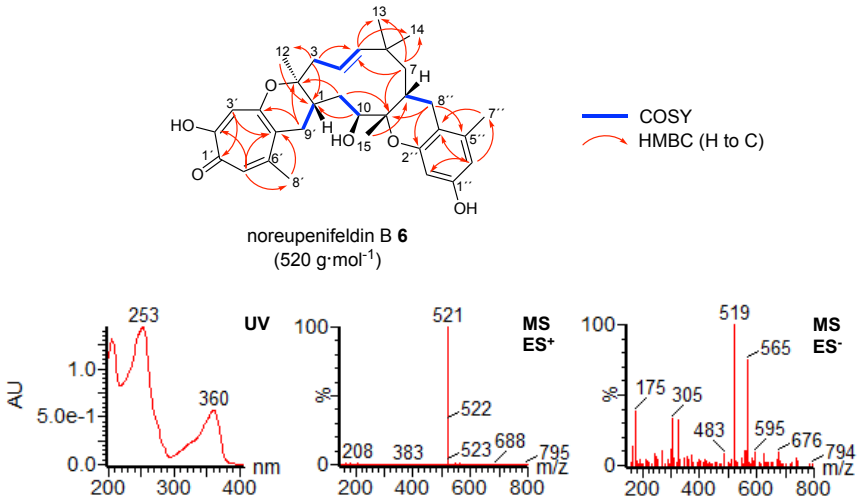



**Figure 3.3:** Chemical characterization of noreupenifeldin B **6** by NMR spectroscopy (key COSY and HMBC correlations highlighted), UV and mass spectra obtained from $ES^+$ and $ES^-$.

**Table 3.2** Chemical shifts of noreupenifeldin B **6** in $CDCl_3$ (500 MHz). Referenced to $CDCl_3$.

| Pos. | $\delta_H$ / ppm (mult, *J* in Hz) | $\delta_C$ / ppm | C-type | HMBC (H to C) | NOESY |
|---|---|---|---|---|---|
| **1** | 2.19 (m) overlap | 41.6 | CH | 9′b | |
| **2** | - | 80.9 | C | - | |
| **3a** | 2.52 (dd, 13.3, 10.7) | 46.3 | $CH_2$ | 4 | 4, 5, 12 |
| **3b** | 2.73 (dd, 13.3, 4.5) | | | 4 | 12 |
| **4** | 5.65 (ddd, 15.9, 10.7, 4.5) | 125.3 | CH | 3, 5 | 3a, 7b, 12, 13 |
| **5** | 5.81 (dd, 15.9, 1.9) | 144.6 | CH | 3, 4 | 3a, 3b, 8 |
| **6** | - | 35.1 | C | - | |
| **7a** | 0.83 (dd, 14.1, 4.5) | 46.3 | $CH_2$ | - | |
| **7b** | 1.79 (d, 14.1) | | | - | 7a, 10 |
| **8** | 1.75 (m) | 32.4 | CH | - | 14, 15, 8a′′ |
| **9** | - | 80.6 | C | - | |
| **10** | 4.19 (d, 11.4) | 71.0 | CH | 11 | 4, 7b, 11a, 12 |
| **11a** | 1.54 (ddd, 14.6, 11.4, 2.4) | 30.4 | $CH_2$ | 10 | |
| **11b** | 2.22 (m) overlap | | | - | |
| **12** | 1.40 (s) | 19.3 | $CH_3$ | - | 3b, 10, 9′b |
| **13** | 1.05 (s) | 29.8 | $CH_3$ | - | 4, 7a, 7b |
| **14** | 1.10 (s) | 27.2 | $CH_3$ | - | 8, 8′′a |
| **15** | 1.09 (s) | 15.8 | $CH_3$ | - | 3b, 8, 8′′a |
| **1′** | - | 172.4 | C | - | |
| **2′** | - | 163.3 | C | - | |
| **3′** | 6.96 (s) | 113.7 | CH | - | |
| **4′** | - | 160.6 | C | - | |
| **5′** | - | 122.9 | C | - | |
| **6′** | - | 150.7 | C | - | |
| **7′** | 7.14 (s) | 124.7 | CH | - | |
| **8′** | 2.41 (s) | 27.5 | $CH_3$ | - | 7′ |
| **9′a** | 2.43 (dd, 18.1, 4.6) | 33.2 | $CH_2$ | 1 | 9′b |
| **9′b** | 3.39 (dd, 18.1, 13.2) | | | 1 | 12, 8′ |
| **1′′** | - | 154.9 | C | - | |
| **2′′** | 6.16 (d, 2.4) | 101.6 | CH | - | |
| **3′′** | - | 153.7 | C | - | |
| **4′′** | - | 111.1 | C | - | |
| **5′′** | - | 139.6 | C | - | |
| **6′′** | 6.33 (d, 2.4) | 110.1 | CH | - | 7′′ |
| **7′′** | 2.16 (s) | 19.6 | $CH_3$ | - | 6′′ |
| **8a′′** | 2.30 (dd, 16.1, 1.7) | 29.7 | $CH_2$ | - | 7a, 8 |
| **8a′′** | 2.71 (dd, 16.1, 4.9) | | | 8 | 8 |

*Identification of dehydroxyeupenifeldin **207***

A third and minor co-metabolite **207** was infrequently observed in the crude extracts of *Phaeosphaeriaceae* sp. CF-150626. The non-polar compound ($t_R$ = 10.3 min) showed identical UV absorption maxima ($\lambda$ = 253 nm and 360 nm) as eupenifeldin **5**. Together with a 16 amu difference in nominal mass (532) this suggested **207** to be a deoxy-analogue of eupenifeldin **5** (548). Isolation and purification afforded 2.8 mg from 2 L submerged cultures. HRMS analysis confirmed the expected molecular formula ($[M]H^+$ calculated $C_{33}H_{41}O_6$ 533.2903, found 533.2912) and NMR analysis confirmed **207** to be dehydroxyeupenifeldin: The C-10 oxygenated carbon in **5** ($\delta_H$ = 4.22 ppm; $\delta_C$ = 70.8 ppm) was replaced by an upfield shifted methylene group in **207** ($\delta_H$ = 1.97/1.61 ppm; $\delta_C$ = 31.7 ppm; Figure 3.4; Table 3.3). Comparison to literature NMR data for **207** revealed identical chemical shifts.[81]

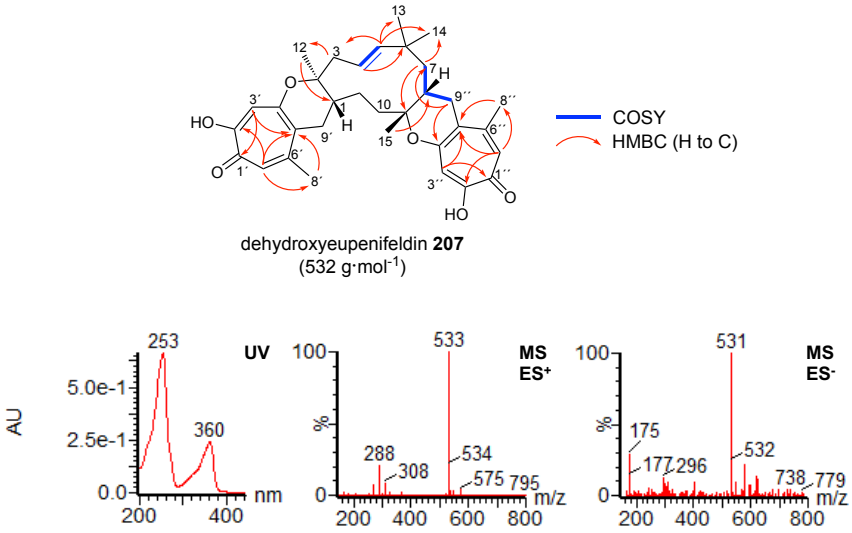


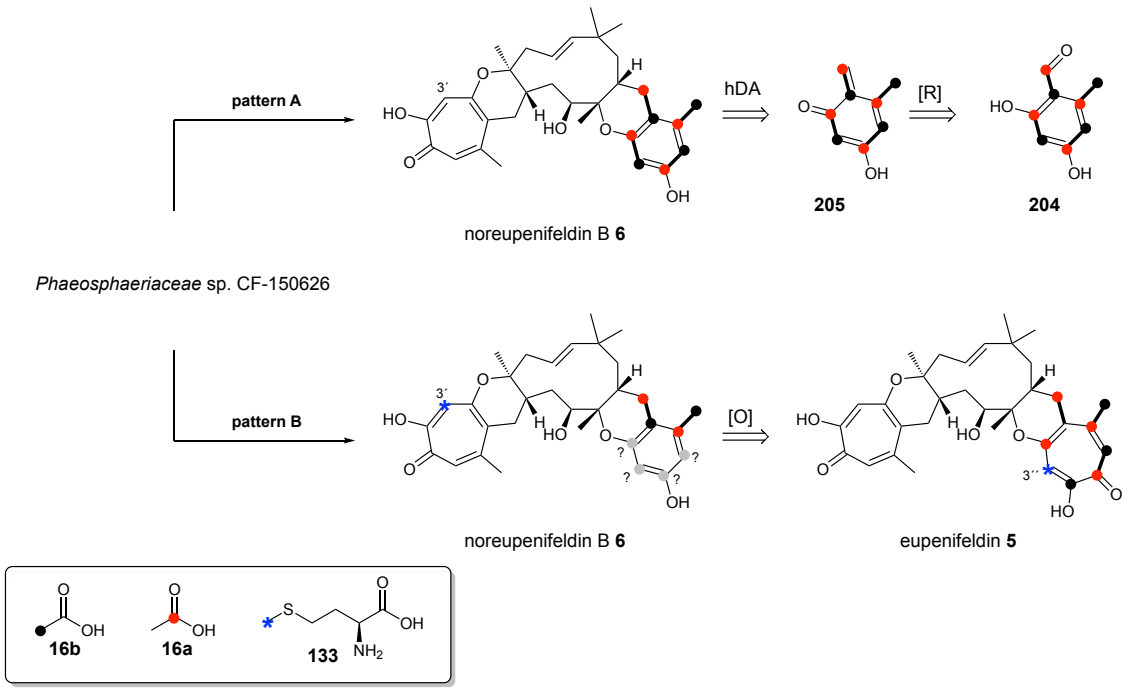

**Figure 3.4** Chemical characterization of dehydroxyeupenifeldin **207** by NMR spectroscopy (key COSY and HMBC correlations highlighted), UV and mass spectra obtained from $ES^+$ and $ES^-$.

**Table 3.3** Chemical shifts of 2.8 mg dehydroxyeupenifeldin **207** in $CDCl_3$ (500 MHz). Referenced to $CDCl_3$.

| Pos. | $\delta_H$ / ppm (mult, *J* in Hz) | $\delta_C$ / ppm | C-type | HMBC (H to C) |
|---|---|---|---|---|
| **1** | 2.19 (m) | 41.5 | CH | - |
| **2** | - | 80.3 | C | - |
| **3a** | 2.68 (m) | 45.5 | $CH_2$ | - |
| **3b** | 2.55 (m) | | | 2, 5, 12 |
| **4** | 5.67 (m) | 124.7 | CH | 3, 5, 6 |
| **5** | 5.88 (d, 16.0) | 144.1 | CH | 3, 6, 13, 14 |
| **6** | - | 35.4 | C | - |
| **7a** | 1.68 (d, 14.6) | 47.1 | $CH_2$ | 6, 8, 9, 14 |
| **7b** | 0.79 (dd, 14.6, 4.4) | | | 5 |
| **8** | 1.85 (m) | 31.7 | CH | - |
| **9** | - | 80.0 | C | - |
| **10a** | 1.97 (m) | 31.7 | $CH_2$ | - |
| **10b** | 1.61 (dd, 13.0, 4.3) | | | 9, 11 |
| **11a** | 2.19 (m) | 24.6 | $CH_2$ | - |
| **11b** | 1.52 (m) | | | |
| **12** | 1.31 (s) | 19.6 | $CH_3$ | 1, 2, 3 |
| **13** | 1.08 (s) | 29.8 | $CH_3$ | 5, 6, 7, 14 |
| **14** | 1.12 (s) | 27.2 | $CH_3$ | 5, 6, 7, 13 |
| **15** | 1.17 (s) | 22.8 | $CH_3$ | 8, 9 |
| **1′** | - | 172.6 | C | - |
| **2′** | - | 163.2 | C | - |
| **3′** | 6.95 (s) | 113.4 | CH | 2′, 5′ |
| **4′** | - | 160.6 | C | - |
| **5′** | - | 120.9 | C | - |
| **6′** | - | 150.3 | C | - |
| **7′** | 7.13 (s) | 124.9 | CH | 1′, 2′, 5′, 8′ |
| **8′** | 2.41 (s) | 27.5 | $CH_3$ | 5′, 6′, 7′ |
| **9′a** | 2.71 (m) | 31.5 | $CH_2$ | - |
| **9′b** | 2.35 (m) | | | - |
| **1′′** | - | 173.4 | C | - |
| **2′′** | - | 162.7 | C | - |
| **3′′** | 6.93 (s) | 113.2 | CH | 1′′, 2′′, 4′′, 5′′ |
| **4′′** | - | 160.9 | C | - |
| **5′′** | - | 118.6 | C | - |
| **6′′** | - | 151.6 | C | |
| **7′′** | 7.14 (s) | 125.3 | CH | 1′′, 2′′, 5′′, 8′′ |
| **8′′** | 2.36 (s) | 27.5 | $CH_3$ | 5′′, 6′′, 7′′ |
| **9′′a** | 2.79 (dd, 16.9, 5.6) | 34.1 | $CH_2$ | 8, 4′′, 5′′ |
| **9′′b** | 2.36 (m) | | | 7, 8, 9 |

### 3.3.3. Isotopic Labelling Studies

To investigate the biosynthetic origin of the monobenzopyranyl moiety in noreupenifeldin B **6** we devised labelling experiments using [1-$^{13}$C]-acetate **16a** and [2-$^{13}$C]-acetate **16b**. The monobenzopyran in tropolone sesquiterpenoids such as **6** was previously proposed to result from a DIELS-ALDER reaction between a mono-tropolone precursor (*e.g.* neosetophome B **129**) and an *o*-quinomethide **205** derived from *o*-orsellinaldeyhde **204**.[176,177] Incorporation of the tetraketide **204** into tropolone sesquiterpenoids should result in a characteristic labelling pattern, as four intact acetate units would be expected to be incorporated (Scheme 3.5; pattern A).[186] A reasonable alternative mechanism would proceed through a ring-contraction of a tropolone precursor (Scheme 3.5; pattern B). In the latter case several labelling patterns are possible – but the ring-contraction is likely to disrupt an intact acetate unit and the labelling pattern should thus be distinguishable from the labelling pattern resulting from incorporation of *o*-orsellinaldeyhde.




**Scheme 3.5** Possible labelling pattern of the monobenzopyranyl-moiety in noreupenifeldin B **6** upon labelling with [1-$^{13}$C]-acetate **16a** and [2-$^{13}$C]-acetate **16b**: pattern A = resulting from the DIELS-ALDER reaction of **129** with **205**; pattern B = resulting from the ring-contraction of eupenifeldin **5**; bold bond = intact acetate unit.

A time course was conducted with the producing fungus, *Phaeosphaeriaceae* sp. CF-150626, to assess the exact timing of tropolone sesquiterpenoid formation (data not shown). Production of eupenifeldin **5** and noreupenifeldin B **6** started at day 4. [1-$^{13}$C]- and [2-$^{13}$C]-acetate **16ab** were therefore pulse-fed in separate experiments to submerged cultures (1 L) of *Phaeosphaeriaceae* sp. on days 3, 4, 5 and 6 to a final concentration of 10 mM. Mass-directed purification to homogeneity afforded [1-$^{13}$C]- and [2-$^{13}$C]-labelled eupenifeldin **5** and noreupenifeldin B **6**.

All compounds were analysed by $^{13}$C-NMR and the obtained spectra were compared to $^{13}$C-NMR spectra of the non-labelled metabolites (hitherto referred to as `natural abundance´ spectra). Peak intensities were then determined using MestReNova 11.0.4 and/or TopSpin 3.5. The obtained data set for each compound (both labelled and unlabelled) was normalized to the C-3´´ (eupenifeldin **5**) and C-3´ (noreupenifeldin B **6**) carbon peak intensity: The carbon at this position is not derived from acetate but rather from *S*-adenosylmethionine and its peak intensity is not expected to be enhanced upon acetate feeding (compare Scheme 3.5):

$$I_{norm} = \frac{I_{C-x}}{I_{C-3}}$$

$I_{norm}$ = normalized peak intensity; $I_{C-x}$ = peak intensity of given carbon; $I_{C-3}$ = peak intensity of carbon C-3´(´)

The resulting peak enhancement (PE) after isotope feeding is defined as the ratio between normalized peak intensities obtained in labelled conditions and normalized peak intensities obtained in natural abundance conditions:

$$PE = \frac{I_{norm-labelled}}{I_{norm-natural\ abundacne}}$$

PE values > 1.8 were regarded as `incorporation events´, meaning, that the given carbon signal was derived from the labelled precursor. Examples of raw data and NMR spectra for incorporation of [1-$^{13}$C]-acetate **16a** into eupenifeldin **5** are listed and depicted in Table 3.4 and Figure 3.5 (for all other data see Appendix 8.2).

**Table 3.4** Summarized NMR data for eupenifeldin **5** labelled with [1-$^{13}$C]-acetate **16a** and non-labelled (natural abundance); norm. = normalized; PE = Peak enhancement; incorporation events are highlighted in green.

| | eupenifeldin (natural abundance) **5** | | | eupenifeldin (labelled) **5** | | | |
|---|---|---|---|---|---|---|---|
| **Pos.** | **$\delta_C$ / ppm** | **Intensity** | **norm to C-3´´** | **$\delta_C$ / ppm** | **Intensity** | **norm. to C-3´´** | **PE** |
| **1** | 41.6 | 177.9 | 0.07 | 41.6 | 207.39 | 0.08 | 1.14 |
| **2** | 80.5 | 3009.01 | 1.22 | 80.6 | 14527.71 | 575 | 4.70 |
| **3** | 46.3 | 1290.71 | 0.52 | 46.3 | 1324.97 | 0.52 | 1.00 |
| **4** | 125.9 | 429.39 | 0.17 | 125.9 | 3353.28 | 1.33 | 7.60 |
| **5** | 144.1 | 2023.44 | 0.82 | 144.1 | 1796.69 | 0.71 | 0.86 |
| **6** | 35.0 | 1760.37 | 0.72 | 35.0 | 10908.87 | 4.32 | 6.03 |
| **7** | 46.6 | 60625 | 0.25 | 46.6 | 4254.73 | 1.68 | 6.83 |
| **8** | 32.0 | 1775.99 | 0.72 | 32.0 | 1650.63 | 0.65 | 0.91 |
| **9** | 82.0 | 4204.58 | 1.71 | 82.0 | 15815.78 | 6.26 | 3.66 |
| **10** | 70.8 | 1839.77 | 0.75 | 70.8 | 1395.66 | 0.55 | 0.74 |
| **11** | 30.2 | 388.89 | 0.16 | 30.2 | 2997.62 | 1.19 | 7.51 |
| **12** | 19.4 | 3294.09 | 1.34 | 19.4 | 3045.26 | 1.20 | 0.90 |
| **13** | 29.7 | 1840.71 | 0.75 | 29.7 | 2112.96 | 0.84 | 1.12 |
| **14** | 27.3 | 415.84 | 0.17 | 27.2 | 556.93 | 0.22 | 1.30 |
| **15** | 16.1 | 1620.74 | 0.66 | 16.1 | 1644.76 | 0.65 | 0.99 |
| **1´** | 172.6 | 2288.98 | 0.93 | 172.5 | 4434.64 | 1.75 | 1.89 |
| **2´** | 163.2 | 2470.32 | 1.00 | 163.2 | 809.90 | 0.32 | 0.32 |
| **3´** | 113.5 | 2709.01 | 1.10 | 113.7 | 2326.78 | 0.92 | 0.84 |
| **4´** | 160.3 | 2369.71 | 0.96 | 160.4 | 6866.19 | 2.72 | 2.82 |
| **5´** | 122.5 | 824.15 | 0.33 | 122.7 | 418.02 | 0.17 | 0.49 |
| **6´** | 150.5 | 2521.18 | 1.02 | 150.6 | 7608.16 | 3.01 | 2.94 |
| **7´** | 124.7 | 2767.78 | 1.12 | 124.8 | 1545.07 | 0.61 | 0.54 |
| **8´** | 27.4 | 4551.74 | 1.85 | 27.5 | 3312.92 | 1.31 | 0.71 |
| **9´** | 33.0 | 1714.44 | 0.70 | 33.1 | 11775.21 | 4.66 | 6.69 |
| **1´´** | 173.3 | 2199.75 | 0.89 | 173.3 | 6660.18 | 2.63 | 2.95 |
| **2´´** | 162.9 | 2045.41 | 0.83 | 162.9 | 958.06 | 0.38 | 0.46 |
| **3´´** | 112.8 | 2461.56 | 1 | 112.9 | 2527.95 | 1 | 1 |
| **4´´** | 159.5 | 2288.51 | 0.93 | 159.5 | 8734.40 | 3.46 | 3.72 |
| **5´´** | 118.7 | 2301.47 | 0.93 | 118.7 | 1124.97 | 0.45 | 0.46 |
| **6´´** | 151.5 | 2774.16 | 1.13 | 151.5 | 10371.30 | 4.10 | 3.64 |
| **7´´** | 125.6 | 2102.25 | 0.85 | 125.7 | 1528.93 | 0.60 | 0.71 |
| **8´´** | 27.5 | 4567.07 | 1.86 | 27.5 | 3549.98 | 1.40 | 0.76 |
| **9´´** | 34.4 | 2074.85 | 0.84 | 34.4 | 13816.01 | 5.47 | 6.48 |

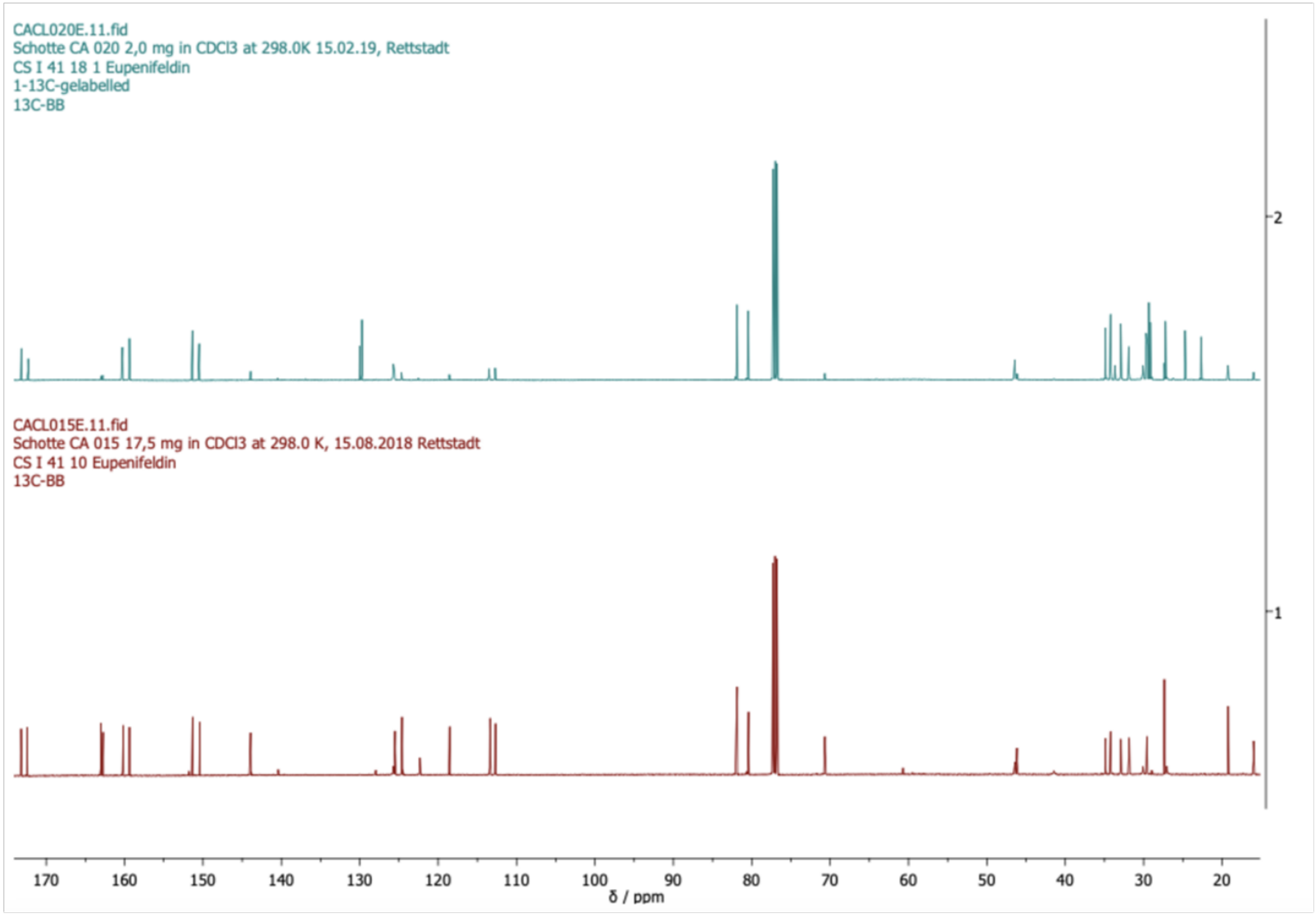
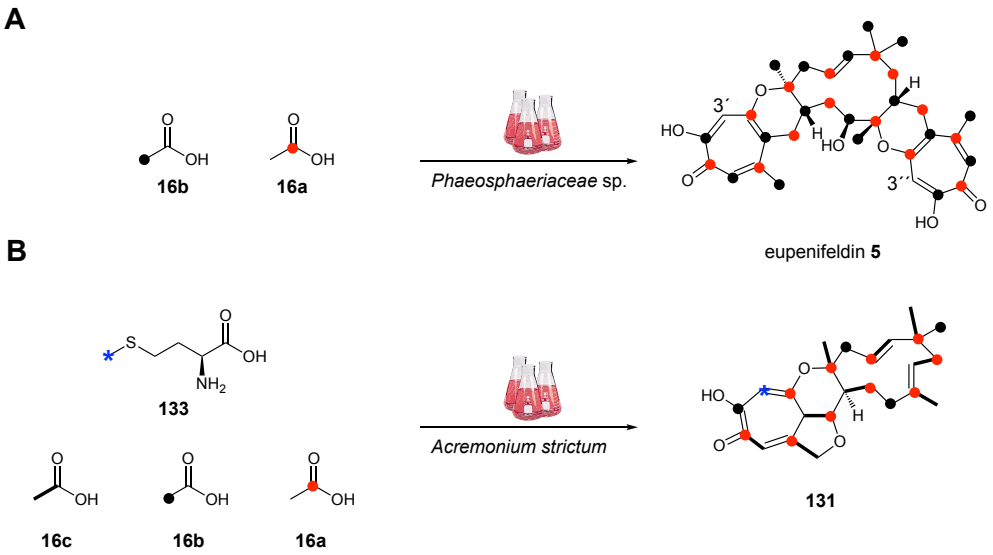



**Figure 3.5** $^{13}$C-NMR spectrum for [1-$^{13}$C]-labelled eupenifeldin **5** (top) compared to natural abundance $^{13}$C-NMR spectrum for **5** (bottom; 500 MHz in $CDCl_3$).

For eupenifeldin **5** 31/33 carbon signals were enhanced either upon feeding with [1-$^{13}$C]-acetate **16a** or [2-$^{13}$C]-acetate **16b** (Scheme 3.6 A). The only carbons not enhanced were carbon C-3′′ and C-3′. While C-3′′ was used for normalization, non-enhancement of the same carbon (C-3′) in the symmetrical second tropolone ring serves as an internal control, confirming that this position is not derived from acetate. The obtained pattern of labelling was identical to the observed pattern for xenovulene B **131** (previously reported by SIMPSON *et al.*; Scheme 3.6 B), both in respect to the tropolone moiety and to the macrocylic terpene core. Together, these findings are in agreement with the expected biosynthesis of eupenifeldin **5** from humulene **7** and stipitaldehyde **167**.

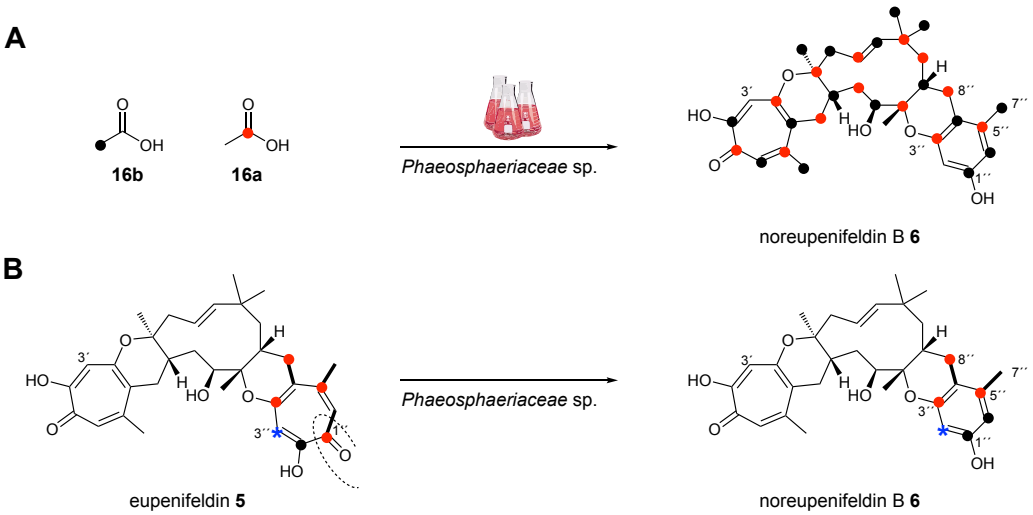



**Scheme 3.6** Isotopic labelling of tropolone sesquiterpenoids: **A,** isotopic labelling of eupenifeldin **5** with sodium [1-$^{13}$C, 2-$^{13}$C]-labelled acetate; **B,** isotopic labelling of xenovulene B **131** with sodium [1-$^{13}$C, 2-$^{13}$C, 1,2,$^{13}C_2$]-labelled acetate and [*methyl*-$^{13}$C]-labelled methionine as described by SIMPSON *et al.*[87]

For noreupenifeldin B **6** 30/32 carbon atoms were enhanced (Scheme 3.7 A). As expected, the observed pattern of label incorporation of both the tropolone ring and the humulene macrocycle was identical to the pattern observed for xenovulene B **131** and eupenifeldin **5**. In respect to the phenol, only two intact acetate units (C-7′′/C-5′′, C-4′′/C-8′′) were identified. Neighbouring carbons C-1′′ and C-6′′ are both derived from [2-$^{13}$C]-acetate **16b**, suggesting that a C-1 unit was excised in the course of **6** biosynthesis consistent with a ring-contraction pathway. Intriguingly, not all carbons within the phenol were found to be enhanced: the peak intensity of the C-2′′ carbon did not change upon feeding with either [1-$^{13}$C]-acetate **16a** or [2-$^{13}$C]-acetate **16b**. The non-acetate origin of carbon C-2′′ is in agreement with this position to be derived from methionine.[127]

Taken together these results eliminate *o*-orsellinaldehyde **204** as a possible precursor of **6**, as the observed labelling pattern contradicts this biosynthetic proposal. Instead, the observed labelling

suggests a ring-contraction from eupenifeldin **5** to noreupenifeldin B **6** that proceeds through excision of the C-1´´ carbon in **5** *via* an as-yet unknown mechanism (Scheme 3.7 B).



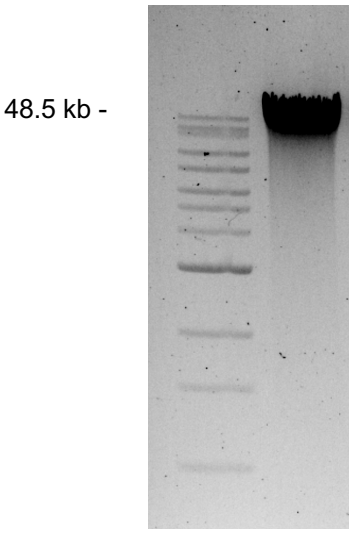

**Scheme 3.7** Isotopic labelling of noreupenifeldin B **6** and possible biosynthesis of **6**: **A,** isotopic labelling of **6** with sodium [1-$^{13}$C, 2-$^{13}$C]-labelled acetate; **B,** putative ring-contraction in *Phaeosphaeriaceae* sp. CF-150626 proceeding *via* excision of the C-1´´ carbon in eupenifeldin **5**.

## 3.4. Genome Sequencing of *Phaeosphaeriaceae* sp. CF-150626

To link the biosynthesis of **5** and **6** to a corresponding gene cluster we sequenced the genome of the producing organism. Genomic DNA (gDNA) of *Phaeosphaeriaceae* sp. CF-150626 was isolated from a 6 d culture in PD medium. For this purpose, fungal mycelia was obtained by Büchner filtration and the material was ground to a fine powder in liquid nitrogen. gDNA was isolated using the GenElute™ Plant Genomic DNA MiniPrep Kit (*Sigma Aldrich*). Combined multiple isolates were further purified using a sodium acetate DNA precipitation and purification protocol. High-molecular gDNA (> 40 kb) was submitted to DR. DANIEL WIBBERG (Centre for Biotechnology [CeBiTec] Bielefeld), who performed the whole genome sequencing and annotation (Figure 3.6 A). In brief, whole-genome-shotgun PCR-free libraries were constructed with the Nextera XT DNA Sample Preparation Kit (*Illumina*) and Illumina sequencing was achieved on the MiSeq platform (2 x 300 bp paired-end sequencing, v3 chemistry). The raw data was assembled with the GS De Novo Assembler software (version 2.8) and eucaryotic gene prediction was performed with the gene prediction tool Genemark v4.3.6.

A draft genome sequence of about 44.7 Mb was obtained. Key assembly statistics were evaluated using the quality assessement tool for genome assemblies (QUAST) and are displayed in Figure 3.6 B.[187] In total, 13702 genes were predicted on 659 scaffolds with an $N_{50}$ of 164515. While the number of scaffolds is high in comparison to the state-of-the art (< 20)[188] the genome quality was sufficient for the purposes of this study.

**A**




**B**

| | |
|---|---|
| Bases in scaffold (bp) | 44,718,648 |
| # of scaffolds | 659 |
| # of scaffolds (>= 50000 bp) | 246 |
| GC (%) | 48.39 |
| $N_{50}$ | 164515 |
| $L_{50}$ | 84 |
| # of predicted genes | 13702 |

**Figure 3.6** Whole genome sequencing of *Phaeosphaeriaceae* sp. CF-150626: **A,** agarose gel assessing the quality of gDNA; **B,** assembly statistics after whole genome sequencing; $N_{50}$ = half the genome size is accounted for by scaffolds >= this size; $L_{50}$ = number of scaffolds required to account for half the genome size.

FungiSMASH 5 beta was used to analyze the obtained genome sequence for the presence of BGC. FungiSMASH is the fungal version of the antibiotics and secondary metabolite analysis shell (antiSMASH), a comprehensive web server widely used in the natural product scientific community for the automated identification of biosynthetic gene clusters.[189,190] A total of 41 biosynthetic gene clusters were predicted that centered around either a core polyketide synthase gene, a non-ribosomal peptide synthetase gene or terpene cyclase gene (Figure 3.7). 12 clusters were directly linked to BGC of known function (as deposited with the minimum information about a biosynthetic gene cluster (MIBiG) server).[14] For example, clusters were linked to the biosynthesis of xenovulene A **1** (42% gene similarity), clavaric acid **208** (100% gene similarity), phyllostictine **209** (50% gene similarity) and emericellin **210** (57% gene similarity).

While under the given producing conditions none of the above-mentioned natural products were identified, the identified clusters represent the biosynthetic potential embedded in the *Phaeosphaeriaceae* genome. In future work a systematic analysis of *Phaeosphaeriaceae* sp. CF-150626 under different culture conditions may unveil further potential of strain CF-150626 to produce new secondary metabolites. As the xenovulene A **1** and eupenifeldin **5** biosynthetic gene clusters were expected to share a high degree of homology in the scope of this study only the cluster related to xenovulene was analyzed in detail.

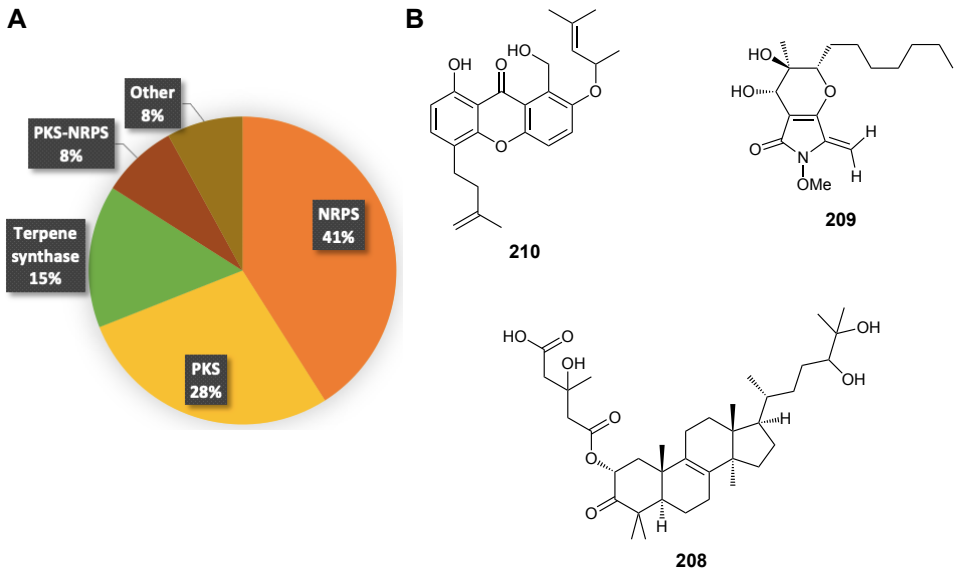



**Figure 3.7** FungiSMASH analysis of the *Phaeosphaeriaceae* genome: **A,** overview of identified clusters based on the cluster´s core gene; **B,** structures of key specialized metabolites linked to BGC in *Phaeosphaeriaceae*.

### 3.4.1. Bioinformatic Analysis of the *eup2* BGC

A detailed *in silico* analysis of the cluster predicted to be related to xenovulene A biosynthesis was performed (Figure 3.8; Table 3.5). Predicted open reading frames (ORF) spanning a ~ 36.6 kbp region around the central non-reducing PKS gene were manually annotated using the basic local alignment search tool (BLASTp) and NCBI´s conserved domain tool.[142,147] In the absence of transcriptomic data the 6-7 genes flanking the core PKS to the left (L) and right (R) were analysed. Analysis of genes more distant than gene L6 and R7 did not return any information that suggested these to be involved in secondary metabolism and they were – preliminarily – assigned to be outside the cluster.

The analysis revealed a central non-reducing PKS (EupPKS) putatively involved in the formation of 3-methylorcinaldehyde **164**. Additional identified ORFs encode: a putative FAD-dependent salicylate hydroxylase (EupL1); a non-heme iron dioxygenase (EupL5); a short-chain dehydrogenase (EupL4); a FAD-dependent monooxygenase (EupL5); a hetero DIELS-ALDERase (EupR1); a humulene synthase (EupR3); a cytochrome P450 (EupR6); two transporter proteins (EupL2, EupR7); two transcription factors (EupR2, EupR4); a putative mannitol-dehydrogenase (EupL6); and one hypothetical protein (EupL3). The biosynthetic potential embedded within the identified cluster was in agreement with the expected genes for tropolone sesquiterpenoid formation and hence we named the cluster *eup2* BGC.

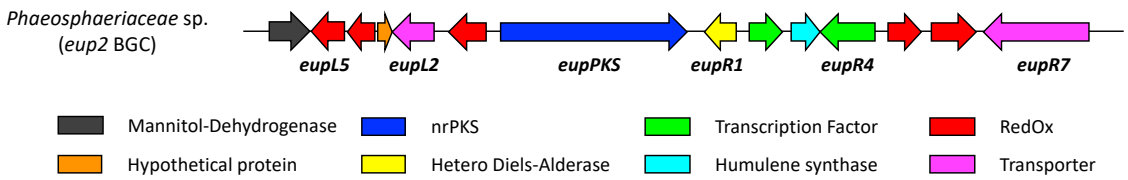



**Figure 3.8:** Schematic overview of the *eup2* biosynthetic gene cluster; genes are scaled in respect to actual size and distance from each other.

**Table 3.5** Annotation of the putative *eup2* BGC; proposed function identified with BLASTp and NCBI´s conserved domain tool.[142,147]

| # | *gene*/protein | bp/aa | proposed function |
|---|---|---|---|
| CF150626g2125.t1 | *eupL6*/EupL6 | 1707/568 | mannitol-dehydrogenase |
| CF150626g2126.t1 | *eupL5*/EupL5 | 1113/334 | non-heme iron dioxygenase |
| CF150626g2127.t1 | *eupL4*/EupL4 | 966/278 | short-chain dehydrogenase |
| CF150626g2128.t1 | *eupL3*/EupL3 | 407/114 | ycil-like protein |
| CF150626g2129.t1 | *eupL2*/EupL2 | 1728/512 | monocarboxylate transporter |
| CF150626g2130.t1 | *eupL1*/EupL1 | 1607/447 | FAD-dependent salicylate hydroxylase |
| CF150626g2131.t1 | *eupPKS*/EupPKS | 8252/2661 | 3-methylorcinaldehydesynthase |
| CF150626g2132.t1 | *eupR1*/EupR1 | 1149/382 | hetero DIELS-ALDERase |
| CF150626g2133.t1 | *eupR2*/EupR2 | 1281/408 | C6 finger domain transcription factor |
| CF150626g2134.t1 | *eupR3*/EupR3 | 1155/384 | humulene synthase |
| CF150626g2135.t1 | *eupR4*/EupR4 | 2315/753 | C6 finger domain transcription facter |
| CF150626g2136.t1 | *eupR5*/EupR5 | 1311/436 | FAD-dependent monooxygenase |
| CF150626g2137.t1 | *eupR6*/EupR6 | 1882/537 | cytochrome P450 |
| CF150626g2138.t1 | *eupR7*/EupR7 | 4722/1466 | ABC-transporter |

To assign putative biosynthetic functions to individual genes in the *eup2* BGC we compared the cluster with the known xenovulene A *asPKS1* BGC and the eupenifeldin *eup* BGC (NCBI accession number: MK4001201.1).[80] The CLINKER & CLUSTERMAP.JS tools were used to calculate a global alignment between the (protein) sequences and to visualize the calculated homologies (Figure 3.9).[143] The homology analysis revealed a very high degree of homology between the *eup2* and *eup* cluster, suggesting these to be very closely related. For every gene within the *eup2* BGC there is a homologous gene within the *eup* BGC encoded at the same position and in the same orientation in regard to the central NR PKS. Global sequence identities of the translated protein sequences were determined to be > 70 % on average (EMBOSS needle).[164]

Comparison with the *asPKS1* xenovulene BGC revealed a lower degree of homology, both in terms of global sequence identities but also in regard to position and orientation of biosynthetic genes relative to the core PKS gene. However, key biosynthetic genes involved in tropolone formation (*eupPKS/asPKS1*, *eupL1/asL1*, *eupL5/asL3*) as well as humulene synthase encoding gene (*eupR3/asR6*) and hetero DIELS-ALDERase encoding gene (*eupR1*/*asR5*) are conserved in both clusters, with sequence identities between 50 - 90 %. Notably, *asR2*, essential for the formation of the tetrahydrofuran ring in xenovulene A **1**, is not conserved in the *eup2* BGC. The two oxidoreductase encoding genes *asL4* and *asL6* have a single homologue in the *eup2* BGC, *eupR5* (34.8 % and 31.8 % sequence identity, respectively). Additionally, the short-chain

dehydrogenase encoding gene *eupL4*/*eupG* and cytochrome P450 encoding gene *eupR6*/*eupH* are unique to the two eupenifeldin biosynthetic gene clusters.

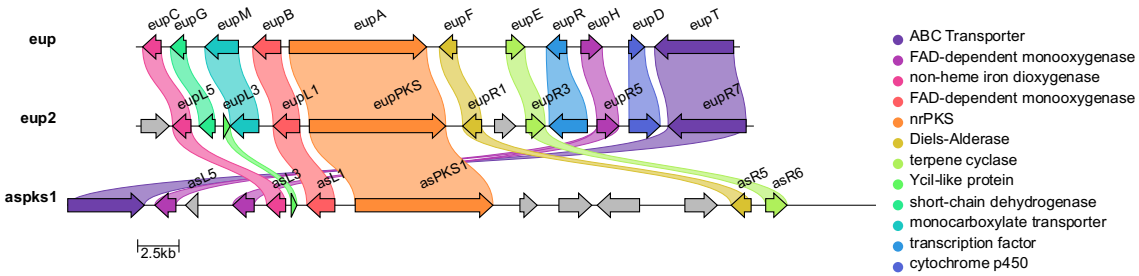
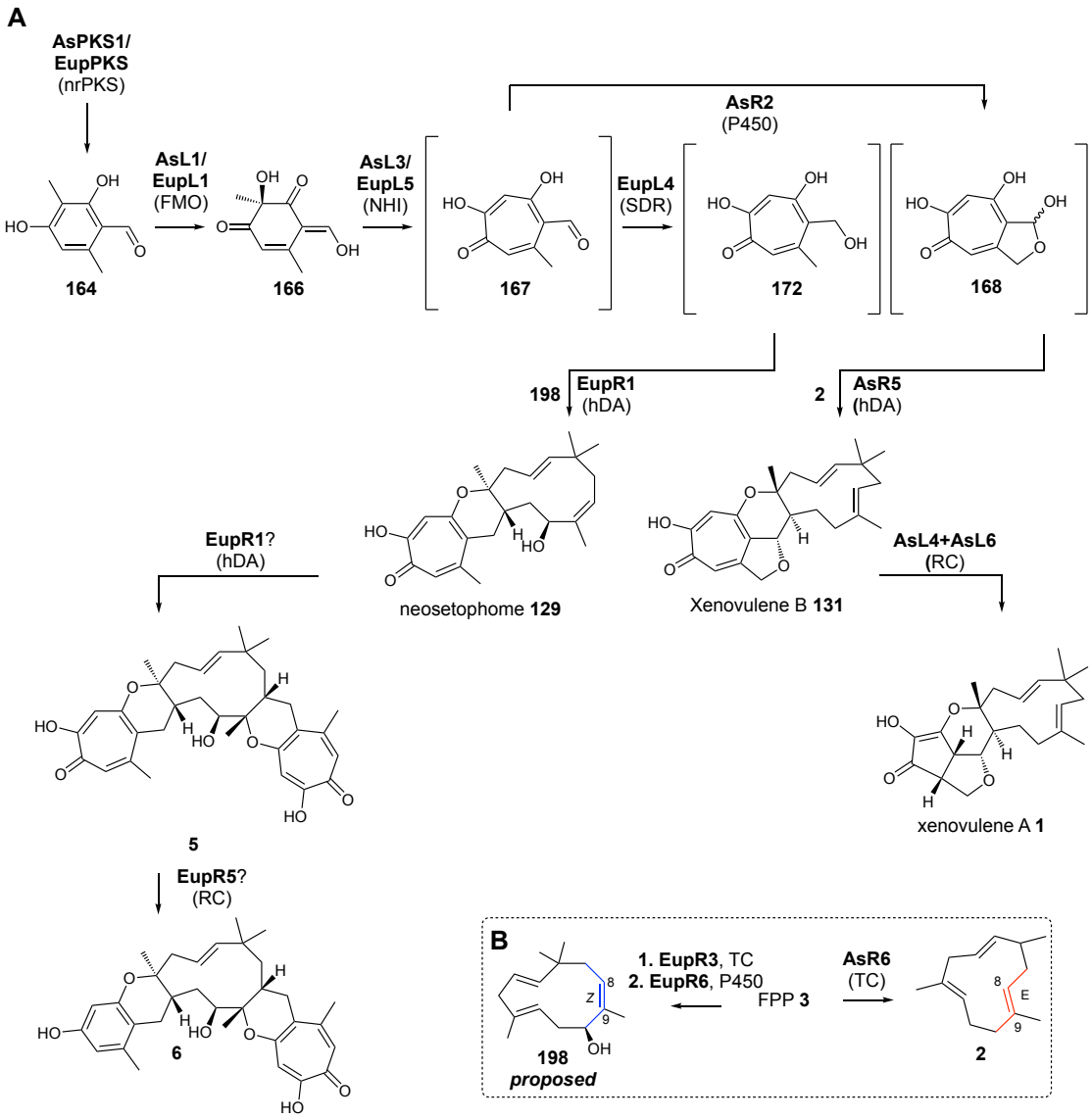



**Figure 3.9** Homology comparison between the putative *eup2* BGC from *Phaeosphaeriaceae* sp., the *eup* BGC from *Phoma* sp. and the *asPKS1* BGC from *A. strictum*; CLINKER & CLUSTERMAPS.JS was used to calculate and visualize the homology.[143]

Based on the data obtained from isotopic labelling, genome sequencing and including the results published by other groups in the course of this work a putative pathway towards eupenifeldin **5** and noreupenifeldin B **6** in *Phaeosphaeriaceae* sp. was devised (Scheme 3.8). The early biosynthetic steps are identical to the biosynthesis of xenovulene A **1** and proceed *via* formation of stipitaldeyhde **167** through the action of the non-reducing PKS EupPKS, the FAD-dependent monooxygenase EupL1 and the non-heme iron dioxygenase EupL5. Stipitaldeyhde **167** constitutes the point of divergence in the biosynthesis of **1** and **5**. In the case of xenovulene A **1**, **167** is *oxidized* at the C-6-methyl group and the characteristic tetrahydrofuran moiety in **1** is formed through intramolecular hemiacetal formation. In contrast, during eupenifeldin **5** biosynthesis stipitaldehyde **167** is *reduced* to the corresponding alcohol stipitol **172**. Then the hetero DIELS-ALDERases AsR5 and EupR1 catalyze *o*-quinomethide (**170** and **174**) formation through dehydration of either **168** or **172**. The subsequent DIELS-ALDER reaction with humulene furnishes the core carbon skeleton of **1** and **5**. Notably, in eupenifeldin **5** biosynthesis we would expect 8*Z*-humulenol **198** instead of α-humulene **2** to react and two DIELS-ALDER reactions must occur.

In the case of xenovulene A **1**, oxidative ring-contractions lead to **1**-formation. Based on the labelling pattern of noreupenifeldin B **6** we expect a similar ring-contraction to occur during **6** formation. Homology of EupR5 to the known ring-contraction enzymes AsL4 and AsL6 marks this enzyme as a likely candidate catalyst.




**Scheme 3.8: A,** Proposed biosynthetic route towards eupenifeldin **5,** noreupenifeldin B **6** and xenovulene A **1**; nrPKS = non-reducing polyketide synthase; FMO = FAD-dependent monooxygenase; NHI = non-heme iron dioxygenase; SDR = short-chain dehydrogenase; P450 = cytochrome P450; hDA = hetero DIELS-ALDERase; TC = terpene cyclase; RC = ring-contraction enzyme; **B,** proposed humulene formation in the respective biosynthetic pathways.

## 3.5. *In vitro* Characterisation of the Terpene Cyclase EupR3

Systematic comparison of the *eup2*, *eup* and *asPKS1* BGC identified the *eupR3* gene as a humulene synthase encoding gene, putatively responsible for the formation of 8*Z*-humulene **7**.

In an initial bioinformatic assessment we compared EupR3 (43.3 kDa) with the other known humulene synthases AsR6 (48.4 kDa), PycR6 (44.1 kDa) and EupE[80] (43.2 kDa). With AsR6 and PycR6 two α-humulene synthases were included in the comparison, whereas EupR3 and EupE were expected to give the structural isomer 8*Z*-humulene. The difference in stereochemistry should be mirrored in differences in the amino acid sequence of the enzymes. However, a global sequence alignment of the four enzymes using the clustal omega multiple sequence alignment

tool[164] (Figure 3.11) revealed a high overall sequence identity between all four enzymes. None of the enzymes appears to possess the aspartate-rich $Mg^{2+}$-binding motifs found in all other known type I terpene cyclases.

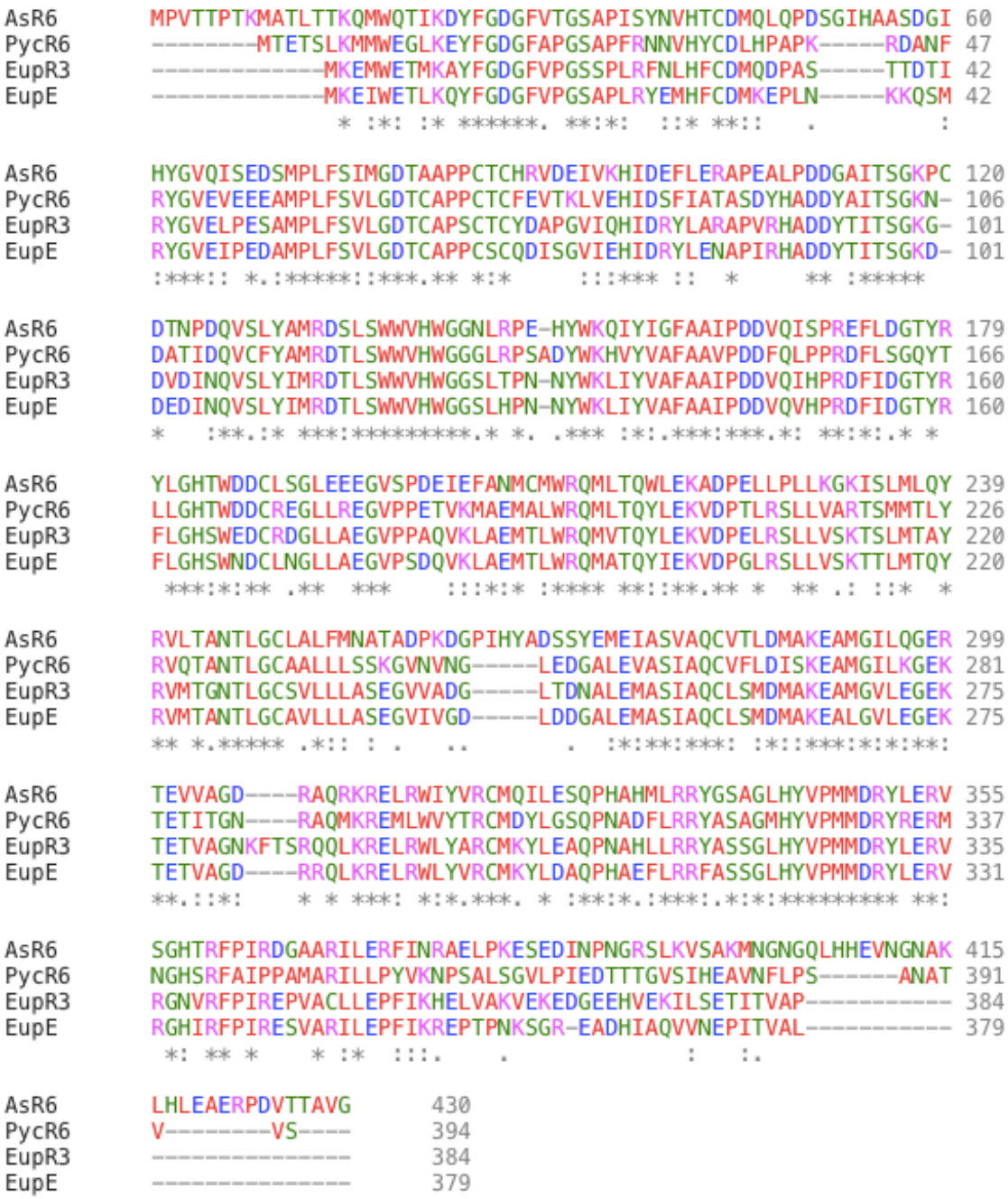


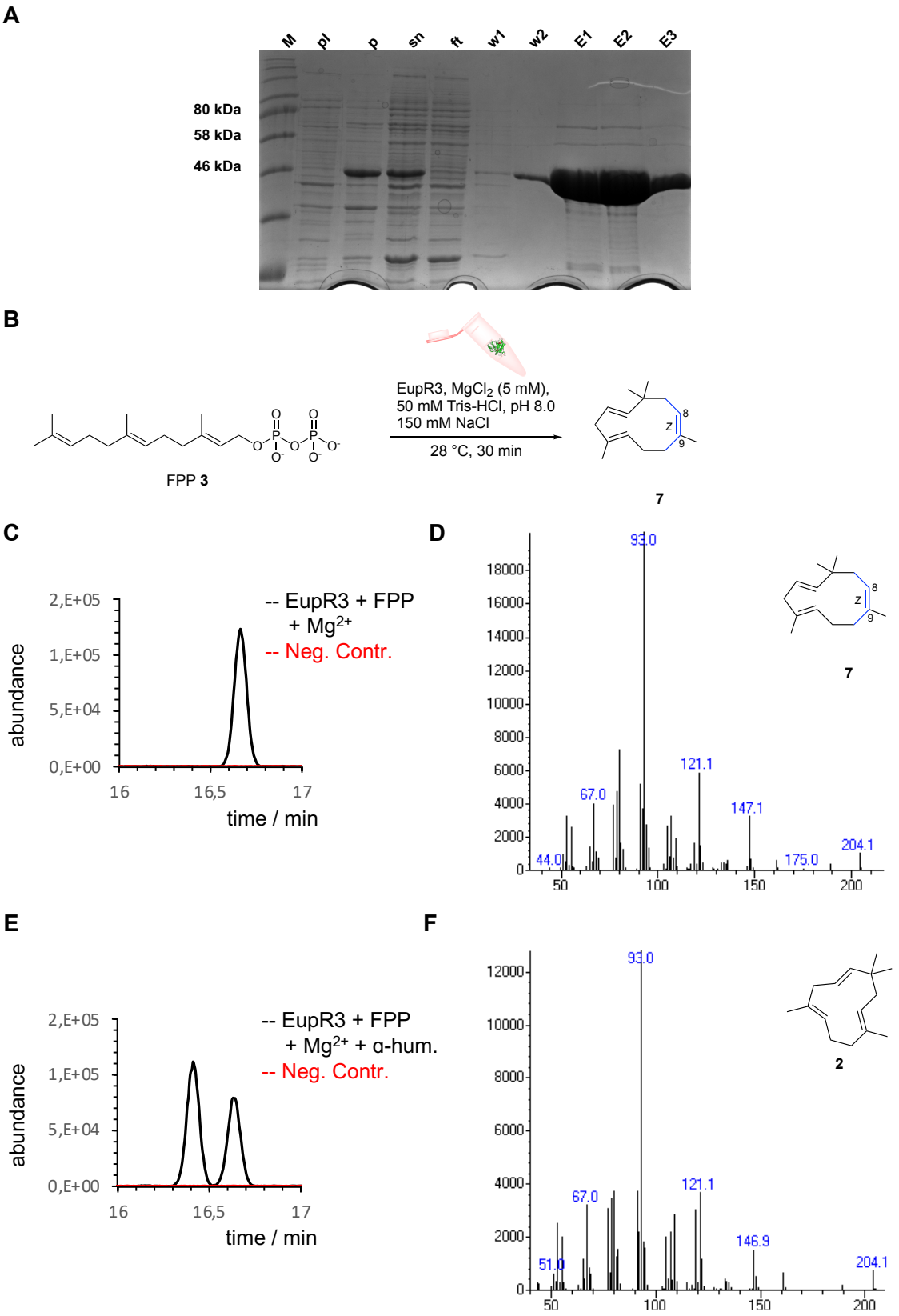

**Figure 3.11** Global sequence alignment of all known humulene synthases.

Manual comparison of sequence identities using EMBOSS Needle[164] (Table 3.6) confirmed that the differences between AsR6/PycR6 and EupR3/EupE were small, with sequence identities between 47 - 77%. Indeed, AsR6 shared higher sequence identity with EupR3 than with PycR6. Based on this preliminary assessment, the differences in product stereochemistry must originate in minor – not obvious – differences in the protein sequence.

**Table 3.6** Sequence comparison of all known AsR6-like humulene synthases. Displayed are global sequence identities.

| | AsR6 | PycR6 | EupR3 | EupE |
|---|---|---|---|---|
| **AsR6** | - | 47.3% | 49.4% | 51.6% |
| **PycR6** | - | - | 58.4% | 57.8% |
| **EupR3** | - | - | - | 77.2% |
| **EupE** | - | - | - | - |

To validate EupR3 as an 8*Z*-humulene synthase, codon-optimized *eupR3* with an *N*-terminally fused $His_6$-Tag was expressed in *E. coli* BL21. Induction with 0.1 mM IPTG and expression at 16 °C for 16-20 h proved best suited to obtain soluble EupR3. EupR3 was isolated using standard IMAC-technology ($Ni^{2+}$ affinity chromatography; Scheme 3.9 A). Aliquots of EupR3 containing 10 % glycerol were shock-frozen in liquid nitrogen and stored at - 80 °C for up to 3 months. The *in vitro* activity of EupR3 was assessed as previously described for AsR6 (Scheme 3.9 B). Incubation of EupR3 with FPP **3** and $Mg^{2+}$ followed by extraction with n-hexane and GCMS analysis identified a single product that was not observed in the negative control lacking EupR3 (Scheme 3.9 C + D). The product had the same molecular mass (*m/z* 204) as α-humulene **2**, but the retention time ($t_R$ = 16.65 min) was clearly different when compared to previous assays with AsR6, where α-humulene **2** eluted at $t_R$ = 16.40 min. To ensure that EupR3 produced a different product to AsR6, co-injection experiments were performed using a commercially obtained standard of α-humulene **2** (*Sigma Aldrich*). After conduction of the EupR3 *in vitro* assay the n-hexane extract was mixed with 0.02 $mg \cdot mL^{-1}$ α-humulene **2** and analysed by GCMS (Scheme 3.9 E + F). The retention times of α-humulene **2** ($t_R$ = 16.4 min) and the EupR3 terpene product ($t_R$ = 16.6 min) were clearly different, demonstrating that EupR3 does not produce α-humulene **2**. However, the EI mass spectra were nearly identical, in agreement with our expectation that EupR3 produces a stereoisomer of α-humulene **2**.




**Scheme 3.9** Characterisation of EupR3: **A,** heterologous expression of *eupR3* in *E. coli* BL21 and purification monitored by SDS PAGE; M = marker (*NEB* Color Prestained Protein Standard, Broad Range (11 - 245 kDa); pI = preinduction; p = pellet; sn = total lysate; ft = flowthrough; w1 = wash step with lysis buffer; w2 = wash step with 50 mM imidazole; E1-E3 = elution fractions (500 mM imidazole); **B**, *in vitro* assay of EupR3; **C,** GCMS (TIC) chromatogram of n-hexane extract of assay containing EupR3; **D,** EI mass spectrum of 8*Z*-humulene **7**; **E,** GCMS (TIC) chromatogram of co-injection of EupR3 *in vitro* assay and α-humulene standard **2**; **F,** EI mass spectrum of α-humulene standard **2**.

To identify the product of EupR3 a biotransformation of FPP **3** by EupR3 was performed. For this, a 50 ml *in vitro* reaction was set-up. To assay-buffer (40 mL) containing 8.6 mg EupR3, FPP **3** was added stepwise (30 mg at $t_0$; 30 mg at $t_{1.5h}$; 20 mg at $t_{3.oh}$). After 6 h the reaction was quenched through addition of 50 ml n-pentane. For total extraction of the unknown product the mixture was stirred overnight. Then the n-pentane was carefully removed, first at atmospheric pressure and 40 °C and finally the last 5 mL under a light continuous stream of nitrogen. After complete removal of the organic solvent no visible or weighable residue remained. Traces (~ < 1 mg) were directly dissolved in NMR solvent (benzene-$d_6$) and analysed by NMR. Full structure elucidation through combined 1D- and 2D-NMR data rapidly confirmed the humulene core structure (Figure 3.11 and Table 3.7). Comparison of the data with a literature report of synthetic 8*Z*-humulene further supported the structure of **7**.[191]

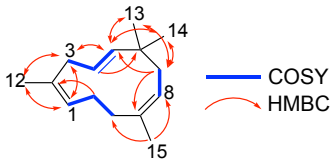



**Figure 3.11:** Chemical characterization of 8*Z*-humulene **7** by NMR spectroscopy (key COSY and HMBC correlations highlighted).

**Table 3.7** NMR data for 8*Z*-humulene **7** in benzene-$d_6$ (600 MHz). Referenced to benzene-$d_6$.

| # | $\delta_H$ / ppm (mult, *J* in Hz) | $\delta_C$ / ppm | C-type | HMBC (H to C) | NOESY |
|---|---|---|---|---|---|
| **1** | 5.05 (m) | 125.3 CH | CH | 3, 12 | 3, 5, 12, 13, 14, |
| **2** | - | 139.7 C | C | - | |
| **3** | 2.53 (d, 7.0) | 41.7 $CH_2$ | $CH_2$ | 1, 2, 4, 5, 12 | 4, 5, 8, 12 |
| **4** | 5.57 (dt, 15.8, 7.0) | 126.5 CH | CH | 3, 5, 6 | 3, 12, 13, 14 |
| **5** | 5.36 (m) | 142.8 CH | CH | 3, 6, 13, 14 | 1, 3, 13, 14 |
| **6** | - | 37.8 C | C | - | |
| **7a** | 1.72 (m) | 43.5 $CH_2$ | $CH_2$ | 5, 8, 9 | |
| **7b** | 1.92 (t, 12.7) | | | 5, 6, 8, 9, 13, 14 | 5, 11a, 13, 14 |
| **8** | 5.24 (m) | 123.4 CH | CH | | 13, 15 |
| **9** | - | 134.8 C | C | - | |
| **10a** | 1.68 (d) | 31.7 $CH_2$ | $CH_2$ | 11 | |
| **10b** | 2.11 (dd, 12.4, 2.9) | | | - | 7b, 12, 13, 14, 15 |
| **11a** | 2.06 (m) | 25.8 $CH_2$ | $CH_2$ | - | 1, 5, 7b, 12, 13, 14, 15 |
| **11b** | 2.21 (m) | | | 1, 2, 10 | 11a, 12, 15 |
| **12** | 1.51 (d, 1.5) | 17.5 $CH_3$ | $CH_3$ | 1, 2, 3 | 3, 4, 5, 15 |
| **13** | 1.05 (s) | 24.1 $CH_3$ | $CH_3$ | 5, 6, 7, 14 | 4, 5, 7b, 8 |
| **14** | 1.06 (s) | 28.8 $CH_3$ | $CH_3$ | 5, 6, 7, 13 | 4, 5, 7b |
| **15** | 1.64 (s) | 23.7 $CH_3$ | $CH_3$ | 8, 9, 10 | 8, 12, 13, 14 |

The alkene configuration of **7** was assessed through 1D NOESY NMR spectroscopy. Among others, the observed key nOe correlation of $CH_3$-15 to CH-8 is in agreement with a C-8/C-9 *Z*-configured olefin. Notably, in α-humulene **2** this correlation should be missing. Indeed, upon obtaining 1D NOESY NMR data for a standard solution of α-humulene **2** in benzene-$d_6$ the particular nOe correlation of $CH_3$-15 to CH-8 is indeed absent. Instead, $CH_3$-15 distinctly correlates with $CH_2$-7, in agreement with an *E*-configured C-8/C-9 alkene (Table 3.8 and Figure 3.13).

**Table 3.8** NMR data for α-humulene **2** in benzene-$d_6$ (600 MHz). Referenced to benzene-$d_6$.

| # | $\delta_H$ (mult, *J* in Hz) | NOESY |
|---|---|---|
| 1 | 4.96 (dd, 10.4, 6.6) | |
| 2 | - | |
| 3 | 2.53 (d, 7.4) | 8, 12 |
| 4 | 5.58 (dt, 15.9, 7.4) | 13/14 |
| 5 | 5.15 (dt, 15.8, 0.9) | 7a/b |
| 6 | - | |
| 7a | 1.96 (d, 7.6) | 13/14, 15 |
| 7b | 1.96 (d, 7.6) | |
| 8 | 4.92 (m) | 13/14 |
| 9 | - | |
| 10a | 2.05-2.08 (m) | |
| 10b | 2.05-2.08 (m) | |
| 11a | 2.05-2.08 (m) | |
| 11b | 2.05-2.08 (m) | |
| 12 | 1.56 (d, 1.5) | 3 |
| 13 | 1.07 (s) | |
| 14 | 1.07 (s) | |
| 15 | 1.38 (m) | 7 |

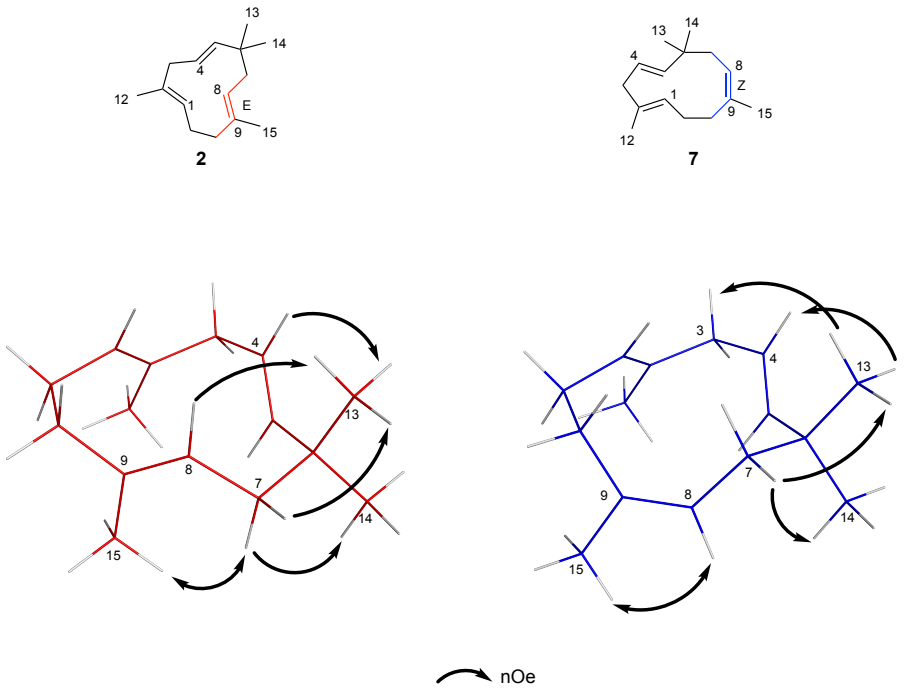



**Figure 3.13** Key nOe-correlations for α-humulene **2** and 8*Z*-humulene **7**: 3D model structures of **2** and **7** were calculated using Spartan 18 and minimised using molecular mechanics.

## 3.6. *In vitro* Characterisation of EupR5

Previous isotopic feeding studies in *Phaeosphaeriaceae* sp. CF-150626 had revealed that the monobenzopyranyl moiety in noreupenifeldin B **6** results from a ring-contraction of a tropolone precursor (Section 3.3.3). The *eup2* BGC was analysed for potential gene candidates that might catalyse such reactions. Since ring-contractions in natural product biosynthesis are often catalysed by enzymes belonging to the class of oxidoreductases (*e.g.,* ZopK during zopfiellin biosynthesis; ForY during formicamycin biosynthesis) we analysed all oxidoreductase genes within the *eup2* BGC.[192,193] The *eup2* BGC harbours five oxidoreductase encoding genes: Biosynthetic roles have been assigned to four of these genes, which are essential for formation of stipitol **172** and

humulenol **198**. The remaining gene candidate, *eupR5*, encoding an FAD-dependent monooxygenase, shows weak sequence homology to the known ring-contraction enzyme encoding genes *asL4* and *asL6*. In the absence of other suitable gene candidates within the *eup2* BGC, *eupR5* was investigated more closely.

To determine the putative biosynthetic function of EupR5 the *eupR5* gene was cloned from cDNA into the expression plasmid pET28a *via* the *Nde*I/*Not*I restriction sites. The successful generation of the expression plasmid pET28a_*eupR5* (Labbook ID: CS I 78 1-13) was confirmed by DNA sequencing (*Eurofins*, Ebersberg). A range of expression conditions varying in the expression temperature (12 °C, 16 °C, ambient temperature) and IPTG concentration (0.1 mM; 1.0 mM) were probed for the production of soluble EupR5 (expected apparent molecular weight: 50.3 kDa; Figure 3.14). Comparison to a preinduction control by SDS-PAGE identified a band corresponding to a protein that was best produced at 16 °C and 1 mM IPTG, that approximately displayed the correct apparent molecular weight. However, attempts to purify the protein from the total lysate *via* IMAC ($Ni^{2+}$ affinity chromatography) failed, as the protein did not bind to the column matrix. Non-binding of the protein to the column matrix may be attributed to either this protein not being EupR5 or the $His_6$-tag being inaccessible and thus incapable of interaction with the column matrix. In absence of purified enzyme, a cell-free extract (total lysate) of the expressed construct was used to probe for any ring-contraction activity. For this, an *in vitro* assay containing eupenifeldin **5** (15 mM), FAD (1 mM) and NADH (1 mM) was conducted at 30 °C for up to 16 h, but no conversion of eupenifeldin **5** was ever observed (data not shown).

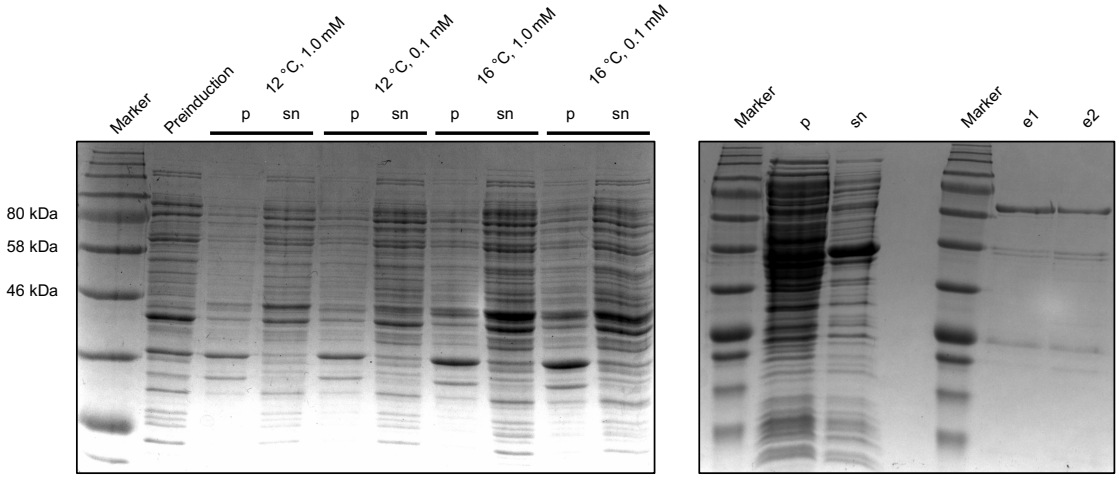



**Figure 3.14** Heterologous expression and attempted purification of *eupR5*/EupR5: SDS-PAGE of key test expression experiments (left); p = pellet; sn = total lysate; SDS PAGE of attempted $Ni^{2+}$ affinity chromatography (right); e1 = elution with 1 ml 500 mM imidazole; e2 = elution with 1 ml 500 mM imidazole; marker = *NEB* Color Prestained Protein Standard, Broad Range (11 - 245 kDa).

As previously observed for AsL4 and AsL6 a putative transmembrane domain was identified *in silico* in the EupR5 protein sequence, that might interfere with correct protein folding. In absence

of soluble preparations of AsL4, AsL6 and EupR5 it was concluded that the (putative) ring-contraction enzymes are not suited for heterologous expression in *E. coli.*

## 3.7. Conclusion, Discussion and Outlook

In this part of the study a prolific producer of eupenifeldin **5** and noreupenifeldin B **6** (*Phaeosphaeriaceae* sp. CF-150626) was characterized on a genetic, chemical and biochemical level. Genome sequencing of the fungus afforded a draft genome sequence that allowed identification of a biosynthetic gene cluster responsible for **5** and **6** biosynthesis (the *eup2* BGC). The cluster showed very high homology to the previously described *eup* BGC (encoded in *Phoma* sp.) that is also responsible for the biosynthesis of **5**.[80] Based on homology comparisons a plausible biosynthetic pathway for the formation of **5** in *Phaeosphaeriaceae* sp. was suggested.

Eupenifeldin **5**, dehydroxyeupenifeldin **207** and noreupenifeldin B **6** share the conserved feature of an eastern *cis*-configured humulene/dihydropyran ring-junction, that was proposed to be derived from 8*Z*-humulene **7**.[175] *In vitro* investigations in this study demonstrated that EupR3 is responsible for the formation of **7**. This contrasts the production of all-*E* α-humulene **2** by the homologous enzymes AsR6 and PycR6. The latter two give rise to the xenovulene-type sesquiterpenoids that harbour an 8*E*-olefin or *trans*-configured substituents in the final pathway products (compare Section 2). It is important to emphasize that all known humulene synthases share a high degree of homology (> 50%), independent of whether they afford **2** or **7** as the terpene cyclase product. Preliminary bioinformatic comparisons of the two types of enzymes did not reveal obvious conserved changes that can explain the differences in stereoselectivity. To identify factors that contribute to the unusual stereoselective fidelity structural data for the two cyclases is required.

Identification of **7** validates the previously proposed classification of tropolone sesquiterpenoids into xenovulene-type and eupenifeldin-type metabolites. The observed eastern *cis*-configuration in **5** and **6** is mirrored in an eastern 8*Z*-configured alkene in monotropolone derivatives such as neosetophome B **129** and the observed stereochemistry is consistent in all members of this subclass. Besides the aforementioned compounds the eupenifeldin-type family of tropolone sesquiterpenoids includes phomanolide A-F **211**-**215** and neosetophome A **216** (Figure 3.15).[81]

The insights gained in this study into the biosynthesis of xenovulene- and eupenifeldin-type tropolone sesquiterpenoids pave the way for the engineering of these pathways. With the *asPKS1*, the *pyc* and the *eup2* BGC three gene clusters have been identified that can serve as a toolbox for the reconstruction of artificial pathways in a heterologous host such as *Aspergillus oryzae*. A possible strategy includes the incorporation of eupenifeldin-type biosynthetic enzymes into the

heterologous expression system established for the xenovulenes (Section 2). With a range of possible tailoring enzymes (*e.g.,* humulene hydroxylase EupR6, FAD-dependent monooxygenase EupR5, short-chain dehydrogenase EupL4) the *eup2* BGC harbours a range of intriguing tailoring enzymes that can be rationally included in the xenovulene biosynthetic pathway.

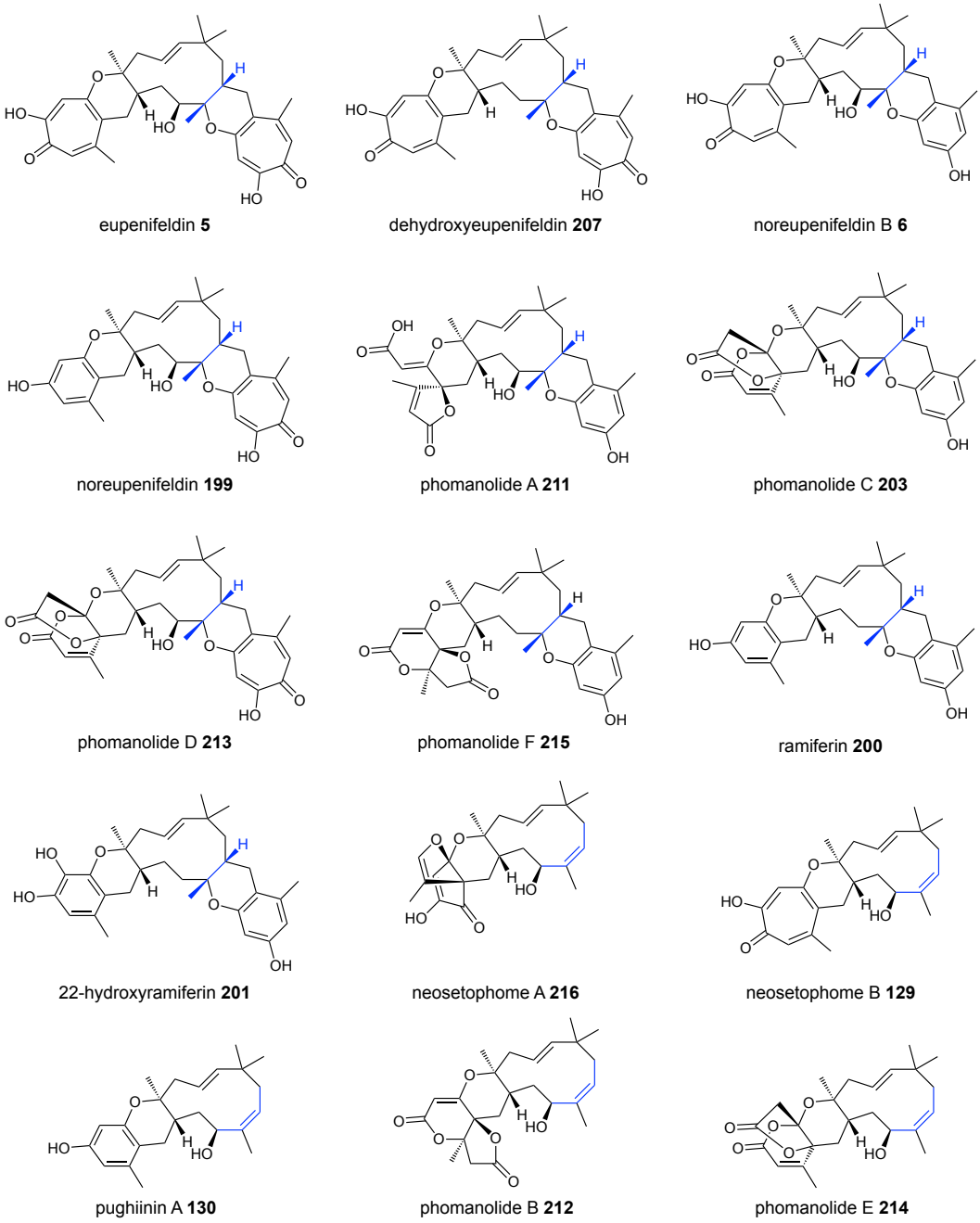



**Figure 3.15:** The eupenifeldin-type family of tropolone sesquiterpenoids.

A particular interest was directed to the biosynthetic origin of the tetrasubstituted benzene rings observed in some tropolone sesquiterpenoids (*e.g.,* **6**, **199** and **200**; Figure 3.15). Isotopic labelling studies using [1-$^{13}$C]- and [2-$^{13}$C]-acetate **16ab** were deployed in this study to address this question.

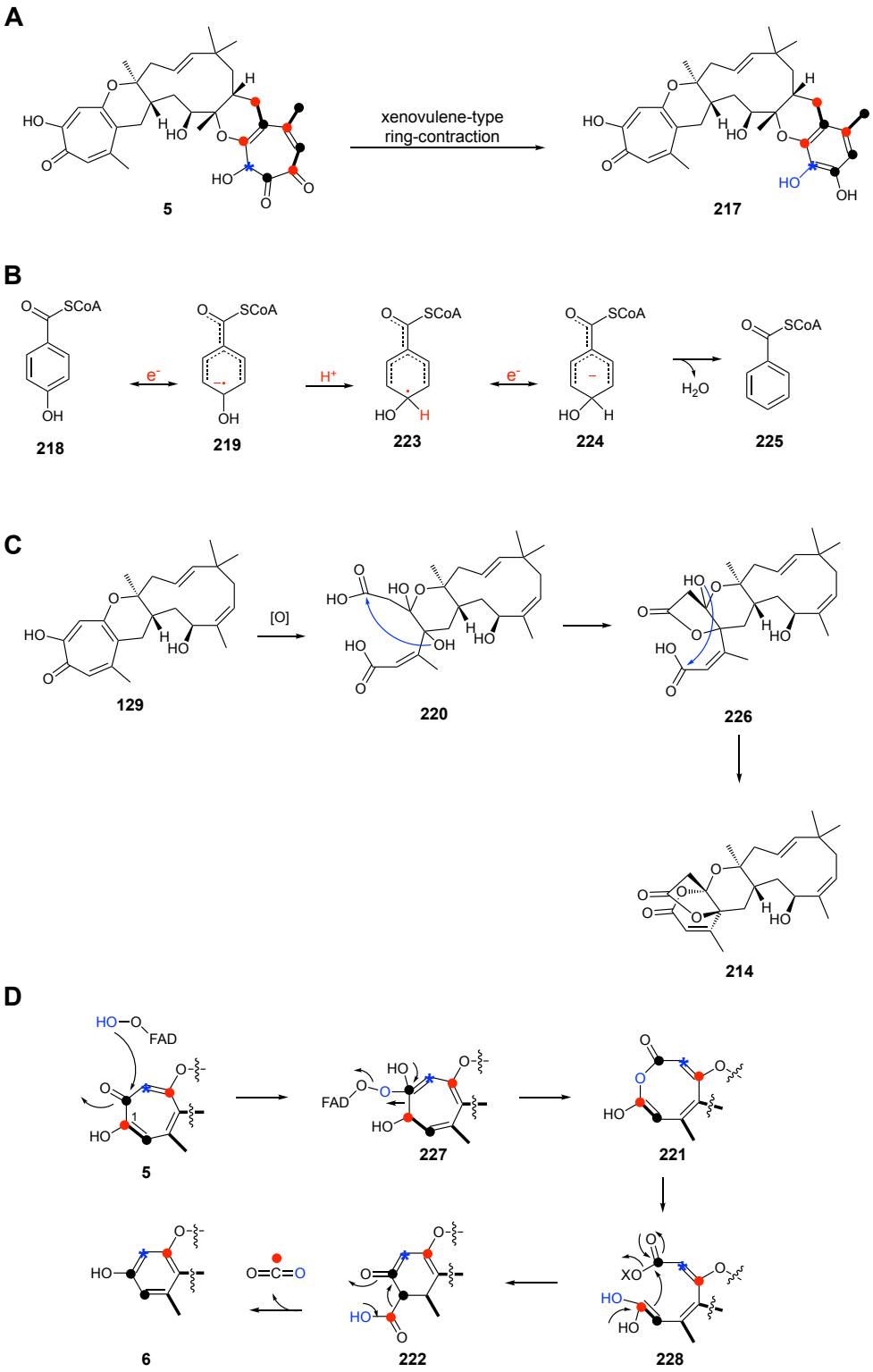
Results of experiments confirmed the meroterpenoid origin of eupenifeldin **5** and noreupenifeldin B **6**. As expected, the observed labelling pattern of the tropolone nucleus and the humulene macrocycle was identical to the pattern described for xenovulene B **131**. Analysis of the labelling pattern of the six-membered ring confirmed that noreupenifeldin B **6** is derived from an as-yet unknown oxidative ring-contraction. Eupenifeldin **5** is the likely biosynthetic precursor.

Interestingly the mechanism of the ring-contraction must differ from the ring-contractions catalysed by AsL4 and AsL6 during the biosynthesis of xenovulene A **1**. The latter mechanism would result in the retention of oxygen on the ring-contracted six-membered ring; however, no evidence for such a hypothetical intermediate **217**, was ever observed (Scheme 3.10 A).

In the case of a xenovulene-type ring-contraction in *Phaeosphaeriaceae* sp. CF-150626, an additional reductive step (a formal aromatic dehydroxylation) would be required. Evidence for enzymatically catalysed aromatic dehydroxylations is rarely reported in the literature. *E.g.*, the enzyme 4-hydroxybenzoyl-CoA reductase was proposed to catalyse the aromatic dehydroxylation of 4-hydroxybenzoyl-CoA **218** in a BIRCH-like fashion.[194] In a first step a one electron transfer generates a radical anion intermediate **219**. Protonation and a second one electron transfer step then lead to the elimination of water (Scheme 3.10 B).[195] However, in tropolone sesquiterpenoids the carbonyl oxygen present in **218**, which contributes to the stabilisation of the radical anion,[195] is absent – additionally, no obvious enzyme candidate catalysing such a reaction is conserved in the known tropolone sesquiterpenoid biosynthetic gene clusters.

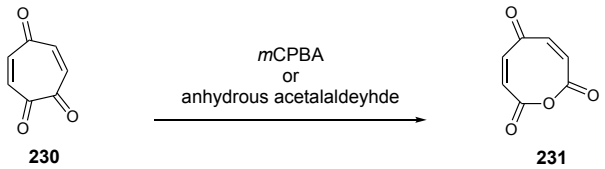
Alternatively, the observed ring-contraction from eupenifeldin **5** to noreupenifeldin B **6** might proceed through a different mechanism. CHE *et al*. first proposed a ring-opening ring-closing mechanism for the biosynthesis of phomanolide E **214**: Oxidative cleavage of the tropolone nucleus in **129** was proposed to give rise to the ring-opened intermediate **220**. Successive intra-nucleophilic attacks would establish the four- and six-membered rings observed in **214** (Scheme 3.10 C).[177]

For noreupenifeldin B **6**-formation, a similar mechanism might be in place: Initial hydroxylation at the C-2-position in **5** could trigger an oxidative ring-expansion to the eight-membered intermediate **221**. Ring-opening might then be followed by a Claisen-condensation like C-C-bond formation, yielding the six-membered intermediate **222**. $CO_2$ elimination would then lead to formation of the monobenzopyranyl-moiety in **6** (Scheme 3.11 D).




**Scheme 3.10** Possible mechanisms for the observed ring-contraction during formation of noreupenifeldin B **6**: **A,** xenovulene-type ring-contraction mechanism would yield unobserved intermediate **217**; **B,** proposed mechanism of aromatic dehydroxylation catalysed by 4-hydroxybenzoyl-CoA reductase; **C,** proposed oxidative rearrangement during formation of phomanolide E **214**;[177] **D,** possible mechanism of noreupenifeldin B **6** formation *via* ring-expansion, ring-opening and $CO_2$-elimination; X = unknown activation.

The initial ring-expansion to the eight-membered ring in **221** is supported by chemical investigations of ITO *et al*.: Treatment of cyclohepta-3,6-diene-1,2,5-trione **230** with either *m*-chloroperbenzoic acid or anhydrous acetaldehyde afforded the eight-membered acid anhydride **231** (Scheme 3.11).[196]




**230** **231**

**Scheme 3.11:** Chemical investigations into ring-expansion of the tropolone analogue **231** by ITO *et al*.[196]

With EupR5, a putative FAD-dependent monooxygenase was identified within the *eup2* BGC that showed sequence homology to the known ring-contraction enzymes AsL4 and AsL6. In the absence of any other conserved gene of unknown function in the *eup* BGC, the *eup2* BGC and the *pyc* BGC *eupR5* is the most promising candidate to catalyse this reaction. Alternatively, as the ring-contracted tropolone sesquiterpenoids are widely encountered in nature it is possible that an enzyme from primary metabolism, conserved in many or all fungi, might catalyse the observed ring-contraction.

# 4. Understanding and Engineering Humulene Formation

## 4.1. Introduction

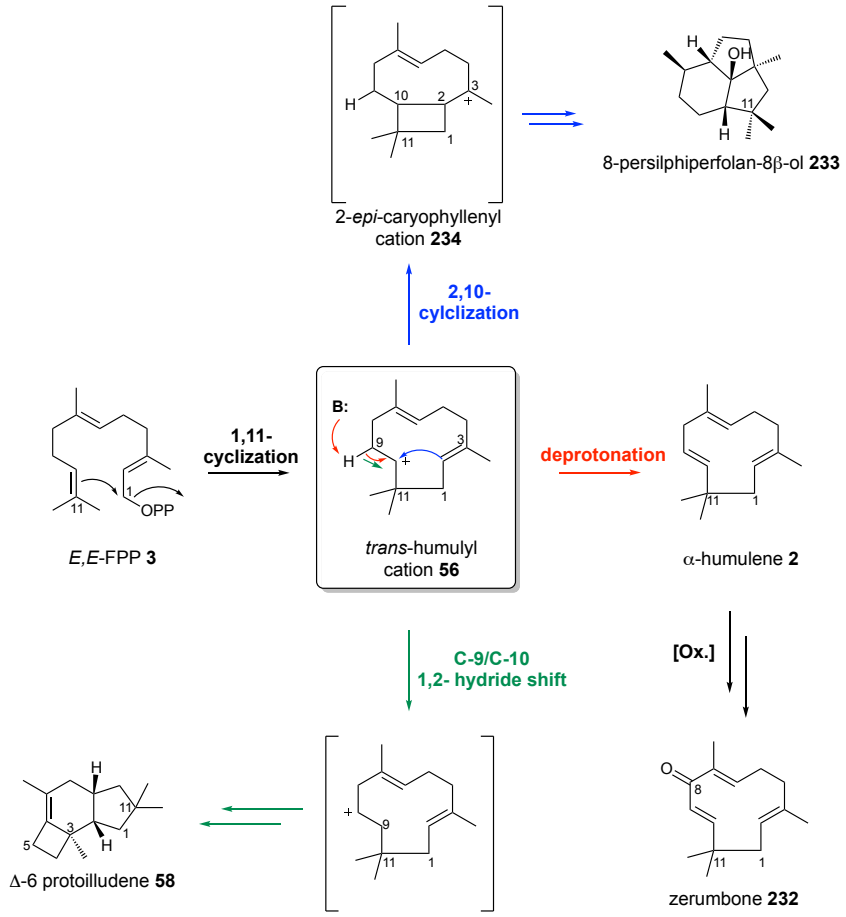
In the course of this study the two non-canonical terpene cyclases AsR6 and EupR3 were identified. The two enzymes catalyse the stereoselective conversion of farnesyl pyrophosphate **3** into α-humulene **2** and 2*Z*-humulene **7**, respectively.[4] Despite a high degree of homology between the two cyclases (~ 50% sequence identity), both enzymes have absolute (!) control over stereoselectivity and produce a single product with distinct stereochemistry at the C-2/C-3 alkene, as was observed upon GCMS analysis in each case (Sections 2.4 and 3.5).

The terpene product of AsR6, α-humulene **2**, is a common component of plant essential oils and yet few enzymes have been linked to the production of **2**. The only *bona fide* humulene synthase identified so far is ZSS1 from *Zingiber zerumbet* Smith. ZSS1 catalyses the first committed step, formation of α-humulene **2**, during zerumbone **232** biosynthesis (Scheme 4.1).[88] However, ZSS1 also produces β-caryophyllene as a minor product (ca. 5%) and shares < 5% sequence identity with AsR6.[88]

In respect to the biosynthesis of α-humulene, **2**-formation is generally assumed to proceed *via* the 1,11-cyclization of farnesyl pyrophosphate **3**. The heterolytic cleavage of the diphosphate moiety ($PP_i$) in **3** and 1,11-closure leads to formation of the *trans*-humulyl cation **56**. Proton elimination from carbon C-9 then affords α-humulene **2** (Scheme 4.1; red arrows).[88] In the case of zerumbone **232** biosynthesis, oxidative chemistry successively establishes the C-8 keto functionality.[165]

The proposed initial formation of the *trans*-humulyl cation **56** (or *E,E*-humulyl cation) is well established and several naturally occurring terpenes are synthesized *via* this intermediate. One such example is presilphiperfolan-8β-ol **233**, that is formed by presilphiperfolan-8β-ol-synthase (PSP; *Armillaria gallica*): After formation of the *trans*-humulylcation **56**, 2,10-cyclization affords the 2-*epi*-caryophyllenyl ion **234**. Successive carbocation rearrangements and the final nucleophilic attack of water afford the pathway product **233** (Scheme 4.1; blue arrows).[197] Likewise, formation of Δ-6 protoilludene **58**, catalyzed by the cyclases Omp6/Omp7 in *Omphalotus olearius*, is assumed to proceed *via* the intermediary *trans*-humulyl cation **56**. After formation of **56** a 1,2-hydride shift and two additional cyclization yield Δ-6 protoilludene **58** (Scheme 4.1; green arrows).[50]

[4] For reasons of simplicity the carbon numbering of humulene was adapted in Section 4 to correspond to the carbon numbering in the precursor FPP **3**.

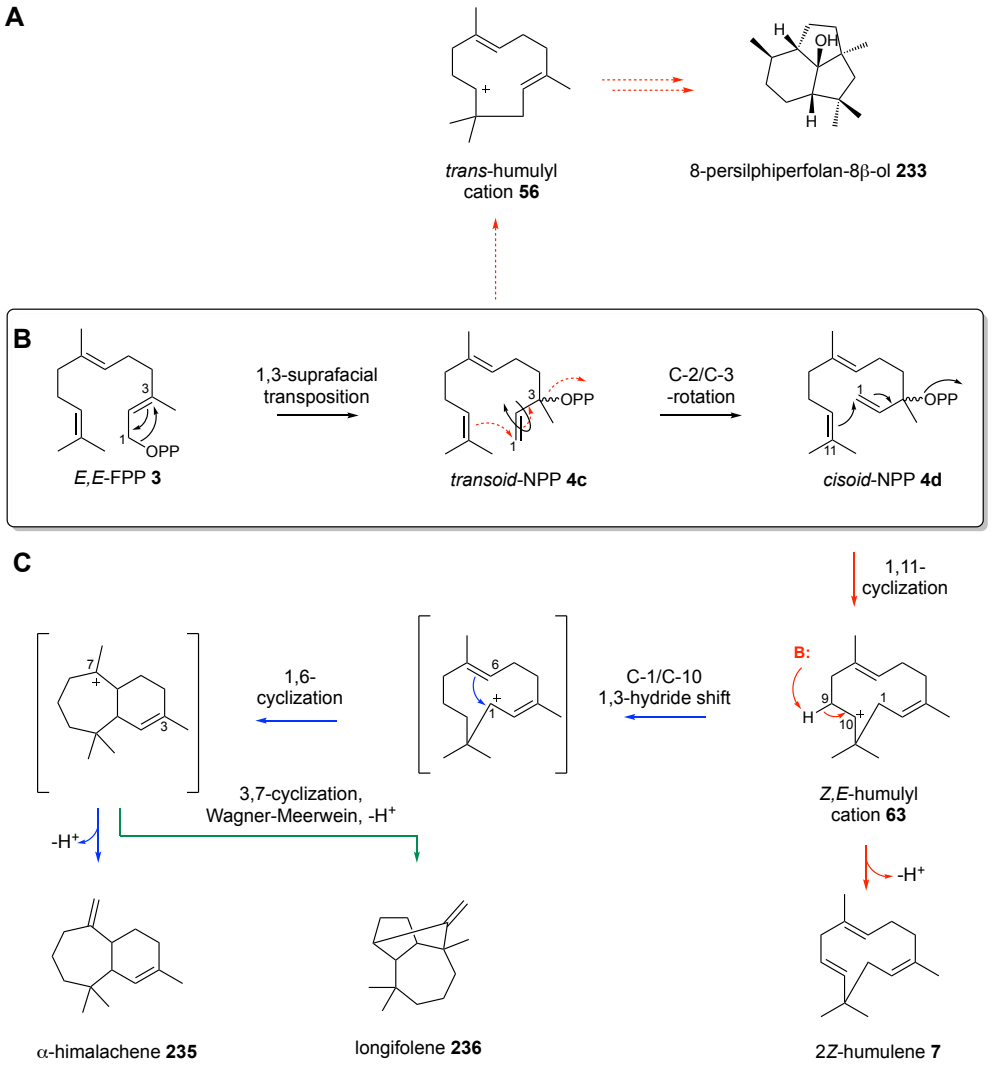



**Scheme 4.1:** Proposed mechanism of α-humulene **2** formation and central role of the *trans*-humulyl cation **56** in diverse sesquiterpenoid pathways; B = general base.

2*Z*-Humulene **7**, the geometric isomer of **2**, has not yet been reported as a naturally occurring terpenoid and no biosynthetic proposal for its formation has yet been described. However, in general, formation of *Z*-alkenes in terpenes requires a formal isomerisation of the *E*-alkene in the universal precursor FPP **3**. This isomerization is usually assumed to proceed *via* an initial 1,3-suprafacial anionotropic rearrangement of the inorganic pyrophosphate, leading to formation of the *transoid* nerolidyl pyrophosphate (NPP) **4c** (Scheme 4.2 B; middle box, black arrows). A conformational change (rotation around the C-2/C-3 bond) then yields the *cisoid* conformer **4d**, which is the proposed precursor of all *Z*-alkene containing sesquiterpenoids.[198]

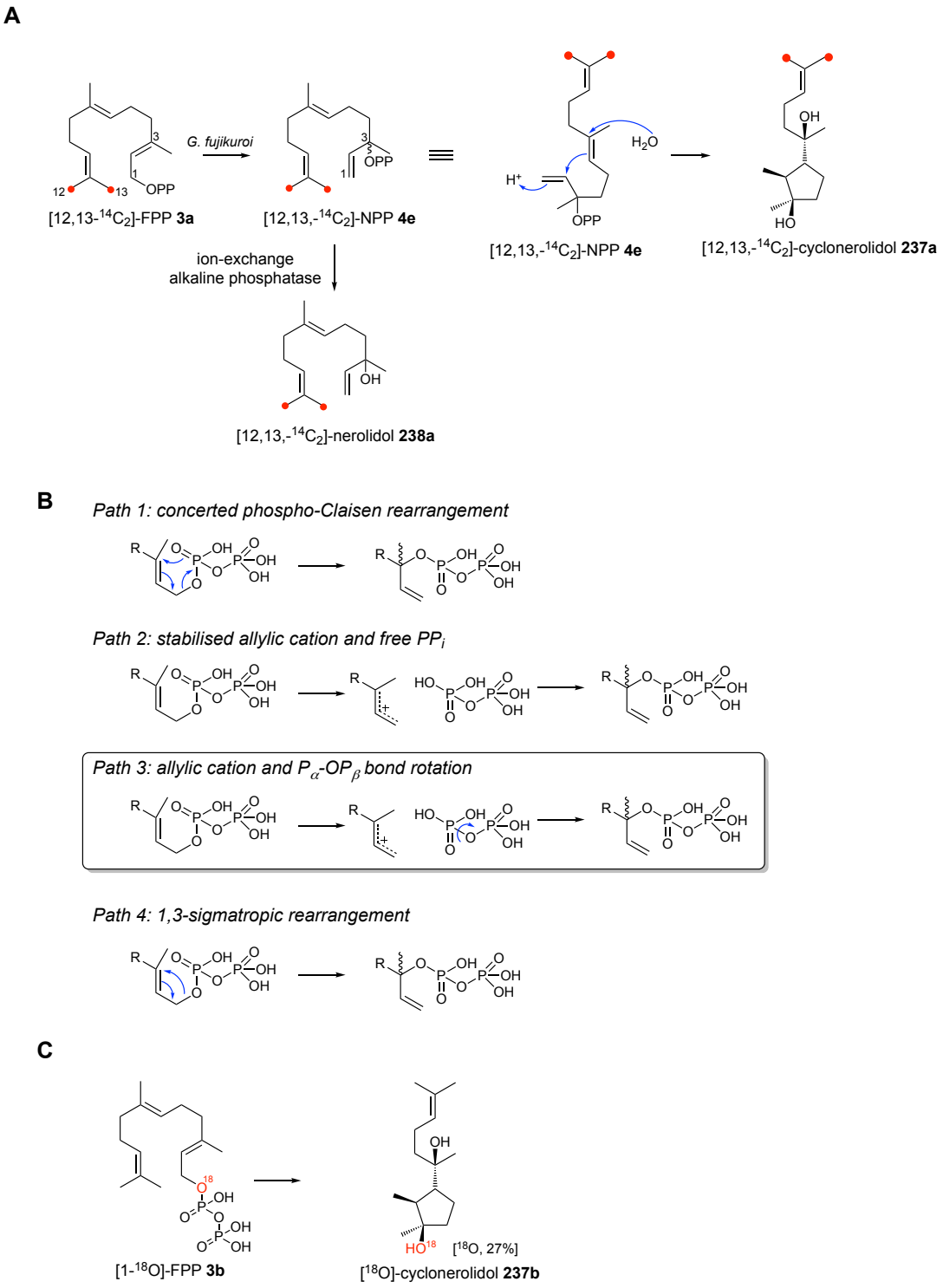
1,11-cyclization of the *cisoid* conformer of NPP **4d**, leads to formation of the *Z,E*-humulyl cation **63** and, as for α-humulene **2** biosynthesis, final elimination of a proton from C-9 could yield 2*Z*-humulene **7** (Scheme 4.2 C; red arrows). While **7**-formation has not yet been investigated, the intermediacy of the *Z,E*-humulyl cation **63** has been proposed during the biosynthesis of several naturally occurring terpenoids, including himalachene **235** and longifolene **236**.[191] These compounds are both produced by the highly promiscuous sesquiterpene cyclase Ag5 from *Abies grandis*. In the case of himalachene **235**, the *Z,E*-humulyl cation **63** is proposed to undergo a 1,3-hydride shift, followed by 1,6-cyclization and deprotonation (Scheme 4.2 C; blue arrows). During

longifolene **236** biosynthesis 1,6-cyclization is proposed to be followed by 3,7-cyclization and WAGNER MEERWEIN rearrangement to yield the pathway product **236** (Scheme 4.2 C; blue and green arrows).[199]




**Scheme 4.2:** Proposed mechanism of 2*Z*-humulene **7** formation and central role of the *Z,E*-humulyl cation **63** in diverse sesquiterpenoid pathways: **A,** possible biosynthetic formation of **233** from NPP; **B,** proposed isomerization of FPP **3** into *cisoid* NPP **4d**; **C,** intermediacy of the *Z,E*-humulyl cation **63** in terpene biosynthesis; B = general base.

It is important to emphasize that the *transoid* conformer of NPP **4c**, is also a possible precursor for formation of α-humulene **2**: 1,11-cyclization of NPP **4c** affords the identical *trans*-humulyl cation **56**, that is also formed from the native substrate FPP **3** (Scheme 4.2 A; red dashed arrows). In fact, the aforementioned terpene cyclase PSP can also convert NPP **4** to presilphiperfolan-8β-ol **233**. Extensive labelling experiments by CANE *et al.*, however, confirmed that NPP is not an intermediate on the biosynthetic pathway to **233** but rather an alternative non-natural precursor (Scheme 4.2 A; red dashed arrows).[197]




**Scheme 4.3** Mechanistic investigations by CANE *et al*. into the FPP/NPP-isomerization: **A,** isotopic labelling of cyclonerolidol **237a** with $^{14}C_2$-labelled FPP **3a**;[200] **B,** mechanistic possibilities of the 1,3-suprafacial transposition; **C,** isotopic labelling of cyclonerolidol **237b** with [$^{18}$O]-labelled FPP **3b** is consistent with path 3.[200]

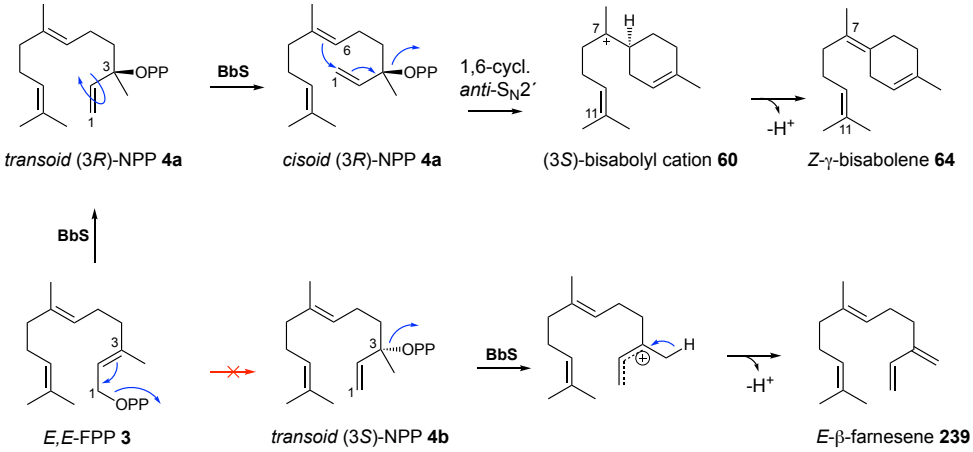
Based on these reported examples of terpene formation, AsR6 is expected to catalyse formation of α-humulene **2** *via* the intermediary *trans*-humulyl cation **56**. EupR3 is expected to first catalyse the isomerization of **4** to NPP **4d** and then catalyse formation of 2*Z*-humulene **7** *via* the intermediary *Z,E*-humulyl cation **63** (as depicted in Scheme 4.1 and 4.2). Quantum chemical

calculations by WANG and TANTILLO suggest that the *E,E*-humulyl cation **56** is energetically favoured over the *Z,E*-humulyl cation **63** ($\Delta E$ = ~9 kcal·mol$^{-1}$).[201] Despite the high homology of AsR6 and EupR3, as yet unknown structural features must therefore enable the latter to catalyse the energetically disfavoured trajectory.

The initial isomerization of FPP **3** into NPP **4** during terpene biosynthesis was first proposed by D. ARIGONI in 1975.[198] Inspiring work by CANE *et al*. on the biosynthesis of cyclonerolidol **237** elucidated the proposed initial anionotropic rearrangement of $PP_i$. Cyclonerolidol **237** was formed *in vitro* using a cell-free enzyme extract from *Gibberella fujikuroi* and FPP **3** (Scheme 4.3 A).[200] Direct evidence for the FPP **3**/NPP **4** isomerization was obtained from elaborate *in vitro* studies with radioactively labelled [12,13-$^{14}C_2$]-FPP **3a**. Short incubation of labelled **3a** with the cell-free extract of *G. fujikuroi*, ion-exchange chromatography and treatment with a phosphatase afforded radioactive [$^{14}$C]-nerolidol **238a** (0.13%; Scheme 4.3 A). As **238a** is the dephosphorylated degradation product of **4** its presence validates NPP **4** as a biosynthetic intermediate on the pathway to cyclonerolidol **237**.[200]

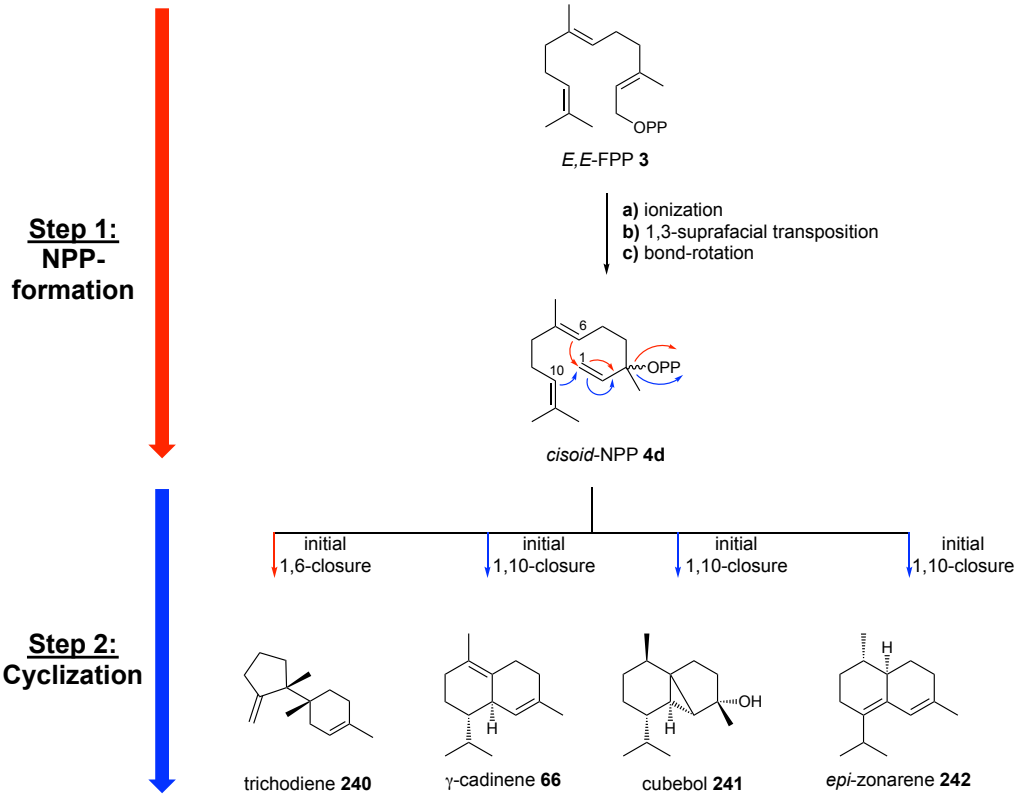
Further labelling experiments shed light on the fate of $PP_i$ during the 1,3-suprafacial transposition. Mechanistically, several possibilities could reasonably explain this allylic rearrangement (Scheme 4.2 B): path **1** proceeds *via* a concerted phospho-Claisen rearrangement and distinct oxygen atoms are involved in the reaction; path **2** proceeds *via* an allylic cation and the intermediate free inorganic pyrophosphate could re-attach *via* any of the six non-bridging oxygen atoms; path **3** proceeds *via* a short-lived ion-pair intermediate that solely allows for rotation around the $P_\alpha$-$OP_\beta$ bond; finally, path **4** proceeds *via* a 1,3-sigmatropic rearrangement.[200] CANE *et al*. distinguished between the pathways by probing the conversion of [1-$^{18}$O]-FPP **3b** (Scheme 4.3 C). The isolated radioactive cyclonerolidol **237b** showed $^{18}$O-enrichment at the C-3 position that equaled one-third of the $^{18}$O-enrichment of the precursor **3b** and thus eliminated paths **1**, **2** and **4**.[200]

Stereochemical aspects of this interconversion add another layer of complexity. DICKSCHAT *et al*. demonstrated that *Z*-γ-bisabolene synthase (BbS) exclusively converts only one enantiomer of NPP, (3*R*)-NPP **4a**, into *Z*-γ-bisabolene **64** (Scheme 4.4). Incubation of BbS with (3*S*)-NPP **4b** results in the selective elimination of the diphosphate and formation of *E*-β-farnesene **239** (Scheme 4.4).[49] Likewise, amorpha-4,11-diene synthase only converts (3*R*)-NPP **4a** into the eponymous amorphadiene.[202] It is noteworthy, however, that very few studies have considered the stereochemistry of NPP specifically, and apart from these BbS studies few other examples systematically assessed both enantiomers.




**Scheme 4.4:** Proposed biosynthetic pathway leading to *Z*-γ-bisabolene **64** and *E*-β-farnesene **239** from the different NPP enantiomers **4a** and **4b**; BbS = *Z*-γ-bisabolene synthase.

The initial isomerization of FPP **3** to NPP **4** significantly expands the chemical space of the terpenome and numerous well-studied pathway products are derived from this. In addition to the aforementioned examples, NPP **4** has been proposed as an intermediate in the biosynthesis of trichodiene **240**,[203] *δ*-cadinene **66**,[51] cubebol **241**,[204] *epi*-zonarene **242**[50] and many others. In general, the isomerization process and the cyclization process are regarded as two *distinct* chemical steps: In the *first* step ionization of FPP **3**, isomerization and reattachment of $PP_i$ yields nerolidyl diphosphate **4c**, or after bond rotation, **4d**. In a separate *second* step NPP **4** undergoes reionization and terpene cyclization (Scheme 4.5).[51]




**Scheme 4.5:** Separate chemical steps catalyse the formation of NPP **4** and subsequent terpene cyclization.

Furthermore, it is important to emphasize that despite the agreement on the chemical mechanism of *Z*-alkene formation, this understanding is not mirrored at a structural level. Although crystal structures have been reported for enzymes that catalyze formation of *e.g.* trichodiene **240**[205] or *δ*-cadinene **66**,[51] no conserved structural features have been identified that drive the initial isomerization process. Investigations into this mechanism are impaired by: **a),** terpene cyclases often producing complicated mixtures of products arising both from FPP **3** and NPP **4**; **b),** a small number of terpene cyclase crystal structures deposited with the protein data base (https://www.rcsb.org); **c),** a general lack of understanding in respect to structure-function relationships in terpene cyclization (with exception of the diphosphate binding mode); and **d),** inconsistent use of NPP **4** and its two enantiomers (3*R*, 3*S*) for the *in vitro* characterisation of new cyclases.

The very high level of stereoselectivity observed in AsR6 and EupR3, giving rise to observation of only one of the two geometric humulene isomers **2** and **7**, in combination with the high overall sequence identity, thus represents an ideal model system to study structure-function relationships that drive the formation of *Z*-alkenes during terpene biosynthesis, and in turn the very early steps of complex pathways such as those involved in presilphiperfolan-8β-ol **233** formation.

## 4.2. Project Aims

In this part of the study the mechanism of humulene formation by AsR6 and EupR3 will be investigated on a chemical and structural level. In initial work the conversion of different isoprenoids (FPP **3**, NPP **4**) will be probed for the wild-type enzymes. Parallel to this, in cooperation with DR. PEER LUKAT (Helmholtz Centre for Infection Research, BLANKENFELDT research group, Braunschweig), we will try to obtain the crystal structure of AsR6 in the unliganded state and in complex with the substrate analogue thiolo-*S*-farnesyl diphosphate. Structural data for AsR6 in combination with a global sequence alignment of terpene cyclases involved in the biosynthesis of tropolone sesquiterpenoids should then identify candidate amino acid residues that are involved in controlling the stereoselectivity during humulene formation. Site-directed mutagenesis will be used to probe the effect of single amino acid mutations on the conversion of FPP- and NPP- substrates alike.

## 4.3. Results – Conversion of FPP and NPP by AsR6 and EupR3

### 4.3.1. Preparation of Recombinant AsR6 and Substrates

For the heterologous expression of *asR6* an *E. coli* codon-optimized gene cassette was obtained as a synthetic construct from *Thermo Fischer Scientific* (plasmid backbone: pET100/D-TOPO). By design the construct encodes an *N*-terminal $His_6$-tag, that facilitates isolation and purification of recombinant *N*-$His_6$-AsR6 by IMAC-technology.[206] Initial expression trials were conducted on an analytical 50 mL scale. A series of expression conditions varying the expression temperature (12 °C, 16 °C, ambient temperature) and IPTG-concentration (0.1 mM – 1.0 mM) were systematically assessed (data not shown). AsR6 was found to be overexpressed in several conditions, with best expression levels observed for 16 °C and low IPTG-concentrations (0.1 mM).

For the preparative production of *N*-$His_6$-AsR6, expression experiments were conducted on a 10 L-scale. Submerged fermentation in 2TY medium was followed by cell harvest and total cell lysis *via* sonification. *N*-$His_6$-AsR6 was then purified from the total lysate by $Ni^{2+}$ NTA affinity chromatography. The protein typically eluted at an imidazole concentration of approximately 200 mM (Figure 4.1 A). Combined elution fractions containing *N*-$His_6$-AsR6 were concentrated and further purified by size-exclusion chromatography (SEC; Superdex 26/600; manually packed). *N*-$His_6$-AsR6 eluted as a single peak with a retention volume of 176 mL (Figure 4.1 B). The purification process was monitored by SDS-PAGE and *N*-$His_6$-AsR6 eluted with an apparent molecular weight of 52 kDa (Figure 4.1 C). Pure *N*-$His_6$-AsR6 (typical yields: 0.8 $mg \cdot L^{-1}$ expression culture after SEC) was concentrated to a final concentration of 4 - 15 $mg \cdot mL^{-1}$ (in SEC-buffer), shock-frozen in liquid nitrogen and stored at - 80 °C for up to three months.

Initial crystallization experiments using the sitting-drop method and a variety of commercially available screens demonstrated that the *N*-terminal polyhistidine tag impaired crystal formation. To overcome this problem fusion tags are often cleaved using site-specific proteolysis prior to crystallization approaches.[207] The initially purified *N*-$His_6$-AsR6 construct contained an enterokinase restriction site (Asp-Asp-Asp-Asp-Lys↓) between the purification tag and AsR6, but attempts to cleave the tag were not successful (data not shown). To enable proteolysis the codon-optimized *asR6* gene cassette was subcloned into the plasmid pETM-11, with a TEV-protease recognition site (Glu-Asn-Leu-Tyr-Phe-Gln↓-Gly/Ser) encoded between the *N*-terminal $His_6$-Tag and *asR6*.[5] Generation of plasmid pETM-11-*asR6* (labbook ID: CS I 135 1-7) was confirmed by

[5] DR. SIMONE HÖFLER (CARLOMAGNO research group, BMWZ Hannover) is kindly thanked for the gift of pETM-11 and recombinant TEV-protease.

DNA sequencing. Heterologous expression and purification by $Ni^{2+}$ NTA affinity chromatography was performed as described for the first AsR6-construct (data not shown). Instead of the SEC purification step *N*-$His_6$-AsR6 obtained from expression of plasmid CS I 135 1-7 was incubated with TEV-protease overnight.[5] Reverse $Ni^{2+}$ NTA affinity chromatography then afforded untagged AsR6 in the flow-through fraction (0 - 40 mL; Figure 4.1 D). AsR6 was combined from all eluting fractions, concentrated to 5 - 15 $mg \cdot mL^{-1}$ (in SEC-buffer) and stored at - 80 °C for up to three months.

To solve the phase problem during elucidation of the AsR6 crystal structure single-wavelength anomalous dispersion (SAD) was used.[208] For this, production of seleno-*L*-methionine derivatized AsR6 was required. Expression of *asR6* in *E. coli* BL21 in the presence of seleno-*L*-methionine was achieved by autoinduction[209] and conducted by UTE WIDOW (Helmholtz Centre for Infection Research; BLANKENFELDT research group, Braunschweig). Cell lysis and purification of *N*-$His_6$-SeMet-AsR6 by $Ni^{2+}$ NTA affinity chromatography and SEC was accomplished in this study (data not shown).

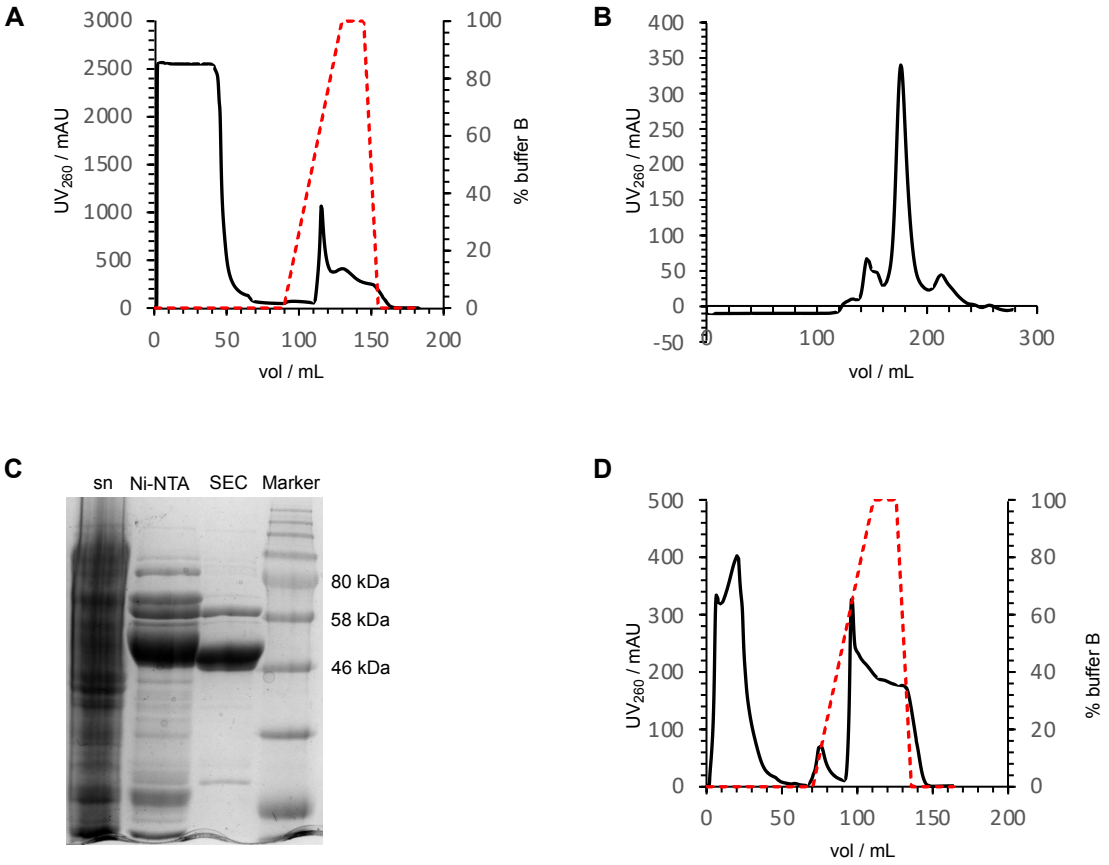
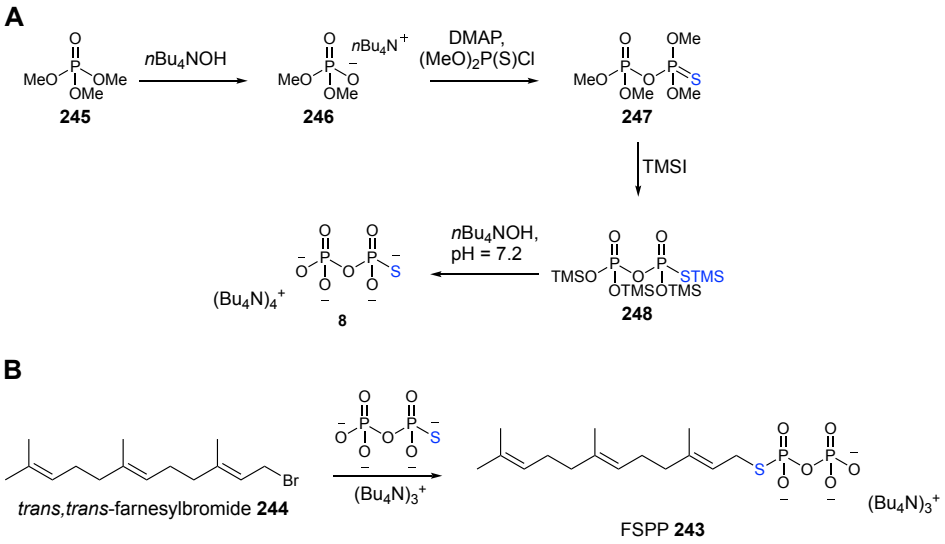



**Figure 4.1** Purification of recombinant AsR6: **A,** UV chromatogram ($\lambda$ = 280 nm) monitoring $Ni^{2+}$ affinity chromatography; dashed red line = % buffer B; **B,** UV chromatogram ($\lambda$ = 280 nm) monitoring SEC; **C,** SDS-PAGE of the purification of AsR6; sn = supernatant; Ni NTA = elution fraction $Ni^{2+}$ NTA; SEC = elution fraction SEC; Marker = Color Prestained Protein Standard, Broad Range (11 - 245 kDa; *NEB*); **D,** UV chromatogram ($\lambda$ = 280 nm) monitoring reverse $Ni^{2+}$ affinity chromatography; dashed red line = % buffer B.

All three AsR6-proteins (*N*-His$_6$-AsR6, AsR6, SeMet-AsR6) were delivered to DR. PEER LUKAT, who performed all work on crystallizing AsR6 and structure elucidation.

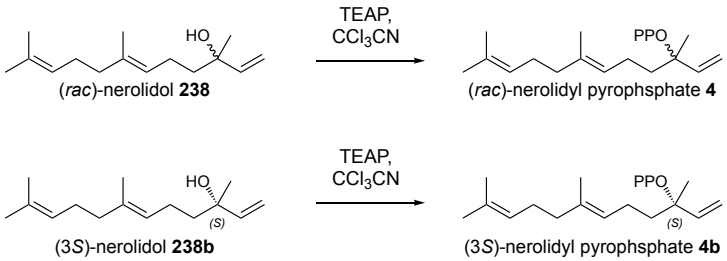
To identify amino acid residues involved in substrate binding/humulene formation a non-hydrolysable substrate analogue, thiolo-*S*-farnesyl diphosphate (FSPP **243**), was synthesized by ADRIAN DEUSCHMANN (master student, COX research group, Hannover). FSPP **243** has previously been demonstrated to be an unreactive substrate analogue of the native substrate FPP **3**, in which the bridging oxygen between isoprenoid and diphosphate is replaced by a sulfur atom. Examples where FSPP **243** was successfully deployed in co-crystallization experiments are *e.g. S. aureus* dehydrosqualene synthase (CrtM) and *Aspergillus terreus* aristolochene synthase (ATAS).[210,211] FSPP **243** was synthesized according to an established protocol of POULTER *et al. via* nucleophilic substitution of *trans*, *trans*-farnesyl bromide **244** with the *tris*(tetra-*n*-butylammonium) salt of inorganic thiolo-*S*-diphosphate **8**.[212,213] The five-step synthesis towards **243** from trimethyl phosphate **244** is outlined in Scheme 4.6.




**Scheme 4.6** Chemical synthesis of thiolo-*S*-farnesyl diphosphate FSPP **243**: **A,** synthesis of thiolo-*S*-diphosphate **8**; **B,** synthesis of FSPP **243**; synthesis performed by ADRIAN DEUSCHMANN.

In addition to the unnatural substrate analogue FSPP **243**, *racemic* (*rac*)-nerolidyl diphosphate **4** and (3*S*)-nerolidyl diphosphate **4b** were synthesized. To do so, either commercial (*rac*)-nerolidol **238** or (3*S*)-nerolidol **238b** were transformed into the corresponding diphosphates. A well-established phosphorylation approach was deployed, which uses *bis*-triethylammonium phosphate in trichloroacetonitrile as phosphorylating reagent (Scheme 4.7).[214] In the absence of commercially available sources of (3*R*)-nerolidol **238c** the corresponding (3*R*)-nerolidyl diphosphate **4a** was not synthesized in the scope of this thesis. The absolute stereochemistry of these compounds was confirmed by measurement of the optical rotation and comparison to

literature data. It is noteworthy that suppliers claims regarding the stereochemical integrity of nerolidol should be carefully checked before use!




**Scheme 4.7:** Chemical synthesis of (*rac*)-/(3*S*)-nerolidyl diphosphate **4** and **4b**;[214] TEAP = *bis*-triethylammonium phosphate solution.

## 4.3.2. Conversion of Diphosphate Substrates by AsR6 and EupR3

*In vitro activity of AsR6*

Initially the activity of AsR6 wild-type enzyme with FPP **3**, (*rac*)-NPP **4** and (3*S*)-NPP **4b** was tested. All assays throughout were performed on a 500 µL-scale with $c_{enz.}$ = 1.0 mg·mL$^{-1}$ (final concentration) and 150 µM of the corresponding diphosphate. The assays were incubated at 28 °C for 30 min and then extracted with n-hexane. All obtained n-hexane extracts were directly analyzed by GCMS. Commercial α-humulene **2** (*Sigma Aldrich*) was used as a reference standard, as described previously in detail in Section 2.5 and 3.5.

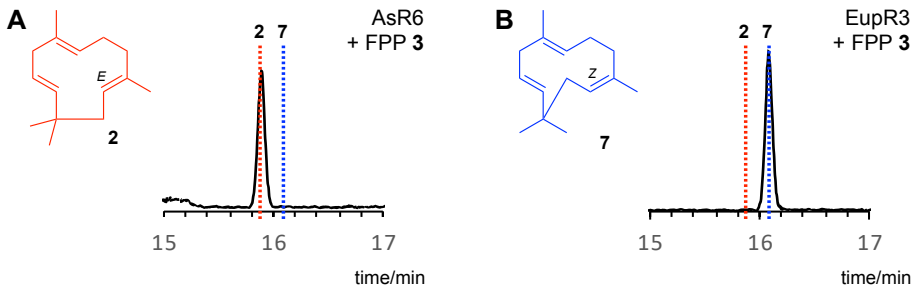



**Figure 4.2** Conversion of FPP **3** by: **A,** AsR6; **B,** EupR3; displayed are the GCMS TIC chromatograms of n-hexane extracts of the corresponding assay; the retention time of the respective humulene isomer **2**/**7** is colour-coded in red (**2**) and blue (**7**).

Incubation of AsR6 with FPP **3** afforded a single product with a retention time $t_R$ = 15.9 min, that was unambiguously confirmed as α-humulene **2** based on its retention time and mass fragmentation pattern (Figure 4.2 A; EI mass spectrum not displayed). In order to compare effects of different diphosphate substrates and point mutations in AsR6, identical *in vitro* assays were performed with EupR3 to establish reference values (retention time, MS-data) for 2*Z*-humulene **7** (for identification and characterization of **7** compare Section 3.5). Incubation of EupR3 with FPP **3** afforded a single product, 2*Z*-humulene **7**, that eluted at $t_R$ = 16.1 min (Figure 4.2 B). The difference in retention time ($\Delta t_R$ = 0.2 min) was sufficient to differentiate between the two

geometric isomers and the assays with FPP **3** and AsR6/EupR3 were always used as a control/reference in the following *in vitro* assays. Due to technical difficulties in the course of this study two different GCMS systems were used. The occasionally deployed second GCMS system led to identical peak resolution of the humulene isomers ($\Delta t_R$ = 0.2 min), but the retention times were shifted to $t_R$ = 21.8 min (α-humulene **2**) and $t_R$ = 22.0 min (2*Z*-humulene **7**; *e.g.,* compare Figure 4.3 A).

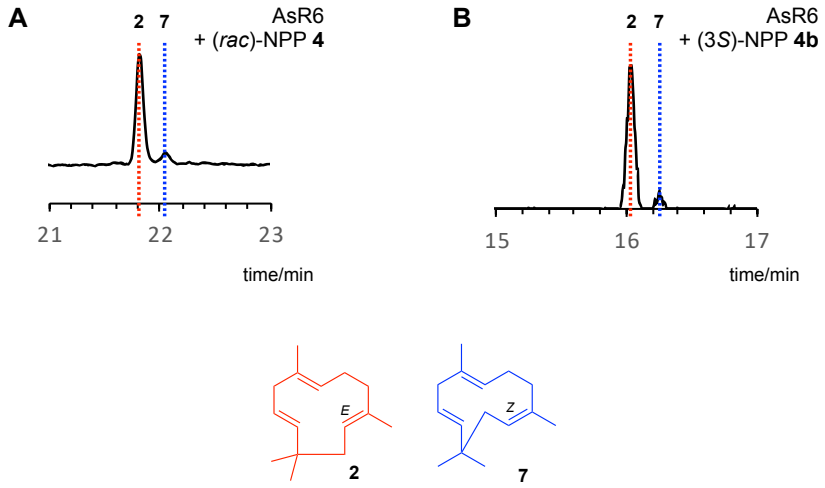



**Figure 4.3** Conversion of: **A,** (*rac*)-NPP **4**; and **B,** (3*S*)-NPP **4b** by AsR6 wild-type enzyme; displayed are the GCMS TIC chromatograms of n-hexane extracts of the corresponding assay; the retention time of the respective humulene isomer **2**/**7** is colour-coded in red (**2**) and blue (**7**).

Having established formation of α-humulene **2** by AsR6 its activity towards the non-native nerolidyl diphosphates **4** and **4b** was tested. First, incubation of AsR6 with (*rac*)-NPP **4** was investigated. Analysis of the assay extracts identified that (*rac*)-NPP **4** is converted to the native product α-humulene **2** as the major product (Figure 4.3 A). However, as a minor byproduct traces of 2*Z*-humulene **7** were identified, that were never previously observed to be produced by AsR6 (Figure 4.3 A). In this case, the chemical nature of NPP **4** must enable formation of the non-native product **7**, albeit in minute quantities. It is important to emphasize that no other products besides **2** and **7** were observed by GCMS. Assays using the enantiomerically pure (3*S*)-NPP **4b** as substrate led to an identical product distribution: AsR6 converts **4b** into the native terpene product α-humulene **2** as the major product but traces of the geometric isomer **7** were detectable by GCMS analysis (Figure 4.3 B).

All sesquiterpene cyclases investigated to date are reported to solely convert one isomer of NPP **4,** usually (3*R*-)-NPP **4a,** into the cyclic reaction product. Conversion of (3*S*)-NPP **4b** by AsR6 thus raised the question whether the enzyme shows an unusual specificity for the (3*S*)-enantiomer – or whether the enzyme can convert both enantiomers of (*rac*)-NPP **4** into humulene. In the absence of synthetic (3*R*)-NPP **4a** the aforementioned *Z*-γ-bisabolene synthase (BbS) was used as a "reporter" enzyme to answer this question: BbS converts (3*S*)-NPP **4b** into the acyclic β-

farnesene **239**, whereas the "reactive" enantiomer, (3*R*)-NPP **4a**, is converted into the cyclic product *Z*-γ-bisabolene **64** (compare Scheme 4.4 in Section 4.1).[49] In this study the specific selectivity of BbS was exploited to deduce whether AsR6 also converts (3*R*)-NPP **4a**: The addition of BbS *after* completion of the reaction between AsR6 and (*rac*)-NPP **4** should reveal if the (3*R*)-enantiomer is converted by AsR6.

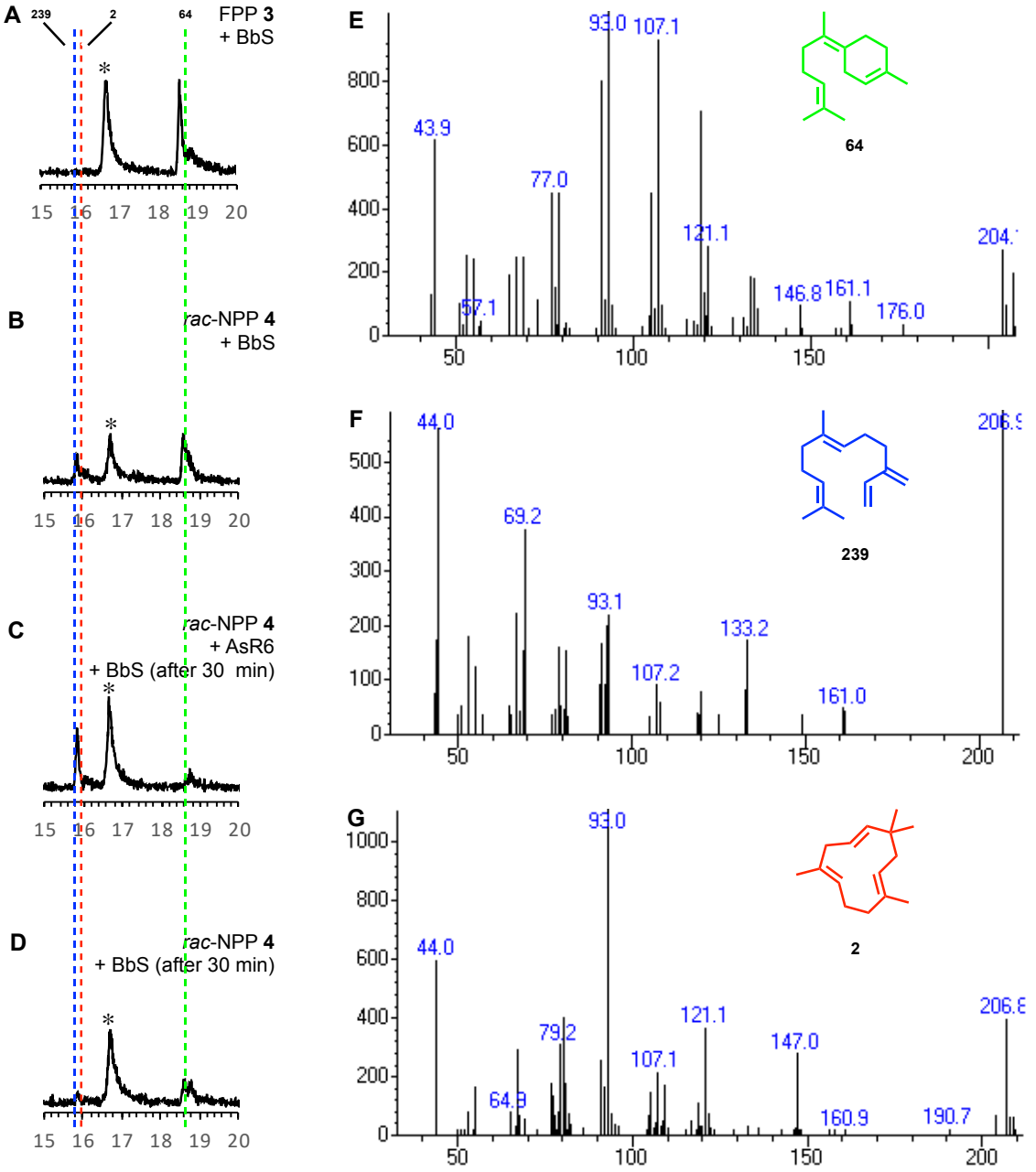



**Figure 4.4** Co-incubation experiments of *rac*-NPP **4** with BbS and AsR6: GCMS TIC chromatograms of assay extracts from the incubation of **A,** FPP with BbS; **B,** (*rac*)-NPP with BbS; **C, (***rac*)-NPP with AsR6 and addition of BbS after 30 min; **D,** (*rac*)-NPP with addition of BbS after 30 min (control); **E,** EI mass spectrum of *Z*-γ-bisabolene **64** from A ($t_R$ = 18.8 min); **F,** EI mass spectrum of *E*-β-farnesene **239** from B ($t_R$ = 15.9 min); **G,** EI mass spectrum of α-humulene **2** from D ($t_R$ = 15.9 min); * = unrelated.

In principle, GCMS analysis of this post-incubation experiment could lead to two possible results:

**A)** AsR6 completely converts both enantiomers of (*rac*)-NPP **4** into α-humulene **2**. In this case delayed addition of the reporter enzyme BbS should not lead to formation of BbS-specific products (**64** or **239**).

**B)** AsR6 converts only (3*S*)-NPP **4b** into α-humulene **2**. In this case delayed addition of the reporter enzyme BbS should lead to formation of *Z*-γ-bisabolene **64**, as the (3*R*)-enantiomer is still present after the reaction between AsR6 and (*rac*)-NPP **4** is complete.

To probe this question the *bbs* gene[6] was expressed in *E. coli* BL21 and recombinant BbS was afforded after immobilized metal affinity chromatography (Figure 4.12; see Section 4.5). Reconstitution of BbS activity *in vitro* reproduced the reported product spectra of this enzyme: Incubation of BbS with FPP **3** led to formation of the single product *Z*-γ-bisabolene **64** ($t_R$ = 18.8 min; Figure 4.4 A; green line). *Z*-γ-bisabolene **64** was identified based on comparison of the EI mass spectrum with reported spectra for **64** (Figure 4.4 E).[49]

Incubation of BbS with (*rac*)-NPP **4** afforded, as expected, a mixture of *Z*-γ-bisabolene **64** and β-farnesene **239** (Figure 4.4 B; identification of β-farnesene **239** [$t_R$ = 15.9 min; blue line] was based on comparison of the EI mass spectrum [Figure 4.4 F] with reported spectra for **239**).[215]

Having established the activity of BbS *in vitro* the post-incubation experiment with AsR6 and BbS was conducted. In this, BbS was added 30 min after the assay with (*rac*)-NPP **4** and AsR6 was started. Extraction with n-hexane after 60 min and analysis by GCMS clearly identified α-humulene **2** ($t_R$ = 15.8 min) as the dominant product (Figure 4.4 C; red line). Since the retention time of α-humulene **2** and β-farnesene **239** are very similar the EI mass spectrum of this peak was analyzed to validate this compound as α-humulene **2** (Figure 4.4 G). Indeed, the mass spectrum clearly identified α-humulene **2**. Key mass fragments corresponding to β-farnesene **239** (*e.g., m/z* = 133.2; *m/z* = 69.2; compare Figure 4.4 F) were not observed. Only traces of *Z*-γ-bisabolene **64** were identified in this assay ($t_R$ = 18.8 min), suggesting that AsR6 converts *both* enantiomers of (*rac*)-NPP **4** to **2**. To rule out possible degradation of (3*R*)-NPP **4a** in the 30 min before addition of BbS a control experiment was performed: Here, (*rac*)-NPP **4** was incubated in the assay buffer for 30 min prior to addition of BbS and then extracted after 60 min. Both products of BbS were identified by GCMS (Figure 4.4 D). Direct comparison with the AsR6/BbS post-incubation experiment showed that with AsR6 present *Z*-γ-bisabolene **64**-production was reduced, indicating that AsR6 had already converted the majority of (3*R*)-NPP **4a** present. Based on these findings it was concluded that AsR6 converts both enantiomers of (*rac*)-NPP **4** into α-humulene **2**.

[6] PROF. JEROEN DICKSCHAT is kindly thanked for the gift of plasmid pYS-*bbs*.

*In vitro activity of EupR3*

The same *in vitro* assays were performed with wild-type EupR3. Formation of 2*Z*-humulene **7** as the sole cyclase product from the incubation with FPP **3** was already established (compare Fig. 4.2 B). Incubation of (*rac*)-NPP **4** with EupR3 and incubation with (3*S*)-NPP **4b** led to identical product distributions and only a single product, 2*Z*-humulene **7**, was identified in all organic extracts (Figure 4.5 A and B). This data confirms EupR3 as a *bona fide* 2*Z*-humulene synthase. Having completely characterized the *in vitro* activity of both wild-type enzymes towards FPP **3** and NPP **4**, structural data for the enzymes was required to identify factors that drive the different stereochemistry in humulene formation.

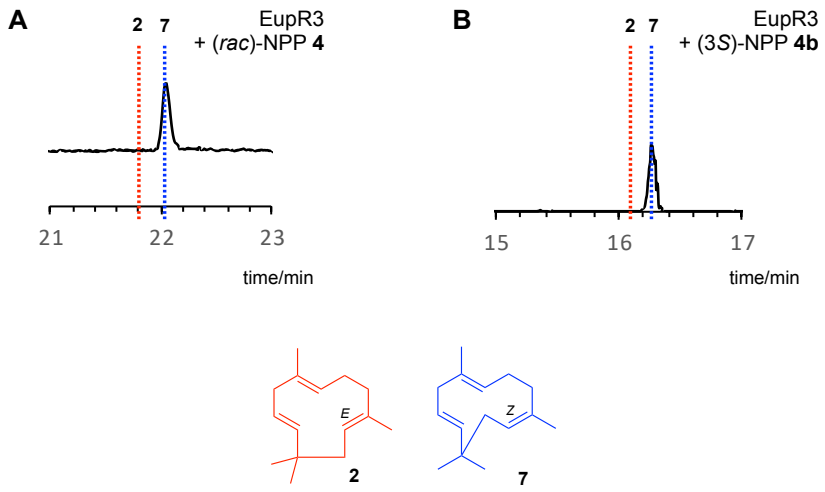



**Figure 4.5** Conversion of: **A,** (*rac*)-NPP **4**; and **B,** (3*S*)-NPP **4b** by EupR3 wild-type enzyme; displayed are the GCMS TIC chromatograms of n-hexane extracts of the corresponding assay; the retention time of the respective humulene isomer **2**/**7** is colour-coded in red (**2**) and blue (**7**).

## 4.4. The AsR6 Crystal Structure

### 4.4.1. Crystal Structure of the Unliganded AsR6

The crystal structure of the unliganded AsR6 was obtained at 2.0 Å resolution from crystals derived from SeMet-*N*-$His_6$-AsR6. AsR6 crystallized as a homodimer in a typical α-helical bundle fold, that is conserved among all reported class I terpene cyclases (Figure 4.6 A).[39] Notably, the first 36aa were found not to be structured but 34 of these residues are artificially introduced from cloning ($His_6$-Tag and TEV-cleavage site). Likewise, the *C*-terminal 22aa did not form a specific secondary structure, suggesting that a *C*-terminally truncated construct might further improve the crystallization process.

Analysis of the electron-density revealed 19 helices that are connected by short linker sequences and form the AsR6 core structure. The structural fold of AsR6 was compared to known protein folds using the DALI protein structure comparison server.[216] Intriguingly the overall fold is most

similar to other class I terpene cyclases, despite AsR6 sharing less than 15% sequence identity with these enzymes. The closest fold relatives are bacterial selinadiene-synthase (PDB: 4okm), trichodiene synthase (PDB: 1jfa) and epi-isozizaene synthase (PDB: 3kb9; Table 4.1). Structural alignment of AsR6 with selinadiene synthase illustrates the similarity as the core helices form homologous structures (Figure 4.6 B).

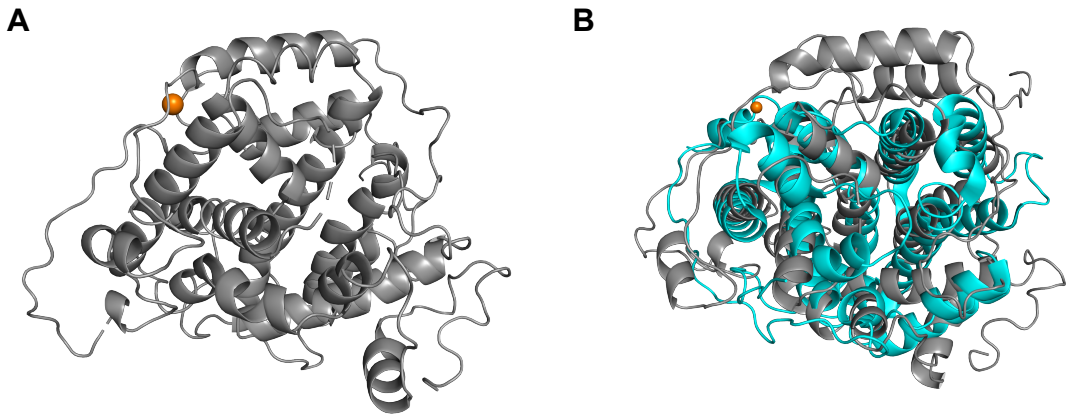



**Figure 4.6** Crystal structure of the unliganded AsR6 humulene synthase: **A,** overview of the AsR6 crystal structure; orange = $Zn^{2+}$ ion; **B,** structural alignment of AsR6 (gray) with the canonical terpene cyclase selinadiene-synthase (cyan; PDB: 4okm); orange = $Zn^{2+}$ ion.

In the unliganded form AsR6 does not bind any $Mg^{2+}$ ions, as previously observed for other class I terpene cyclases.[205] However, a Zinc binding site was identified that is located distal from the active site. A single $Zn^{2+}$ ion is coordinated by Cys43, Cys85, Cys87 and His41 (Figure 4.7 A). As the $Zn^{2+}$-binding cluster is located away from the active site it is unlikely to be involved in the catalytic mechanism but might contribute to overall protein stability. Sequence alignment of the known humulene synthases AsR6, PycR6, EupR3 and EupE[80] revealed that the four residues involved in $Zn^{2+}$ coordination are highly conserved (Figure 4.7 B). No other terpene cyclase with a $Zn^{2+}$ binding cluster has yet been reported. Sequence and structural alignment of AsR6 with canonical cyclases (*e.g.,* fold-relatives listed in table 4.1) support the uniqueness of this feature in AsR6 (comparison not shown).

**Table 4.1** Closest fold relatives of AsR6 based on DALI protein structure comparison server:[216] DALI analysis performed by DR. PEER LUKAT; RMSD-values of backbone atoms were calculated with PyMOL 2.3.0; RMSD = root mean square deviation.

| # | PDB access. | numb. of resid. | % seq. ident. | Fold relative | RMSD |
|---|---|---|---|---|---|
| **1** | 4okm-A | 346 | 12 | selinadiene synthase[217] | 4.43 Å |
| **2** | 1jfa-A | 354 | 8 | trichodiene synthase[205] | 5.64 Å |
| **3** | 3kb9-A | 340 | 10 | epi-isozizaene synthase[218] | 6.06 Å |
| **4** | 6ggi-B | 300 | 8 | cyclooctat-9-en-7-ol synthase[219] | 4.75 Å |
| **5** | 4xlx-A | 288 | 5 | *ent*-kaurene synthase[220] | 4.46 Å |

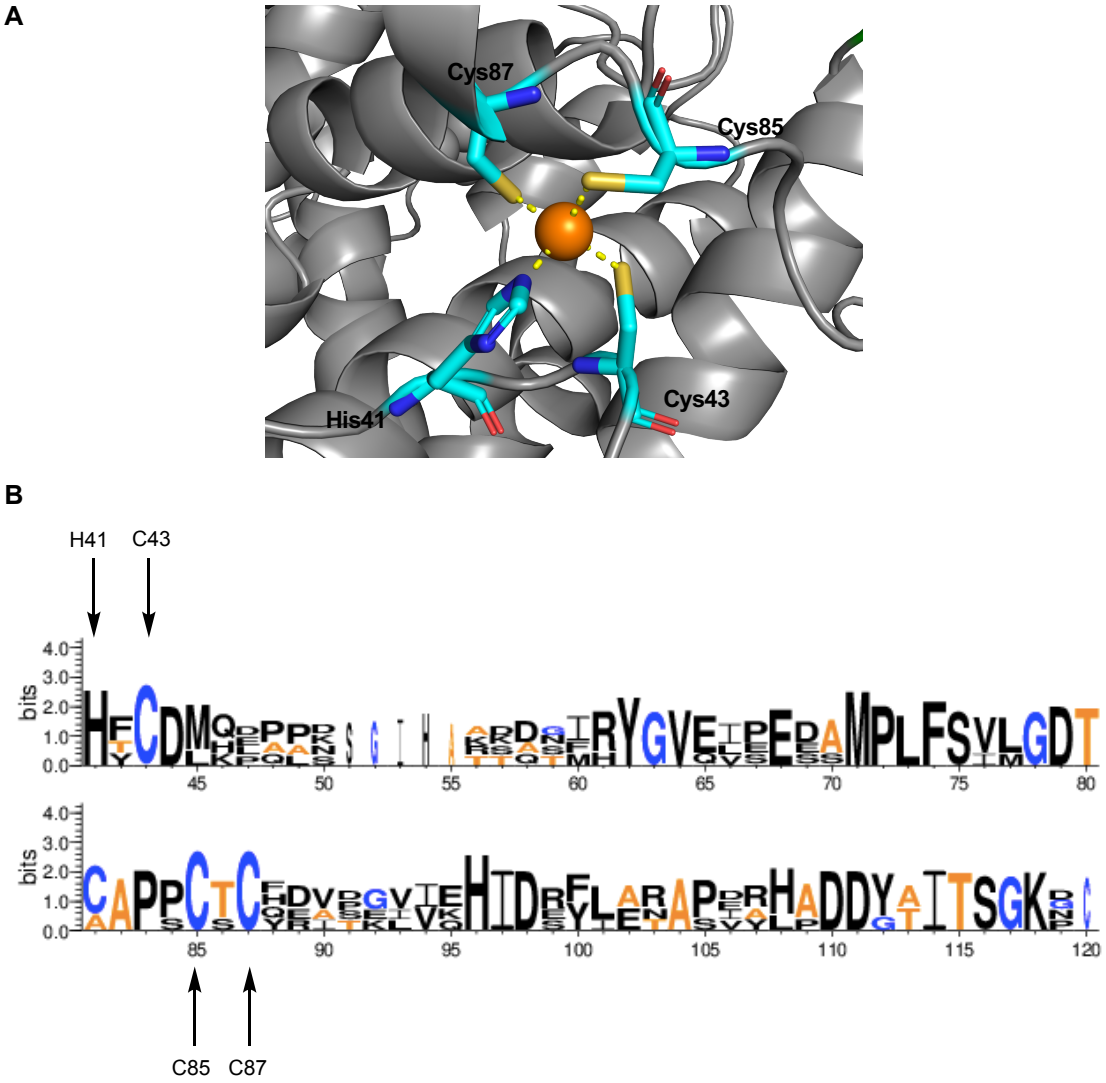



**Figure 4.7** Conserved $Zn^{2+}$ binding site in humulene synthases: **A,** $Zn^{2+}$ binding site in the AsR6 crystal structure; **B,** conserved His41, Cys43, Cys85, Cys87 motif in all AsR6-type humulene synthases; figure created with WebLogo 3;[221] sequence alignment of AsR6, PycR6, EupR3 and EupE performed with ClustalW.[164]

### 4.4.2. Crystal Structure of AsR6 in Complex with FSPP

The crystal structure of AsR6 ($His_6$-Tag cleaved with TEV protease prior to crystallization) in complex with thiolo-*S*-farnesyl diphosphate **243** was obtained at 1.7 Å resolution. Surprisingly FSPP **243** did not crystallize as an intact molecule. Analysis of the electron density revealed a cleaved inorganic thiolo-*S*-diphosphate ($SPP_i$) moiety in the active site (Figure 4.8 A and B). Additional density best fitted to the *in crystallo* cyclized reaction product, α-humulene **2**. However, an unambiguous attribution of the electron density to **2** was not possible and likely a mixture of humulene **2** and non-cyclized FSPP **243** is present in the active site, with humulene **2** being the major component. Similar *in crystallo* cyclized reaction products have also been reported during the co-crystallization of the diterpene cyclase CotB2 and 2-fluoro-geranylgeranyl diphosphate (FGGDP).[219]

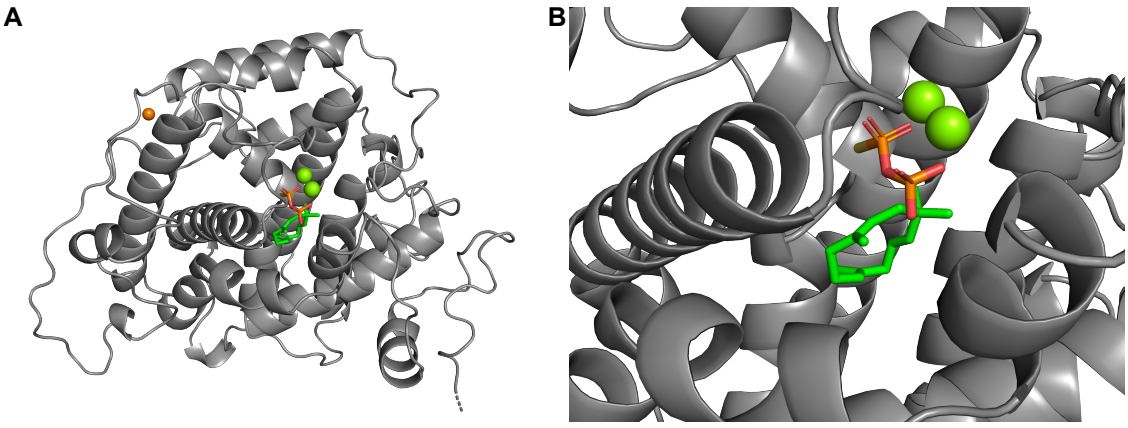



**Figure 4.8** Crystal structure of AsR6 in complex with thiolo-*S*-diphosphate and an *in crystallo* cyclized reaction product: **A,** overview of the AsR6 crystal structure; **B,** overview of the AsR6 active site cavity; green spheres = $Mg^{2+}$ ions, orange spheres = $Zn^{2+}$ ions.

The unintended cyclization of the unreactive substrate mimics is most likely triggered by the high-energy X-ray source deployed in the diffraction experiment. Unfortunately, only a single crystal was obtained from co-crystallization of AsR6 and FSPP **243** and the experiment could not be repeated using lower energies.

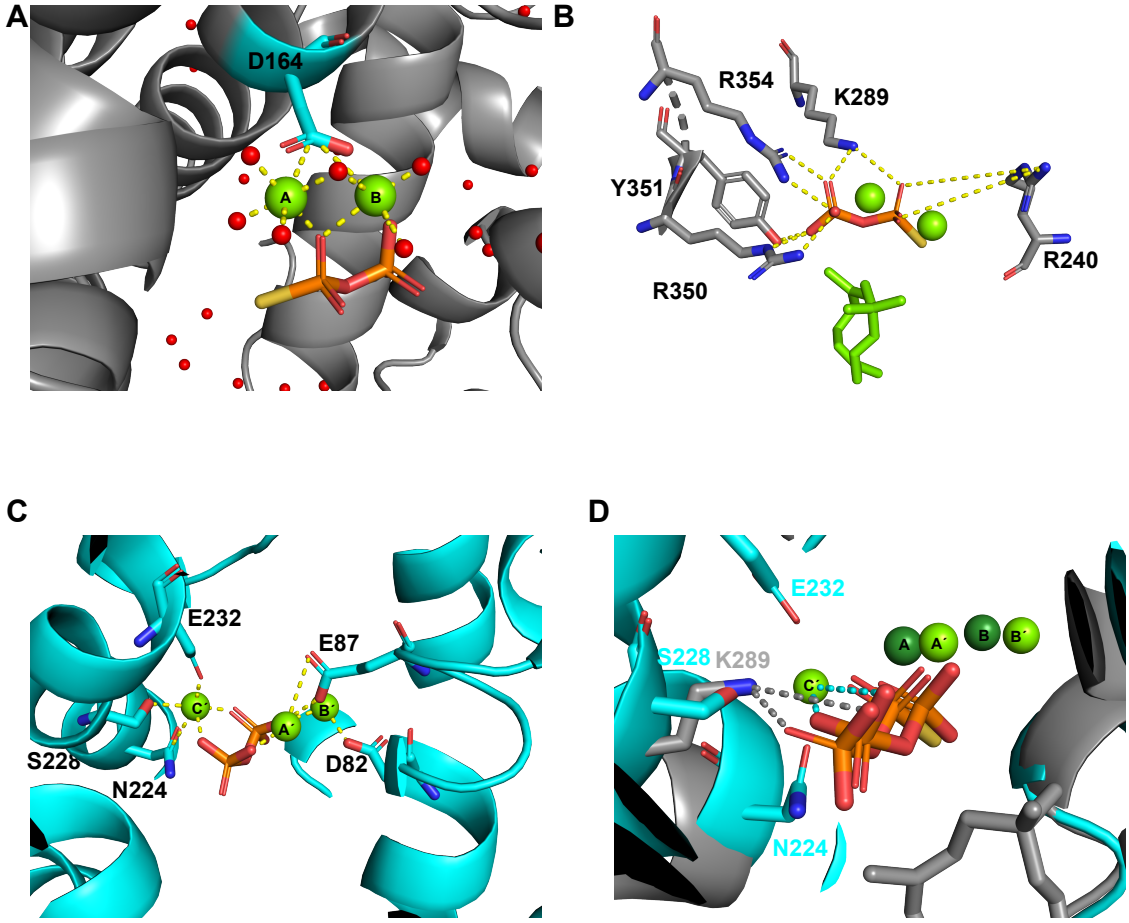



**Figure 4.9** Substrate binding site motif in AsR6: **A,** coordination of $Mg_A^{2+}$ and $Mg_B^{2+}$ by the carboxylate of D164 in AsR6; **B,** coordination of $SPP_i$ in the active site of AsR6; **C,** canonical pyrophosphate binding motif in selinadiene-synthase (PDB: 4okm); **D,** structural alignment of selinadiene-synthase (cyan) and *holo*-AsR6 (grey); dark green spheres = $Mg^{2+}$ ions of AsR6; light green spheres = $Mg^{2+}$ ions of selinadiene-synthase.

It is important to note that the exact orientation of the cyclized reaction product was also not unambiguously determined, and the displayed orientation is the one in best agreement with the determined electron density, ligand-protein distances and biosynthetic considerations.

Analysis of the electron density revealed the presence of two bound $Mg^{2+}$ ions ($Mg_A^{2+}$ and $Mg_B^{2+}$) at the entrance of the active site (Figure 4.8 B). Both ions are coordinated by the carboxylate moiety of a single aspartate (D164; Figure 4.9 A). The $Mg_A^{2+}$ ion is further coordinated by four water molecules in the active site and an oxygen atom of the $SPP_i$-moiety. Likewise, three water molecules and two oxygen atoms of the $SPP_i$-moiety coordinate the $Mg_B^{2+}$ ion (Figure 4.9 A). The single aspartate D164 is located in an aspartate-rich region IP**D**$_{164}$DVQI, that differs substantially from the typical DDxx(x)D/E-motif encountered in all other reported class I terpene cyclases.[39] Further analysis of the electron density revealed that five residues are involved in coordination of the $SPP_i$-moiety (Figure 4.9 B): R240, K289, R350, Y351 and R354, together with the two $Mg_{AB}^{2+}$ ions and water molecules, coordinate the thiolo-*S*-diphosphate at the active site entrance (Figure 4.9 B).

Notably, in AsR6 the pyrophosphate is only coordinated by two magnesium ions rather than three as found in all other class I terpene cyclases. Furthermore, the second conserved $Mg^{2+}$ binding motif, the NSE-triad, is absent. In other class I terpene cyclases, such as selinadiene synthase (SdS; PDB: 4okm) and trichodiene synthase (PDB: 1jfa), the NSE triad complexes a third $Mg_C^{2+}$ ion. In SdS, for instance, the first aspartate-rich motif **D**82DxxxE87 coordinates $Mg_{A'}^{2+}$ and $Mg_{B'}^{2+}$, while **N**224, **S**228, and **E**232 coordinate $Mg_{C'}^{2+}$ (bold letters indicate metal binding; Figure 4.9 C). Structural alignment of *holo*-AsR6 and SdS reveals that in AsR6 the complementary position of $Mg_{C'}^{2+}$ is occupied by the ε-ammonium of K289 (Figure 4.9 D). Sequence alignments of AsR6, PycR6, EupR3 and EupE[80] reveal that both motifs (D164D165; K289) are strictly conserved in all AsR6-type humulene synthases. This represents a novel substrate binding motif in class I terpene cyclases.

In order to identify amino acid residues putatively involved in catalysis of humulene formation all residues within 5 Å of the *in crystallo* cyclized reaction product were annotated. A total of 19 residues were identified (Figure 4.11 A). Mapping of these 19 residues onto the global sequence alignment of humulene synthases presented in Figure 4.10 revealed that 16/19 are strictly conserved in all four cyclases. Analysis of these 16 residues revealed homologies to residues in previously reported terpene cyclases, highlighting the structural relationship of AsR6 to canonical enzymes. *E.g.*, the conserved *C*-terminal residues R350 and Y351 form a conserved RY-dimer that is widely present in other (bacterial) cyclases (Figure 4.11 B).[222] Furthermore, the active sites of terpene cyclases are often lined with aromatic residues such as tyrosine, phenylalanine

and tryptophan that may play important roles in stabilizing intermediary carbocations thorough carbocation-$\pi$ interactions.[223] In AsR6 and homologs key tryptophan residue W138 is situated directly underneath the *in crystallo* cyclized reaction product (Figure 4.11 B; 3.3-6 Å average distance). Further analysis also revealed a conserved tyrosine Y344 that might be involved in the final deprotonation of the humulyl-cation.

```
AsR6       MPVTTPTKMATLTTKQMWQTIKDYFGDGFVTGSAPISYNVHTCDMQLQPDSGIHAASDGI      60
PycR6      --------MTETSLKMMWEGLKEYFGDGFAPGSAPFRNNVHYCDLHPAPK-----RDANF      47
EupR3      -------------MKEMWETMKAYFGDGFVPGSSPLRFNLHFCDMQDPAS-----TTDTI      42
EupE       -------------MKEIWETLKQYFGDGFVPGSAPLRYEMHFCDMKEPLN-----KKQSM      42
                        * :*: :* ******. **:*:  ::* **::   .         :

AsR6       HYGVQISEDSMPLFSIMGDTAAPPCTCHRVDEIVKHIDEFLERAPEALPDDGAITSGKPC      120
PycR6      RYGVEVEEEAMPLFSVLGDTCAPPCTCFEVTKLVEHIDSFIATASDYHADDYAITSGKN-      106
EupR3      RYGVELPESAMPLFSVLGDTCAPSCTCYDAPGVIQHIDRYLARAPVRHADDYTITSGKG-      101
EupE       RYGVEIPEDAMPLFSVLGDTCAPPCSCQDISGVIEHIDRYLENAPIRHADDYTITSGKD-      101
           :***:: *.:*****::***.** *:*     ::*** ::  *     ** :*****

AsR6       DTNPDQVSLYAMRDSLSWWVHWGGNLRPE-HYWKQIYIGFAAIPDDVQISPREFLDGTYR      179
PycR6      DATIDQVCFYAMRDTLSWWVHWGGGLRPSADYWKHVYVAFAAVPDDFQLPPRDFLSGQYT      166
EupR3      DVDINQVSLYIMRDTLSWWVHWGGSLTPN-NYWKLIYVAFAAIPDDVQIHPRDFIDGTYR      160
EupE       DEDINQVSLYIMRDTLSWWVHWGGSLHPN-NYWKLIYVAFAAIPDDVQVHPRDFIDGTYR      160
           *   :**.:* ***:*********.* *. .*** :*:.***:***.*: **:*:.* *

AsR6       YLGHTWDDCLSGLEEEGVSPDEIEFANMCMWRQMLTQWLEKADPELLPLLKGKISLMLQY      239
PycR6      LLGHTWDDCREGLLREGVPPETVKMAEMALWRQMLTQYLEKVDPTLRSLLVARTSMMTLY      226
EupR3      FLGHSWEDCRDGLLAEGVPPAQVKLAEMTLWRQMVTQYLEKVDPELRSLLVSKTSLMTAY      220
EupE       FLGHSWNDCLNGLLAEGVPSDQVKLAEMTLWRQMATQYIEKVDPGLRSLLVSKTTLMTQY      220
           ***:*:** .**  ***    :::*:* :**** **::**.** *  ** .: ::*  *

AsR6       RVLTANTLGCLALFMNATADPKDGPIHYADSSYEMEIASVAQCVTLDMAKEAMGILQGER      299
PycR6      RVQTANTLGCAALLLSSKGVNVNG-----LEDGALEVASIAQCVFLDISKEAMGILKGEK      281
EupR3      RVMTGNTLGCSVLLLASEGVVADG-----LTDNALEMASIAQCLSMDMAKEAMGVLEGEK      275
EupE       RVMTANTLGCAVLLLASEGVIVGD-----LDDGALEMASIAQCLSMDMAKEALGVLEGEK      275
           ** *.***** .*:: : .   ..       .  :*:**:***: :*::***:*:*:**:

AsR6       TEVVAGD----RAQRKRELRWIYVRCMQILESQPHAHMLRRYGSAGLHYVPMMDRYLERV      355
PycR6      TETITGN----RAQMKREMLWVYTRCMDYLGSQPNADFLRRYASAGMHYVPMMDRYRERM      337
EupR3      TETVAGNKFTSRQQLKRELRWLYARCMKYLEAQPNAHLLRRYASSGLHYVPMMDRYLERV      335
EupE       TETVAGD----RRQLKRELRWLYVRCMKYLDAQPHAEFLRRFASSGLHYVPMMDRYLERV      331
           **.::*:    * * ***: *:*.***. * :**:*.:***:.*:*:********* **:

AsR6       SGHTRFPIRDGAARILERFINRAELPKESEDINPNGRSLKVSAKMNGNGQLHHEVNGNAK      415
PycR6      NGHSRFAIPPAMARILLPYVKNPSALSGVLPIEDTTTGVSIHEAVNFLPS------ANAT      391
EupR3      RGNVRFPIREPVACLLEPFIKHELVAKVEKEDGEEHVEKILSETITVAP-----------      384
EupE       RGHIRFPIRESVARILEPFIKREPTPNKSGR-EADHIAQVVNEPITVAL-----------      379
                 *: ** *    * :*  :::.    .             :   :.

                      AsR6       LHLEAERPDVTTAVG   430
                      PycR6      V--------VS*---   394
                      NorR3      ---------------   384
                      EupE       ---------------   379
```

**Figure 4.10** Multiple sequence alignment of humulene synthases involved in the biosynthesis of tropolone sesquiterpenoids: bold residues = active site residues within 5 Å of the cyclized reaction product; cyan = residues involved in $Zn^{2+}$ binding; yellow = residues possibly involved in carbocation-$\pi$ interactions; green = aspartate-rich region involved in $PP_i$ coordination; purple = single conserved residue that differs in the two types of cyclases; red = conserved Y344; blue = RY-dimer; sequence alignment performed with ClustalW;[164] AsR6/PycR6 = α-humulene-synthases; EupE/EupR3 = 2*Z*-humulene synthases.

Two of nineteen residues in the active site are strictly conserved in at least three out of four humulene cyclases and only one residue, L285, is consistently different in the two classes of cyclases that give rise to either α-humulene **2** (AsR6, PycR6) or 2*Z*-humulene **7** (EupR3, EupE). While L285 is also conserved in PycR6, both EupE and EupR3 harbour a conserved M261 at this

position (Figure 4.10). Structurally, L285 points into the active site cavity and is in close vicinity to the bridging sulfur atom in FSPP (4.0 Å) and the C-2 (4.6 Å) and C-3 (3.9 Å) carbon atoms in the reaction product, respectively (Figure 4.11 C). In the absence of any other conserved changes in the active site it is reasonable to propose that L285/M261 are involved in controlling the stereochemical outcome of humulene formation.

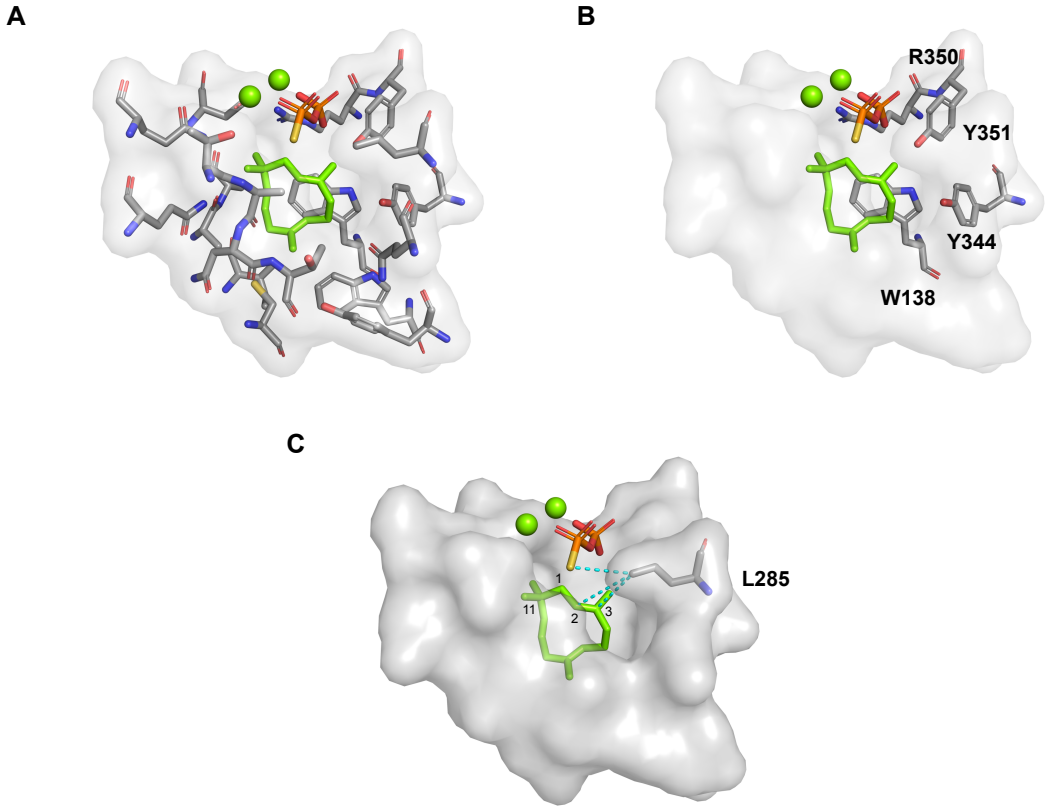



**Figure 4.11 A,** Overview of active site residues within 5 Å of the *in crystallo* cyclized reaction product; **B,** key amino acid residues with known/proposed function based on homology to canonical class I terpene cyclases; **C,** visualisation of key residue L285 in the active site.

## 4.5. Site-directed Mutagenesis

In order to link structural features to a catalytic function a series of site-directed mutagenesis experiments were conducted. Plasmid mutations were directed to investigate **a),** the new $PP_i$-binding motif in AsR6; **b),** the stereoselectivity of humulene formation; and **c),** possible residues involved in the final deprotonation to yield **2**/**7**. In total, eight point mutants were generated (Table 4.2). The desired mutations were either obtained as synthetic constructs from *Thermo Fischer Scientific* or manually introduced using site-directed ligase-independent (SLIM) mutagenesis.[224] All manually introduced mutations were verified by DNA sequencing of the corresponding expression plasmid (Figure 4.12 A).

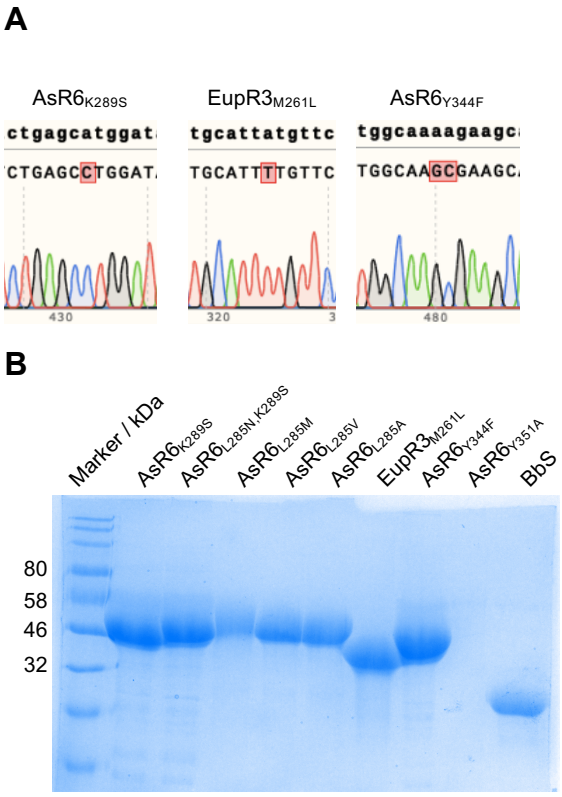



**Figure 4.12** Generation and purification of AsR6 and EupR3 point mutants: **A,** sequencing results of expression plasmids carrying manually introduced point mutations; **B,** SDS-PAGE analysis of purified recombinant AsR6 and EupR3 point mutants and bisabolene synthase (BbS); Marker = Color Prestained Protein Standard, Broad Range (11 - 245 kDa; *NEB*); displayed are samples prepared from the 500 mM elution fractions after $Ni^{2+}$ affinity chromatography.

All point mutants were expressed in *E. coli* BL21 on a 1 L-scale using identical conditions as for wild-type AsR6/EupR3 and purified in good yields from the crude lysate *via* $Ni^{2+}$ NTA (Figure 4.12 B; Table 4.2). Seven out of the total eight mutants were found to be soluble under the given expression conditions. $AsR6_{Y351A}$ was expressed in *E. coli* BL21 but found to be insoluble. All purified point mutants were incubated with suitable diphosphate precursors (FPP **3** or NPP **4**) and analyzed by GCMS.

**Table 4.2** Overview of site-directed mutagenesis experiments conducted in this work; n.d. = not determined.

| Mutation | Origin | Solubility | Yield | Probed catalytic activity |
|---|---|---|---|---|
| $AsR6_{K289S}$ | SLIM[224] | ✓ | 3.2 mg·$L^{-1}$ | $PP_i$-binding site |
| $AsR6_{L285N,K289S}$ | *Thermo Fischer* | ✓ | 3.1 mg·$L^{-1}$ | $PP_i$-binding site |
| $AsR6_{L285M}$ | *Thermo Fischer* | ✓ | 2.7 mg·$L^{-1}$ | *E*/*Z* selectivity |
| $AsR6_{L285V}$ | *Thermo Fischer* | ✓ | n.d. | *E*/*Z* selectivity |
| $AsR6_{L285A}$ | *Thermo Fischer* | ✓ | n.d. | *E*/*Z* selectivity |
| $EupR3_{M261L}$ | SLIM[224] | ✓ | 8.0 mg·$L^{-1}$ | *E*/*Z* selectivity |
| $AsR6_{Y344F}$ | SLIM[224] | ✓ | 7.5 mg·$L^{-1}$ | final deprotonation |
| $AsR6_{Y351A}$ | *Thermo Fischer* | x | - | final deprotonation |

### 4.5.1. Mutations Targeting the $PP_i$-Binding Site

Structural comparison of AsR6 with canonical terpene cyclases revealed a new pyrophosphate binding motif that consists of a binuclear magnesium cluster and a conserved K289 residue. In canonical cyclases, such as the previously described selinadiene synthase (SDS; PDB: 4okm), the position of K289 is occupied by a serine residue, S232, that is involved in coordination of the third magnesium atom, $Mg_C^{2+}$ (together with N228 and E236; Figure 4.13 A).

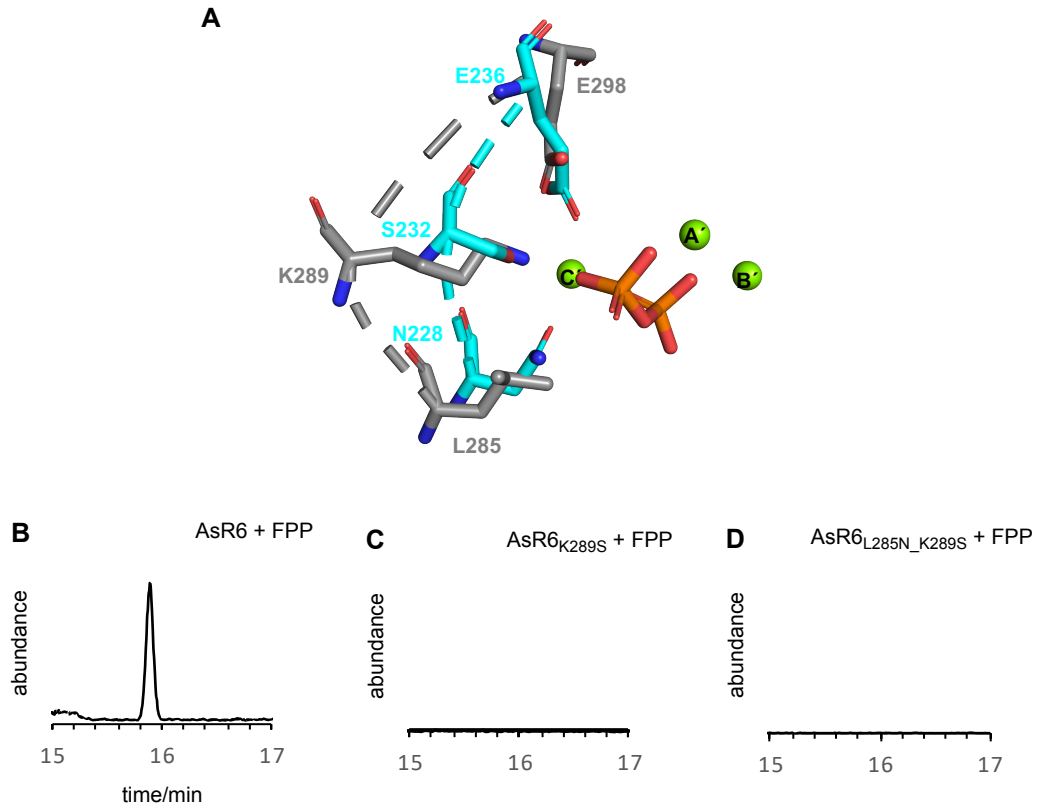



**Figure 4.13** Mutations targeting the $PP_i$ binding site in AsR6: **A,** canonical NSE triad coordinating $Mg_C^{2+}$ in selinadiene synthase (PDB: 4okm; cyan) and new binding motif in AsR6 (grey); **B-D** *in vitro* activity of AsR6, $AsR6_{K289S}$ and $AsR6_{L285N\_K289S}$ towards FPP **3**; displayed are GCMS TIC chromatograms of assay extracts from the incubation of FPP **3** with **B,** AsR6; **C,** $AsR6_{K289S}$; and **D,** $AsR6_{L285N\_K289S}$.

To verify that K289 replaces the function of $Mg_C^{2+}$ in AsR6 the K289S mutant was prepared and analyzed. Incubation of $AsR6_{K289S}$ with FPP **3** did not lead to formation of α-humulene **2** or any new product and confirmed that K289 is essential for catalysis (Figure 4.13 B and C).

In a follow-up experiment the canonical NSE-motif was reintroduced into AsR6, to assess whether $Mg_C^{2+}$-binding could be restored in the K289S mutant, and cyclase activity could be regained. Structural comparison with the NSE-triad in selinadiene synthase (Figure 4.13 A) revealed that in AsR6 the complementary positions of N228, S232 and E236 are occupied by L285, K289 and E298. Notably, the required glutamate residue that complements the NSE-triad is already present in the complementary position in native AsR6 (E298; compare Figure 4.13 A).

However, creation of the double mutant AsR6$_{L285N_K289S}$ and incubation of AsR6$_{L285N_K289S}$ with FPP **3** did not result in formation of α-humulene **2** or any other cyclase product (Figure 4.13 D). Together the two mutations (K289S; L285N_K289S) validate that K289 is essential for humulene formation and that removal of K289 likely prevents correct substrate recognition or activation.

### 4.5.2. Mutations Targeting Stereoselectivity

Global sequence alignment of the active site residues in AsR6, PycR6, EupE and EupR3 identified a single amino acid position that was consistently different in the two classes of cyclases that produce either α-humulene **2** or 2*Z*-humulene **7**. Position 285 (in AsR6) is the only active site residue which is not highly conserved. In fact, it appears to vary consistently with product stereochemistry. In AsR6 and PycR6, which produce 2*E*-humulene **2**, this position is leucine. In EupR3 and EupE, which produce 2*Z*-humulene **7**, this position is methionine. In the absence of other conserved changes this position was considered as the most likely candidate to control the observed stereoselectivities.

To start our investigation into this position a structural model of EupR3 was created using the Phyre2 One-2-One threading tool.[148][7] This model structure was aligned with AsR6 (Figure 4.14 A). Given the high sequence homology between AsR6 and EupR3, and because AsR6 was used as the template for modeling, not surprisingly both structures are practically identical (RMSD value = 1.27 Å over 352 residues). L285 and M261 take up similar conformations in the respective structures, pointing into the active site cavity and being located directly beneath the $SPP_i$ cleavage site (Figure 4.14 B).

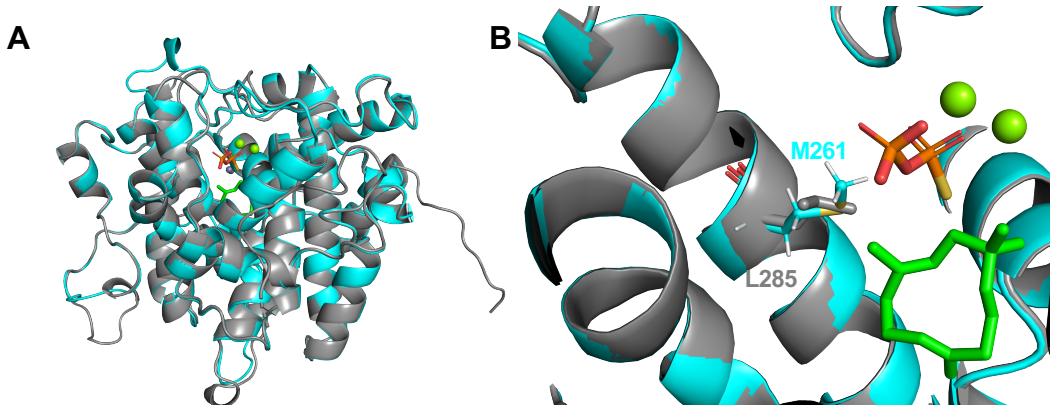



**Figure 4.14** Comparison of the AsR6 crystal structure and structural model of EupR3: **A,** structural alignment of AsR6 (grey) and EupR3 (cyan); **B,** visualization of L285/M261 in the active site.

To investigate the influence of this position cross-convergent point mutants of AsR6 and EupR3 were created: L285 in AsR6 was mutated to methionine and M261 in EupR3 was mutated to

[7] Template for modelling: AsR6; Model built by DR. PEER LUKAT.

leucine, yielding AsR6$_{L285M}$ and EupR3$_{M261L}$, respectively. Both mutants were then systematically assayed with FPP **3**, (*rac*)-NPP **4** and (3*S*)-NPP **4b**.

*In vitro activity of AsR6$_{L285M}$*

Incubation of AsR6$_{L285M}$ with FPP **3** led to formation of the non-native product 2*Z*-humulene **7** as the dominant product (> 80%; based on peak integration). Production of the native product α-humulene **2** was significantly reduced (Figure 4.15 A and B). The single (!) introduced point mutation, L285M, thus led to a near complete switch in stereoselectivity. Next, the activity of AsR6$_{L285M}$ was probed with (*rac*)-NPP **4**. Analysis of assay extracts showed, as for FPP **3**, major production of the non-native product 2*Z*-humulene **7**. α-humulene **2** was only observed in traces (Figure 4.15 C). A similar product distribution was observed for conversion of (3*S*)-NPP **4b**, although no traces of the native product α-humulene **2** were detected here (Figure 4.15 D). However, as conversion of the deployed batch of (3*S*)-NPP **4b** was poor in general it has to be noted that trace production of α-humulene **2** would probably not have been detected by GCMS.

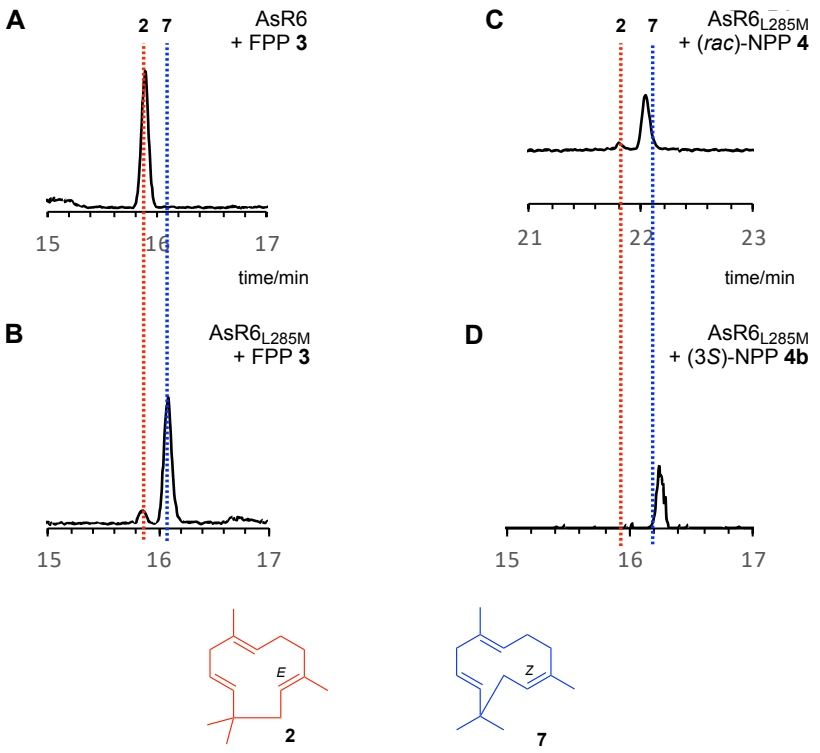



**Figure 4.15** Conversion of: **A,** FPP **3** by AsR6; **B,** FPP **3** by AsR6$_{L285M}$; **C,** (*rac*)-NPP **4** by AsR6$_{L285M}$; **D,** (*3S*)-NPP **4b** by AsR6$_{L285M}$; displayed are the GCMS TIC chromatograms of n-hexane extracts of the corresponding assays; the retention time of the respective humulene isomer **2**/**7** is colour-coded in red (**2**) and blue (**7**).

*In vitro activity of EupR3$_{M261L}$*

Identical *in vitro* assays were performed with the cross-convergent mutant EupR3$_{M261L}$. First, the activity towards the substrate FPP **3** was assessed. EupR3$_{M261L}$ exclusively produces the non-native terpene product α-humulene **2** (Figure 4.16 A and B). Production of the native product, 2*Z*-

humulene **7**, is completely abolished in this mutant and no traces of **7** were ever observed in the extracts of these assays. Again, the single introduced point mutation is sufficient to switch the stereoselectivity.

Next, the activity of EupR3$_{M261L}$ towards nerolidyl pyrophosphate was tested. Incubation with (*rac*)-NPP **4** affords a mixture of α-humulene **2** and 2*Z*-humulene **7**, with **2** and **7** being produced at similar levels (~ 1 : 1 mixture; Figure 4.16 C). Obviously, the chemical nature of NPP **4** compensates the formerly lost capacity to produce the native terpene product 2*Z*-humulene **7**. Incubation with the single enantiomer (3*S*)-NPP **4b** affords – as already observed for all other enzymes in this study – an identical product distribution as for incubation with (*rac*)-NPP **4**. Intriguingly, the introduced M261L mutation into EupR3 leads to more severe changes in the product distribution than the introduction of the cross-convergent mutation L285M in AsR6. While L285M retains at least some capacity to convert FPP **3** to the native product, α-humulene **2**, M261L is completely deficient in the production of 2*Z*-humulene **7** and only by changing the chemical nature of the substrate the natural activity can be partially regained. Together the assays identify the L285/M261 position to have gatekeeping control over the stereoselectivity observed in humulene formation.

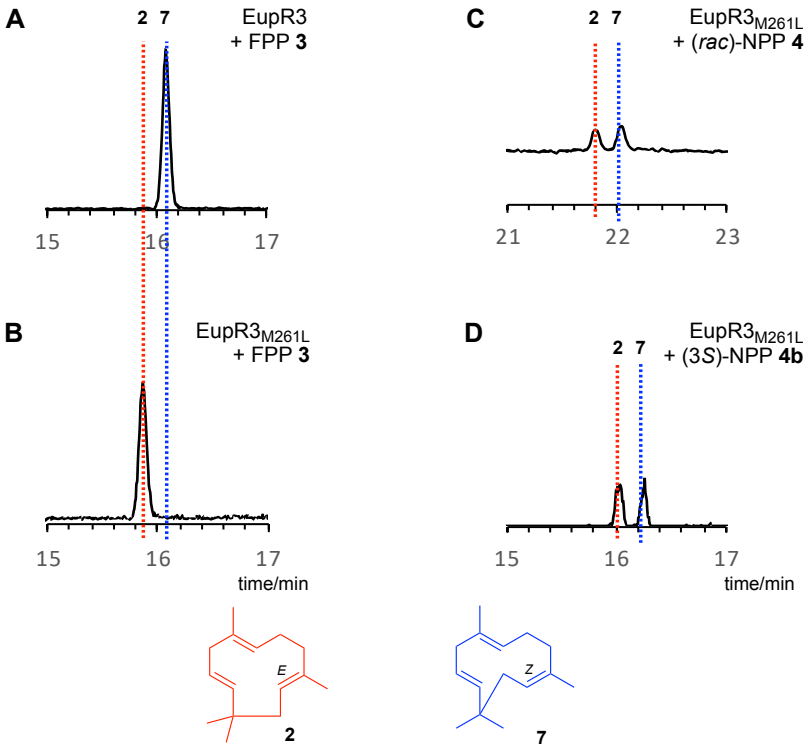



**Figure 4.16** Conversion of: **A,** FPP **3** by EupR3; **B,** FPP **3** by EupR3$_{M261L}$; **C,** (*rac*)-NPP **4** by EupR3$_{M261L}$; **D,** (3*S*)-NPP **4b** by EupR3$_{M261L}$; displayed are the GCMS TIC chromatograms of n-hexane extracts of the corresponding assay; the retention time of the respective humulene isomer **2**/**7** is colour-coded in red (**2**) and blue (**7**).

*Additional mutations*

Two additional mutations at the key L285 position in AsR6 were probed, to test whether exchange of leucine for other residues than methionine has similar effects on the stereoselectivity. The two point mutations AsR6$_{L285A}$ and AsR6$_{L285V}$ were prepared and tested with the native substrate farnesyl diphosphate **3**. Both mutants exclusively convert **3** into the native product α-humulene **2** (Figure 4.17 A and B). No other products were observed by GCMS and no significant differences in production titres were observed when compared to the AsR6 wild-type. The cross-convergent mutations in EupR3 were planned but not completed in the scope of this thesis.

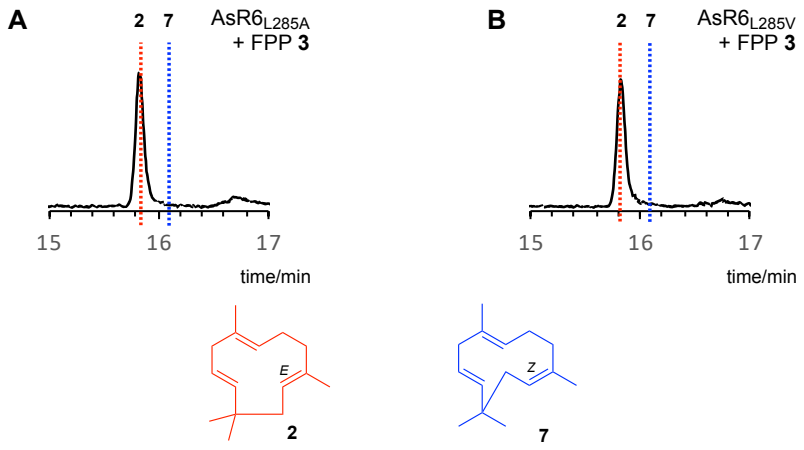


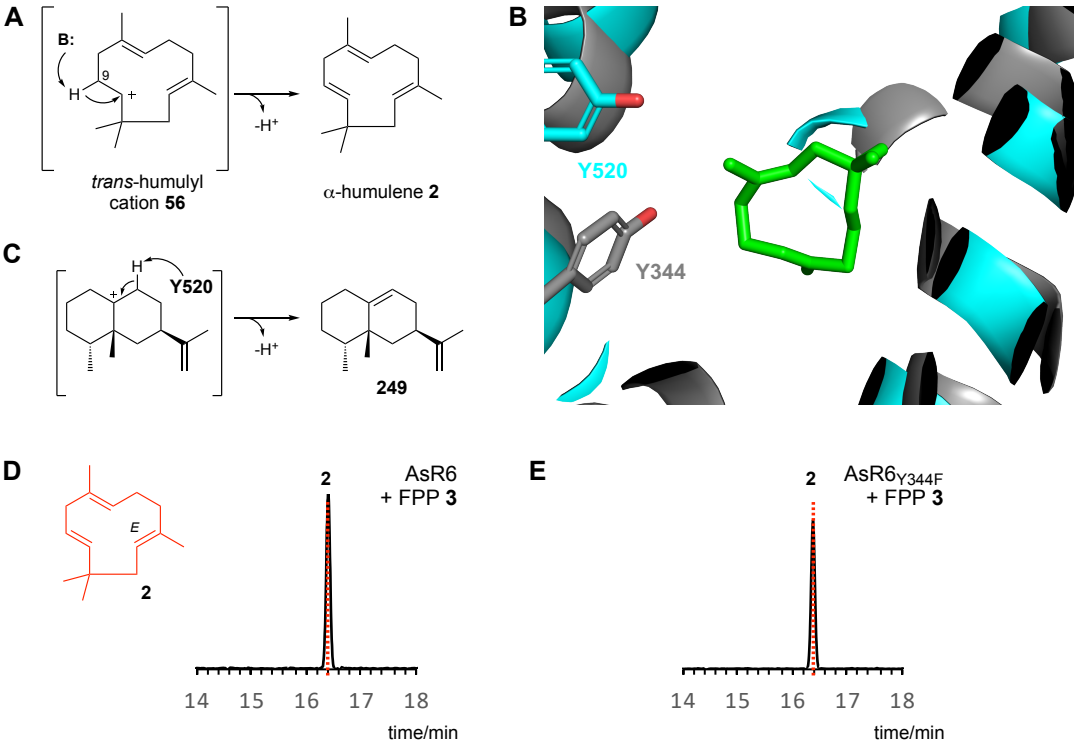

**Figure 4.17** Conversion of: **A,** FPP **3** by AsR6$_{L285A}$; and **B,** FPP **3** by AsR6$_{L285V}$; displayed are the GCMS TIC chromatograms of n-hexane extracts of the corresponding assay; the retention time of the respective humulene isomer **2**/**7** is colour-coded in red (**2**) and blue (**7**).

### 4.5.3. Mutation of Candidate Residues Involved in Final Deprotonation

To complete our investigations into the mechanism of humulene formation, considerations were made regarding possible residues in AsR6 that could catalyse the final deprotonation required to form α-humulene **2**. Mechanistically, the intermediary *E,E*-humulylcarbocation **56** needs to undergo selective deprotonation at C-9 (Scheme 4.8 A). As carbocations are highly acidic by nature, thought to be as acidic as sulfuric acid, only a weak base is required to trigger this deprotonation.[225]

Investigations into the final deprotonation step were not of primary interest in the course of this study. However, during analysis of the AsR6 crystal structure a structural alignment was performed with *epi*-aristolochene synthase (PDB: 5eat; Scheme 4.8 B; RMSD backbone atoms: 4.69 Å). SIEGEL *et al*. proposed a key tyrosine Tyr520 as the likely base to catalyse the final deprotonation during the biosynthesis of *epi*-aristolochene **249** by tobacco *epi*-aristolochene synthase (Scheme 4.8 C).[226] The structural alignment of AsR6 with *epi*-aristolochene synthase identified a similar tyrosine residue in the AsR6 active site (Y344). In the absence of other obvious residues that could function as a base in the AsR6 active site this residue was chosen for

mutation. However, while expression of *asR6*$_{Y344F}$ afforded soluble polyhistidine-tagged enzyme, incubation with FPP **3** did not result in any differences in the product distribution when compared to wild-type AsR6 (Scheme 4.9 D and E).




**Scheme 4.8** Final deprotonation in the biosynthesis of α-humulene **2**: **A,** deprotonation at C-9 is required for **2**-formation; **B,** structural alignment of *epi*-aristolochene synthase (cyan; PDB: 5eat) and AsR6 (grey) with respective Y520/Y344 highlighted; **C,** final deprotonation in the biosynthesis of *epi*-aristolochene **249**; **D,** incubation of AsR6 with FPP **3**; **E,** incubation of AsR6$_{Y344F}$ with FPP **3**; displayed are GCMS TIC chromatograms of assay extracts.

## 4.6. Conclusion, Discussion and Outlook

In this work the mechanism of humulene formation in AsR6 and EupR3 was investigated. Together with the Helmholtz Institute for Infection Research the first crystal structure of a *bona fide* humulene synthase (AsR6) was obtained and described as a new type of class I terpene cyclase with a unique pyrophosphate binding motif. Structure-function guided site-directed mutations furthermore revealed a key amino acid position (L285) that controls the stereochemical outcome of humulene cyclization.

### 4.6.1. The AsR6 Crystal Structure

The crystal structure of AsR6 was obtained in the unliganded state and in complex with thiolo-*S*-diphosphate **8** and an *in crystallo* cyclized reaction product. While the overall structural fold was found to be similar to previously described class I terpene cyclases (all α-helical bundle fold),

several structural features of AsR6 differ. AsR6 (and related humulene cyclases) are unique in respect to **a),** a new $PP_i$-binding motif that comprises a *bi*nuclear magnesium cluster and a K289-residue in which the ε-ammonium appears to replace the third $Mg^{2+}$ ion; and **b),** a $Zn^{2+}$ binding cluster of unknown function.

The new $PP_i$-binding motif is the most intriguing feature of AsR6. To-date, all reported structures of class I terpene cyclases share the common feature of a *tri*nuclear magnesium cluster.[39] This feature is conserved across sesquiterpene cyclases from bacteria (*e.g., Streptomyces coelicolor epi*-isozizaene synthase; PDB: 4lz0; Figure 4.18 A),[227] fungi (*e.g.*, *Fusarium sporotrichiodes* trichodiene synthase; PDB: 2q9y; Figure 4.18 B)[228] and plants (*e.g., Gossypium arboreum δ*-cadinene synthase; PDB: 3g4f; Figure 4.18 C)[51] - but also found in other terpene cyclases such as the diterpene cyclase CotB2 (PDB: 6ggi; Figure 4.18 D).[219]

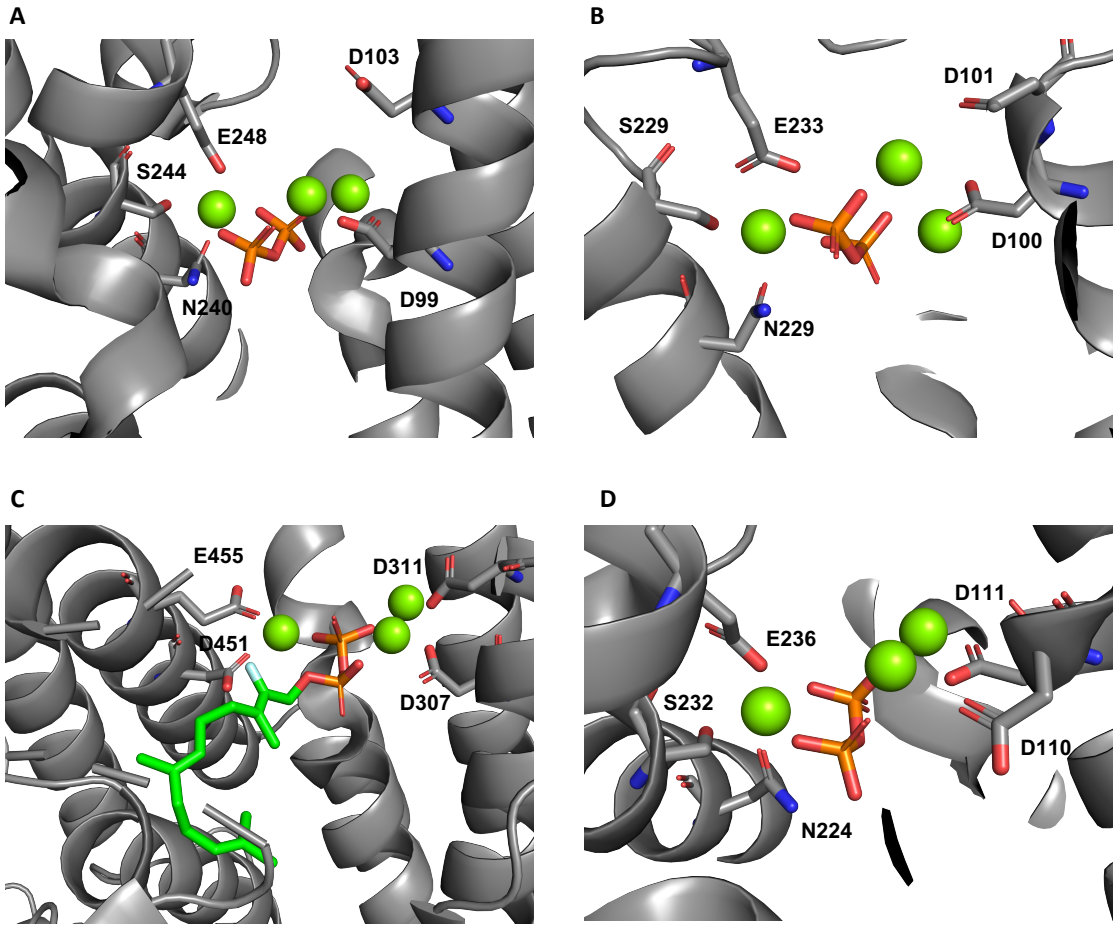



**Figure 4.18** Conserved trinuclear $Mg^{2+}$ clusters in class I terpene cyclases: **A,** crystal structure of *epi*-izozizaene synthase in complex with $PP_i$ (PDB: 4lz0); **B,** crystal structure of trichodiene synthase in complex with $PP_i$ (PDB: 2q9y); **C,** crystal structure of *δ*-cadinene synthase in complex with 2-fluorofarnesyl diphosphate (PDB: 3g4f); **D,** crystal structure of the diterpene cyclase CotB2 in complex with $PP_i$ (PDB: 6ggi); green spheres = $Mg^{2+}$ ions.

While the *tri*nuclear cluster is generally strictly conserved in class I terpene cyclases, minor modifications in the $Mg^{2+}$ ion binding motifs are well-documented. For instance, the diterpene

cyclase CotB2 harbours a modified DDxx(x)D/E motif, which is `abbreviated´ to a **D**110**D**111xD-motif (Figure 4.18 D).[219] CHRISTIANSON *et al*. reported that in (+)-*δ*-cadinene synthase (DCS) the second aspartate-rich region (the `NSE/DTE´ triad) is replaced by a second **D**451Dxx(x)**E**455 motif, that coordinates the third magnesium atom, $Mg_C^{2+}$ (Figure 4.18 C).[51] Likewise, minor modifications of the NSE motif are also well-documented. PETERS *et al.* for instance reported on a modified NSE motif in the abietadiene synthase from *Abies grandis*, in which the conserved central serine residue is exchanged for a glycine residue.[229]

With a *bi*nuclear $Mg^{2+}$ ion cluster AsR6 thus significantly expands the structural space of class I terpene cyclases. Notably, both binding sites of the $Mg^{2+}$ ions are substantially different in AsR6 when compared to all other class I terpene cyclases: The DDxx(x)D/E motif is replaced by a **D**164DVQIS motif (missing the final D/E residue); and the NSE motif, that normally coordinates the third $Mg^{2+}$ ion, is replaced by the ε-ammonium of a lysine residue, K289. Site-directed mutagenesis was successfully deployed to confirm the essential role of K289 in catalysis – based on the structural data obtained for AsR6 and in combination with the introduced mutations it is plausible to assume that K289 replaces the third $Mg^{2+}$ ion typically encountered in class I terpene cyclases and is essential for substrate coordination.

The newly identified $PP_i$-binding motif also explains the difficulty in predicting the function of AsR6 *in silico* in the first place (Section 2.1). Terpene cyclases in general show a low degree of sequence homology at the primary sequence level and *in silico* identification of these enzymes strongly relies on recognition of the conserved aspartate rich regions - that are not present in AsR6. For example, the sesquiterpene cyclases trichodiene synthase and aristolochene synthase (both fungal canonical class I terpene cyclases) only share 4.3% sequence identity, but can be identified by conserved domain algorithms such as NCBI´s conserved domain tool, whereas AsR6 is not.[147]

Regarding the second unique feature of AsR6, the $Zn^{2+}$ cluster, sequence alignments with other class I terpene cyclases revealed that the cluster is only encoded within humulene synthases involved in tropolone sesquiterpenoid biosynthesis and not a common feature of class I terpene cyclases (compare Figure 4.10). However, in the absence of mutational studies the function of the $Zn^{2+}$ cluster remains elusive. As its location is distal from the active site it is reasonable to assume that it is not involved in the catalytic mechanism or substrate coordination. Alternatively, $Zn^{2+}$ clusters have been reported to contribute to the overall protein stability and can majorly influence physical parameters such as the melting point.[230] It is plausible to assume that in AsR6 the $Zn^{2+}$ ion might also be contributing to overall protein stability – but, particularly from an evolutionary viewpoint, it is intriguing how and why this class of cyclases acquired this unusual feature. In

future work site-directed mutation of the binding motif (3x Cys, 1x His) should thus investigate the role of this cluster or assay AsR6 in presence/absence of $Zn^{2+}$.

### 4.6.2. Stereochemical Aspects of Humulene Formation

A particular focus of this work was directed to elucidate factors that control the stereoselective formation of either α-humulene **2** or its geometric isomer 2*Z*-humulene **7**. Initially, structure-sequence comparisons of AsR6 and other tropolone sesquiterpenoid humulene synthases identified a key position in α-humulene **2** producing synthases (L285), that was consistently changed to a methionine M261 in humulene synthases producing **7** (EupR3). Cross-convergent mutations in AsR6 and EupR3 (yielding the mutants $AsR6_{L285M}$ and $EupR3_{M261L}$) and extensive *in vitro* assays probing the conversion of FPP **3**, (*rac*)-NPP **4** and (3*S*)-NPP **4b** were then used to investigate the mechanism. The results of the conducted experiments are summarized in Table 4.3 and, based on this, a possible mechanism of humulene formation is discussed (*vide infra*).

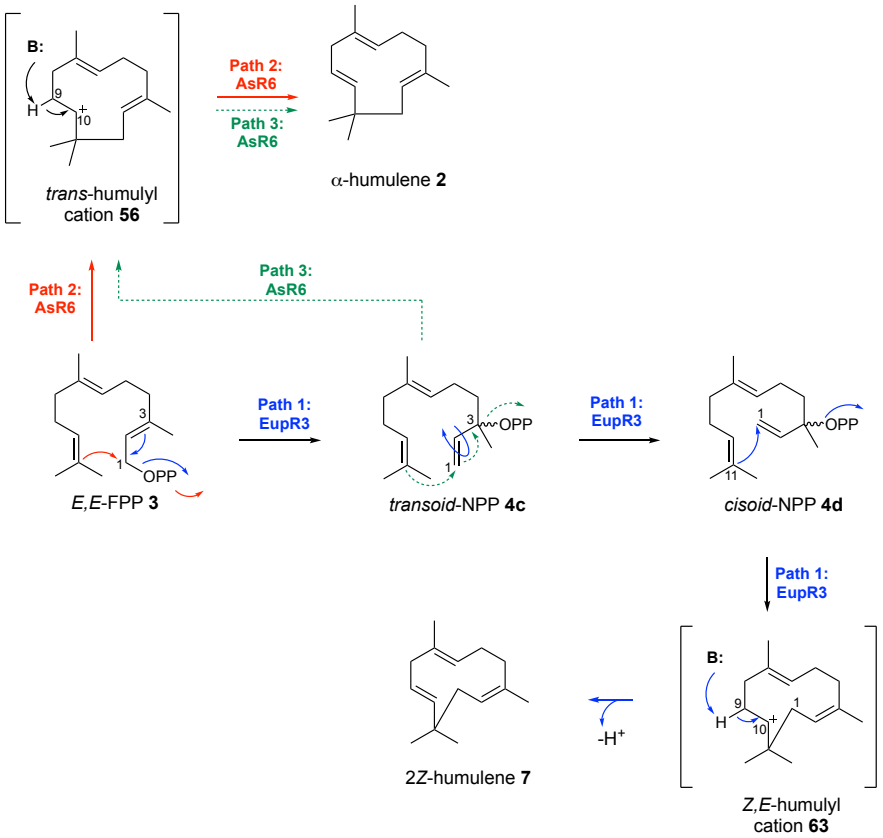
**Table 4.3** Summary of results from the conversion of FPP **3**, (*rac*)-NPP **4**, (3*S*)-NPP **4b** by wild-type enzymes and mutants; 1 % = trace production.

| Enzyme | aa at position 261/285 | Conversion of FPP 3 to | | Conversion of (*rac*)-NPP 4 to | | Conversion of (3*S*)-NPP 4b to | |
|---|---|---|---|---|---|---|---|
| | | α-humulene **2** | 2*Z*-humulene **7** | α-humulene **2** | 2*Z*-humulene **7** | α-humulene **2** | 2*Z*-humulene **7** |
| AsR6 | L | 100 % | - | 99 % | 1 % | 99 % | 1 % |
| EupR3 | M | - | 100 % | - | 100 % | - | 100 % |
| $EupR3_{M261L}$ | L | 100 % | - | 50 % | 50 % | 50 % | 50 % |
| $AsR6_{L285M}$ | M | 20% | 80 % | 1 % | 99 % | 1 % | 99 % |

Initially the wild-type enzymes AsR6 and EupR3 were probed with the native substrate FPP **3** and (*rac*)-NPP **4**. In the case of EupR3 both FPP **3** and NPP **4** were converted to the same (and only) product, 2*Z*-humulene **7**. This is in agreement with EupR3 catalysing a two-step process, where FPP **3** is first isomerized to the *cisoid* form of NPP **4d**, and then cyclized to 2*Z*-humulene *via* an intermediary *Z*,*E*-humulyl cation **63** (Scheme 4.9; Path 1 - blue arrows). This proposed pathway accommodates both the findings in this study and the previously reported findings on the generation of *Z*-alkene containing terpenoids.[48]

α-Humulene **2** formation by AsR6 is more complex as two possible routes to **2** are conceivable. The first possibility proceeds *via* the initial 1,11-cyclization of FPP **3**, resulting in the direct formation of the *trans*-humulyl cation **56**. Deprotonation from C-9 then affords α-humulene **2** (Scheme 4.9; Path 2 – red arrows). Alternatively, AsR6 could – like EupR3 – catalyse the isomerization of FPP **3** to NPP **4** in a first step. 1,11-Cyclization from the *transoid* form of NPP **4c** would then also result in formation of the *trans*-humulyl cation **56** (Scheme 4.9; Path 3 – green

arrows). In the case of Path 3, where NPP **4** is a true intermediate, incubation of AsR6 with FPP **3** and NPP **4** should lead to identical product distributions (only production of the native product α-humulene **2**). However, upon incubation of AsR6 with (*rac*)-NPP **4**, traces of the non-native product 2*Z*-humulene **7** were observed. The small but distinctive difference in the product distributions suggests, that NPP **4** is a substrate for AsR6, but not an intermediate in the conversion of FPP **3** to α-humulene **2**. In AsR6, α-humulene **2** formation thus likely proceeds *via* Path 2 and *via* the direct cyclization of FPP **3** without the intermediacy of NPP **4**. It is important to emphasize – again - that the *trans*-humulyl cation **56** is energetically favoured over the *Z,E*-humulyl cation **63** and EupR3 is thus required to actively prevent formation of **56**. Also, conversion of both FPP **3** and NPP **4** into the native products by both AsR6 and EupR3 demonstrates that the stereochemical outcome of the reaction is decided by the enzyme and not the chemical nature of the substrate.




**Scheme 4.9:** Proposed mechanism of α-humulene **2**- and 2*Z*-humulene **7**-formation by AsR6 and EupR3; B = general base.

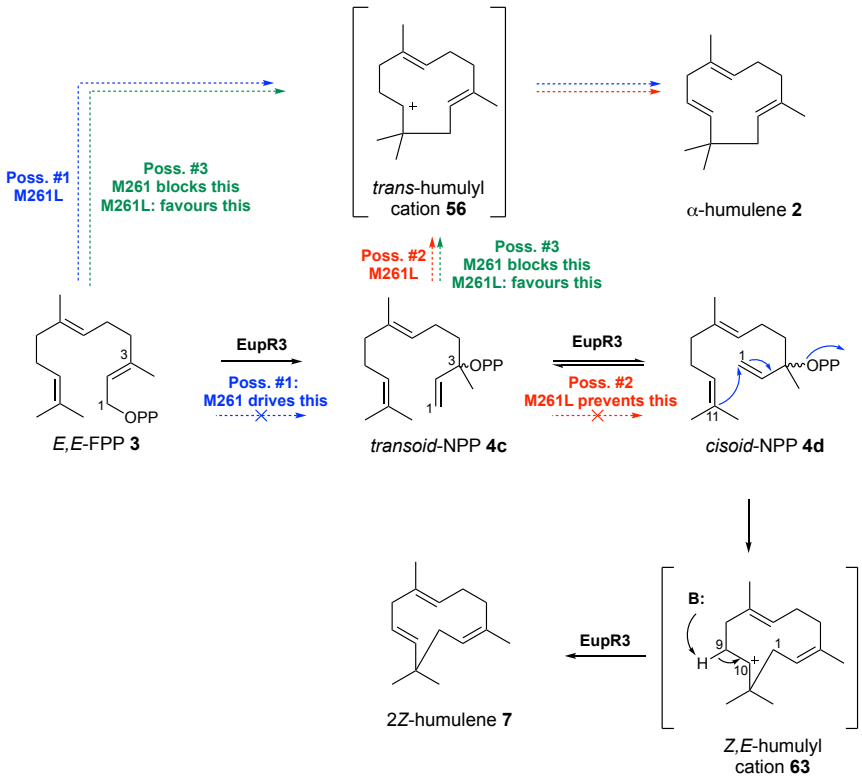
Next, the effect of the cross-convergent mutations has to be taken into consideration. Incubation of EupR3$_{M261L}$ with FPP **3** led to the exclusive production of the non-native product α-humulene **2** and production of 2*Z*-humulene **7** was entirely abolished. It is worthwhile to marvel at the fact that a single (!) point mutation is sufficient to lead to a complete switch in the observed stereoselectivity. The amino acid residue at this position thus has – in the case of EupR3 – absolute

(gatekeeping) control over the stereoselectivity of humulene formation. Three reasonable explanations for the function of the M261-residue were considered in this thesis:

a) M261 is involved in the first distinct chemical step of 2*Z*-humulene **7** formation and controls the isomerization of FPP **3** into NPP **4** (Scheme 4.10; Poss. #1 – blue arrows). In absence of methionine the inserted mutation M261L would then lead to α-humulene formation *via* 1,11-cyclization of FPP **3**, as observed for native AsR6.

b) M261 is involved in controlling the interconversion of *transoid* NPP **4c** into *cisoid* NPP **4d** (*e.g.,* by controlling the equilibrium between **4c** and **4d**). The inserted mutation M261L would then block this process/favour formation of **4c** and humulene formation proceeds *via* 1,11-cyclization of *transoid* NPP **4c** (Scheme 4.10; Poss. #2 – red arrows)

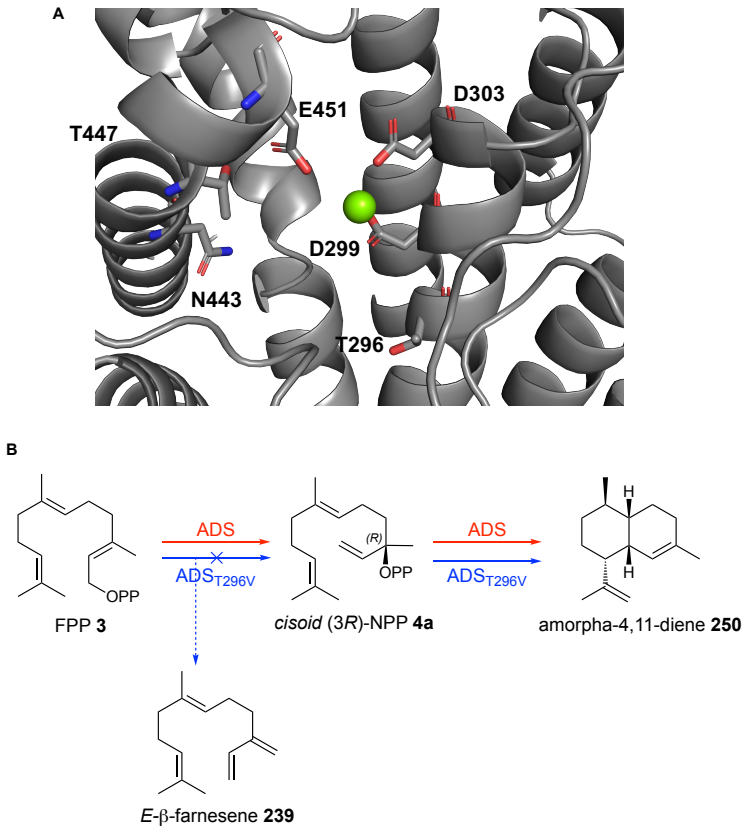
c) M261 blocks formation of the energetically favoured *trans*-humulyl cation **56**, thus directing the pathway towards 2*Z*-humulene **7**-formation. The inserted mutation M261L then leads to formation of α-humulene **2** as the energetically favoured product (Scheme 4.10; Poss. #3 – green arrows).




**Scheme 4.10:** Possible modes of action of the key M261 residue in EupR3.

Precedence in the literature suggested that the M261 residue might be involved in the initial isomerization process (Poss. #1 or Poss. #2): CANE *et al*. investigated the formation of amorpha-4,11-diene **250** by amorphadiene-synthase (ADS). A single point mutation, T296V, successfully

abolished **250**-formation and instead incubation with FPP **3** exclusively afforded acyclic reaction products, such as β-farnesene **239**. Although no crystal structure for ADS has yet been reported, modelling of ADS suggests that T296 is situated directly beneath the first aspartate-rich motif, DDxx(x)D/E (Scheme 4.11 A), and therefore close enough to the substrate to affect ionization steps. Cyclase activity of the T296V mutant was only regained upon incubation with (3*R*)-NPP **4a** (Scheme 4.11 B).[202] The findings by CANE *et al*. suggested that $ADS_{T296V}$ was defective in the initial allylic isomerization of FPP **3** to NPP **4**.[202]

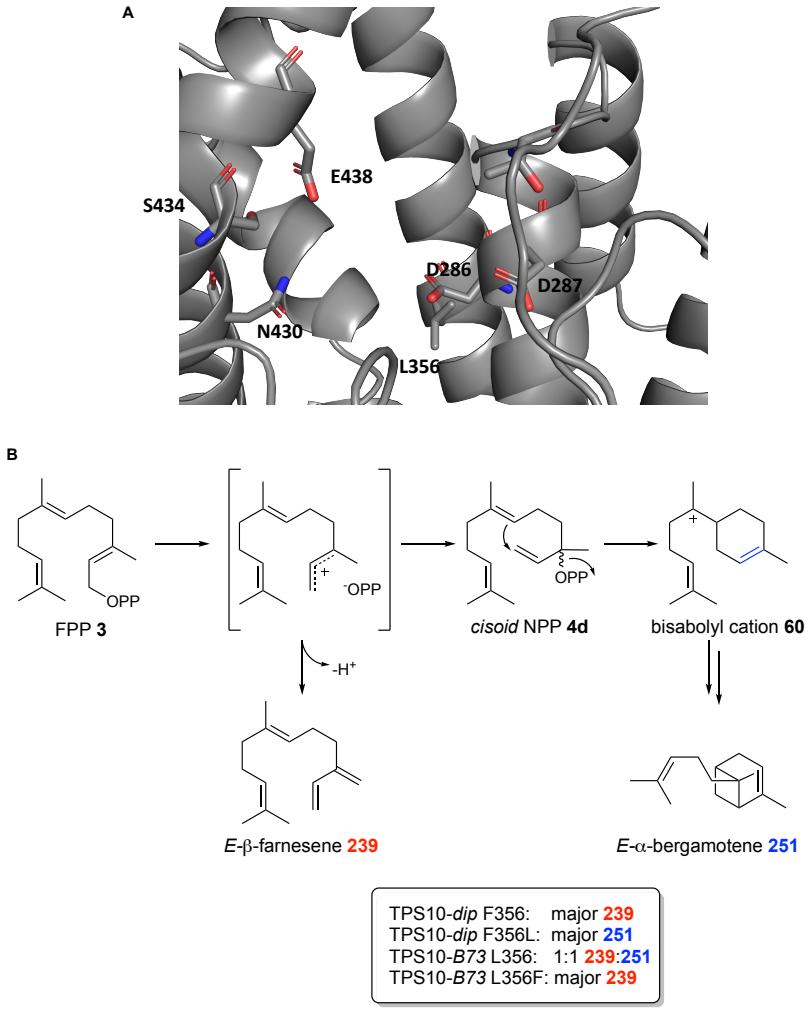



**Scheme 4.11** Formation of amorpha-4,11-diene **250** by ADS: **A,** structural model of ADS; green sphere = $Mg^{2+}$ ion; SWISS-Model[149] was used to create the ADS structural model using 5-epi-aristolochene synthase (PDB: 4rnq) as template; QMEAN value of the model = -2,06 and GMQE value = 0.76; indicating adequate reliability of the model; **B,** *in vitro* activity of ADS wildtype and $ADS_{T296V}$.

However, identical assays in this study using NPP **4** as an alternative substrate with $EupR3_{M261L}$ did not lead to reconstitution of the wild-type enzyme activity. Instead of reconstituted formation of 2*Z*-humulene **7**, $EupR3_{M261L}$ converted (*rac*)-NPP **4** to a ~ 1 : 1 mixture of 2*Z*-humulene **7** and α-humulene **2**. These findings eliminated the possibility in which M261 is controlling the isomerization of FPP **3** to *transoid* NPP **4c** (Poss. #1), as in this case incubation with NPP **4** should have reconstituted wild-type activity. However, all assays conducted in this study do not allow differentiation between Poss. #2 (controlling the interconversion of *transoid* and *cisoid* NPP **4cd**; *e.g.,* by changing the equilibrium of both species in the active site) and Poss. #3 (controlling the

carbocation trajectories; *e.g.,* by changing the relative rates of cyclization of *cisoid* and *transoid* NPP **4cd**).

Only one other similarly influential residue has been reported in the literature. DEGENHARDT *et al.* investigated formation of *E*-α-bergamotene **251** and *E*-β-farnesene **239** in several species of maize plants.[231] *E*-β-farnesene **239** is a degradation product of the *transoid* FPP **3** and derived from **3**-dephosphorylation and subsequent deprotonation (Scheme 4.12 B). *E*-α-bergamotene **251**, on the other hand, is derived from the bisabolyl cation **60** and biosynthetically **60**-formation must proceed through the *cisoid* NPP **4d** intermediate (Scheme 4.12 B).




**Scheme 4.12** Stereochemical control of *E*-α-bergamotene biosynthesis in maize terpene synthase: **A,** structural model of maize terpene synthase (TPS10); SWISS-Model[149] was used to create the TPS10 structural model using 5-*epi*-aristolochene synthase (PDB: 4di5) as template; QMEAN value of the model = -2,16 and GMQE value = 0.78, indicating adequate reliability of the model; **B,** influence of the key F356/L356 amino acid residue in maize terpene synthase TPS10.

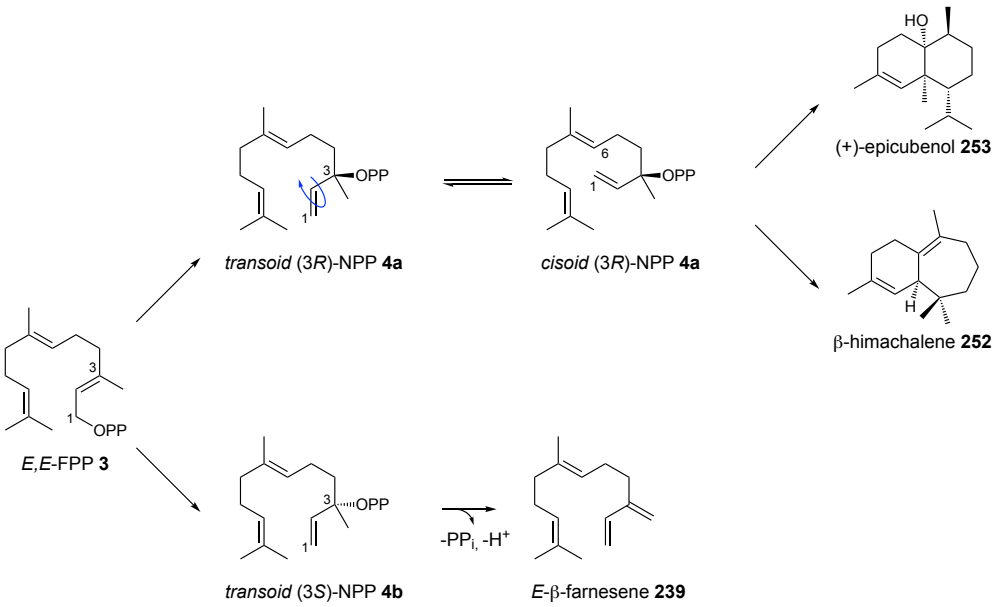
Different orthologous maize terpene synthases (TPS10) afford these two compounds in varying ratios, ranging from 0.1:1.0 to 1.4:1. Site-directed mutagenesis of TPS10-*dip* (*Zea diploperennis*) and TPS10-*B73* (*Zea mays mays*) afforded the cross-convergent mutants F356L and L356F, that

were each sufficient to invert the product ratios of **239**:**251** in the respective TPS10 ortholog.[231] No crystal structure exists for any of the investigated TPS10 orthologs but homology models suggest that the key residue is in close vicinity to the DDxx(x)D/E motif (Scheme 4.12 A).

Intriguingly, no conserved amino acid motif is observed in the reported residues controlling the stereochemical outcome in ADS, TPS10 and AsR6, suggesting that no universal mechanism can be identified that controls/drives the transformation of FPP **3** into NPP **4**. However, in all discussed examples the gatekeeping amino acid residue is situated in close vicinity to the $PP_i$ binding motif.

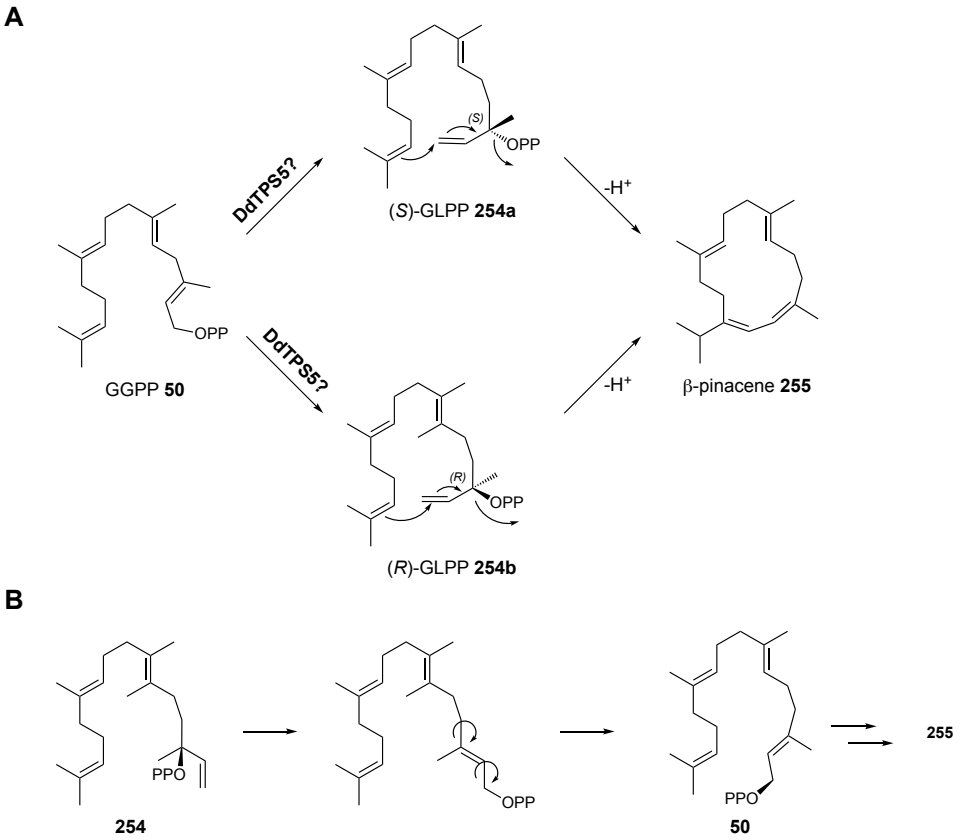
The question of how AsR6 and EupR3 confer stereoselectivity is further complicated by the fact that these cyclases (at least AsR6) appear to convert *both* enantiomers of (*rac*)-NPP **4**: Conversion of (3*S*)-NPP **4b** was probed with all enzymes and mutants in this study and yielded identical product distributions as observed for (*rac*)-NPP **4**. Conversion of (3*R*)-NPP **4a** was confirmed indirectly by post-incubation experiments with AsR6 and bisabolene synthase. In future experiments these findings should be validated directly through synthesis of (3*R*)-NPP **4a**. While the synthetic strategy pursued in this study is not viable to assess (3*R*)-NPP **4a** (no commercial source of (3*R*)-nerolidol **238c**) there are alternative strategies to synthesize (3*R*)-NPP **4a** from farnesol.[49]

This use of both enantiomers of NPP (or related 3-*O*-diphosphates) is highly unusual in terpene biosynthesis. As outlined before, the majority of terpene cyclases that produce terpenoids with *Z*-alkene configurations solely convert one of the two NPP enantiomers into cyclic terpenoids (*e.g.,* bisabolene synthase, amorpha-diene synthase). However, it is also true that very few experiments have been performed to examine the stereoselectivity of TCs for 3-*O*-diphosphate substrates. DICKSCHAT *et al*. recently reported that β-himalachene synthase from *Cryptosporangium arvum* converts (3*R*)-NPP **4a** into β-himalachene **252** whereas (3*S*)-NPP **4b** is only converted into the acyclic *E*-β-farnesene **239**.[232] Another example is epicubenol synthase for which TANDON and CANE demonstrated that only (3*R*)-NPP **4a** but not (3*S*)-NPP **4b** is converted into epicubenol **253** (Scheme 4.13).[233]




**Scheme 4.13:** Stereoselective formation of (+)-epicubenol **253** and β-himalachene **252** from (3*R*)-NPP **4a.**

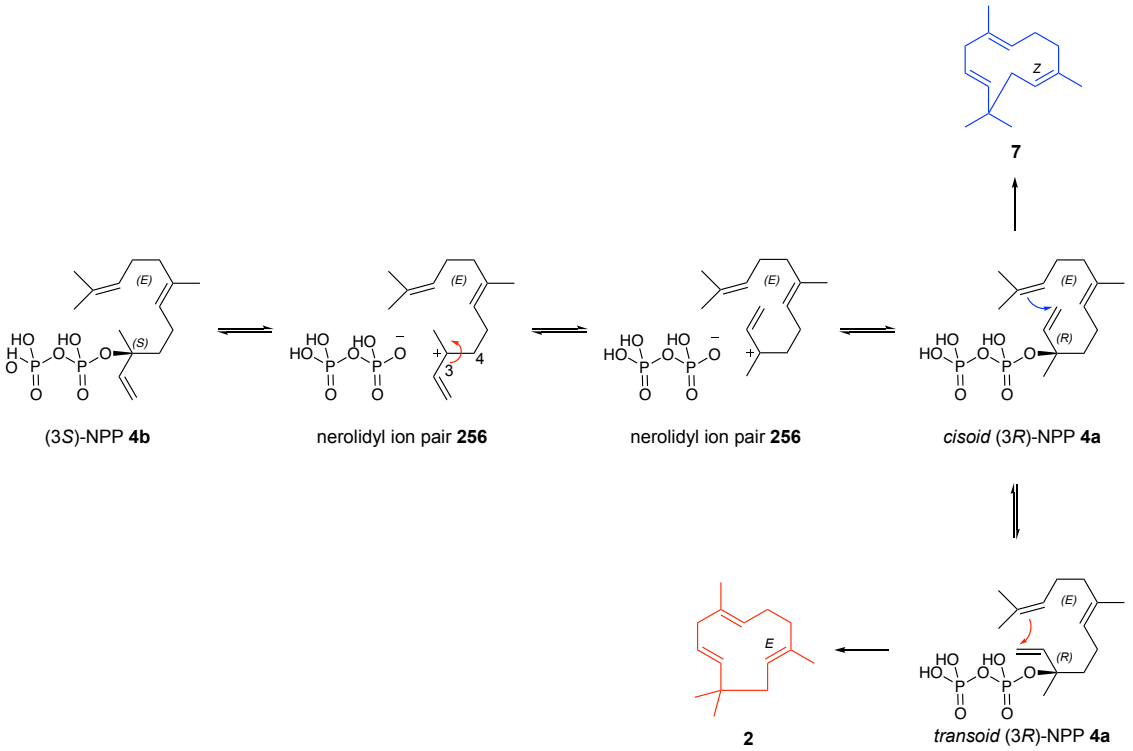
The only example of a terpene cyclase converting both enantiomers of a 3-*O*-diphosphate into the same cyclic product was reported by DICKSCHAT *et al*. for a diterpene synthase.




**Scheme 4.14: A,** proposed biosynthetic pathway leading to **261**; **B,** possible mechanism of isomerization of the unreactive GLPP-enantiomer in the active site of DdTPS5.

Similar to the isomerization of FPP **3** to NPP **4** during sesquiterpene formation, in diterpene biosynthesis formation of *Z*-configured alkenes has been proposed to proceed through the initial allylic isomerization of the native precursor geranylgeranyl diphosphate (GGPP) **50** into the 3-*O*-diphosphate geranyllinalyl diphosphate (GLPP) **254**. The diterpene cyclase β-pinacene synthase (DdTPS5) from *Dictyostelium discoideum* was shown to convert both enantiomers of geranyllinalyl diphosphate to β-pinacene **255** (Scheme 4.14 A). The authors proposed that the non-native GLPP enantiomer binds to the active site but may react back to GGPP **50** *via* an inverse allylic isomerization. Subsequent conformational changes could enable the successive 1,14-ring closure (Scheme 4.14 B).[234]

Based on the results obtained in this study there is no evidence that would allow us to suggest a validated mechanism on how AsR6 can convert both NPP enantiomers. One possibility, however, is a mechanism where AsR6 can interconvert (3*R*)-NPP **4a** into the reactive (3*S*)-NPP **4b** (or *vice versa*, dependent on which enantiomer is the reactive one). Mechanistically, such an interconversion would require a formal rotation of the vinyl side chain around the C-3/C-4 bond and the intermediacy of a `nerolidyl ion pair **256**´, as proposed in Scheme 4.15.




**Scheme 4.15:** Possible interconversion of (3*S*)-NPP **4b** into (3*R*)-NPP **4a**.

Concluding investigations were made into possible residues within the AsR6 active site that catalyze the final deprotonation step. However, *in vitro* analysis of AsR6$_{Y344F}$, with Y344 occupying a structurally equivalent position as the base Y520 in aristolochene synthase, did not affect α-humulene **2** formation. A second tyrosine residue, that is conserved in the active site of

all humulene synthases, was investigated in this study and was targeted for mutation (Y351A). However, expression of *asR6*$_{Y351A}$ did not lead to the expression of functional enzyme. Insolubility of the Y351A mutant suggests that the introduced mutation negatively influences protein folding. Exchange of Y351 for a sterically more equivalent amino acid (such as phenylalanine in the mutant Y344F) should be considered for future experiments. In the absence of other amino acid candidates to catalyze the final deprotonation step it is also possible that pyrophosphate acts as a general base in this process. This has been suggested in a range of well investigated class I terpene cyclases, where site-directed mutation was not successful in linking any amino acid to base activity.[39]

Together, the findings in this study reveal AsR6 as a unique new member of class I terpene cyclases. The work conducted here paves the way for further investigations into this unusual system. The focus of future experiments should be directed to investigate the unusual mechanism of how both (3*S*)-NPP **4b** and (3*R*)-NPP **4a** are accepted as substrates and converted into humulene. Further structural studies using *e.g.* nerolidol **238** as a ligand could shed light on the binding mode and further help elucidate this mechanism. Nerolidol **238** was already successfully used as a co-crystallization agent in the structure elucidation of hedycaryol synthase.[235] Additionally, site-directed mutagenesis should also be used to systematically assess the function of other residues within the active site. This could help identify factors that control the fidelity of the investigated humulene synthases (only one product each). In this context it would be very interesting to understand, mechanistically, how AsR6 (and related cyclases) prevent the humulyl cation **56**/**63** from undergoing additional carbocation trajectories, as observed for many other cyclases. Rational engineering of the active site cavity might also lead to the production of novel terpenes: *e.g.*, in germacrene A synthase (producing germacrene A **59** by initial 1,10-cyclization) a single point mutation was sufficient to turn the synthase into an α-humulene synthase (producing α-humulene **2** by initial 1,11-cyclization) and it would be very interesting to probe, whether reverse engineering of homologous residues in AsR6 would lead to similar success.[236]

# 5. Combinatorial Biosynthesis of New-to-Nature Tropolone Sesquiterpenoids

Part of the presented work has been published in *Angew. Chem. Int. Ed.* **2020**, *59*, 23870-23878.[122]

## 5.1. Introduction

The importance of natural products as lead structures for drugs and as drugs themselves is undisputed.[237] Several reasons contribute to the outstanding position that natural products hold in modern drug development. Most importantly, when compared to synthetic substances, natural products typically occupy a complementary region within chemical space (*e.g.* in terms of overall complexity but also in size and stereochemistry).[238] Natural products have also '*evolved in a biological context'* and evolution has honed highly specialized metabolites with discrete biological targets that can be exploited for drug development.[238]

Production of these compounds by microbial hosts is an additional attractive feature. The intricate structural frameworks of natural products are typically derived from very simple bioavailable building blocks, *e.g.* acetate or amino acids. Consequently, and opposed to synthetic approaches, expensive or toxic solvents and catalysts can be avoided and a *green* total biosynthesis is often possible.[238]

Over time several strategies have evolved to further diversify existing natural product scaffolds. These strategies were devised in order to optimize drug leads with respect to potency, bioavailability or emerging resistances. In addition to total- or semisynthesis of derivatives, combinatorial biosynthesis has proven to be an effective tool to systematically expand the chemical space around existing scaffolds. According to HEINZ G. FLOSS combinatorial biosynthesis can be defined as "*the application of genetic engineering to modify biosynthetic pathways to natural products in order to produce new and altered structures using nature´s biosynthetic machinery"*.[239] Pioneering work by HOPWOOD *et al*. in 1985 confirmed the feasibility of this concept. The transfer of genes between bacterial strains producing the isochromanequinone antibiotics actinorhodin, dihydrogranaticin and medermycin led to the production of novel hybrid antibiotics.[240]

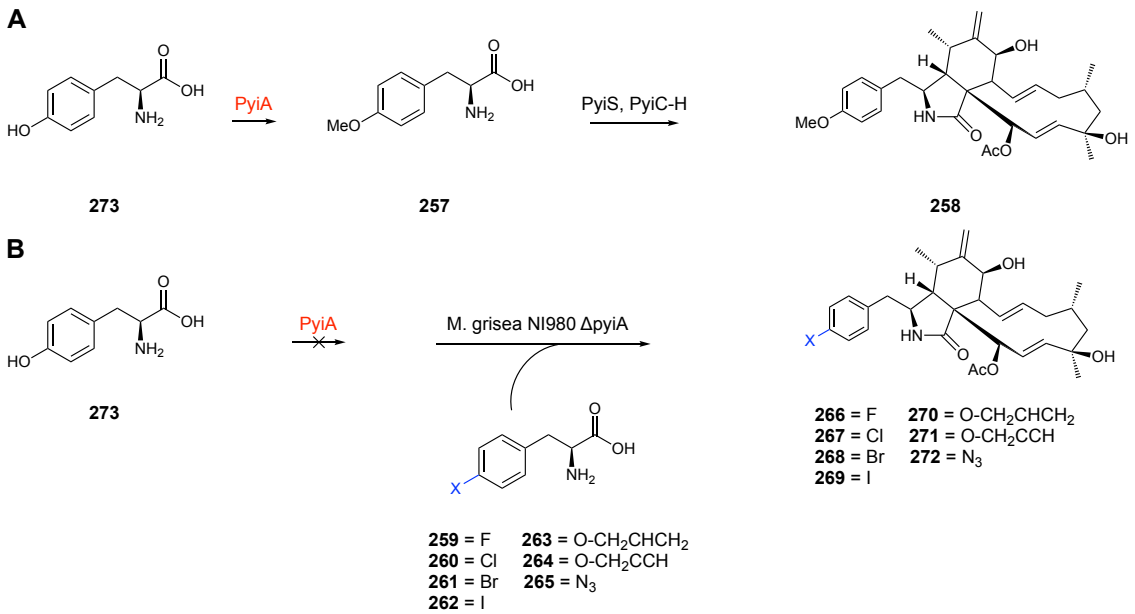
Building on HOPWOOD´s work several strategies have arisen that: **A,** exploit the promiscuity of individual enzymes for multiple substrates (*e.g.* mutasynthesis); **B,** use engineered enzymes for modified biosynthetic transformations (*e.g.* domain swaps in modular PKS/NRPS); or **C,** re-

combine biosynthetic genes from multiple pathways in a heterologous host (pathway-level engineering).[241]

### A) Mutasynthesis

Mutasynthesis constitutes the simplest approach to the engineering of natural product analogues. It does not require sophisticated knowledge of the biosynthetic pathway itself. In this process cultures of a producing organism are supplemented with precursor analogues. If successful, the substrate promiscuity of individual enzymes leads to the generation of novel compounds.[242]

The power of mutasynthesis in the generation of natural product analogues was recently demonstrated for the fungal cytochalasin pathway. Cytochalasins are natural products of hybrid PKS-NRPS origin that have reported antibacterial, antiviral and antitumor activity.[243] The broad spectrum of bioactivities makes them intriguing subjects to systematically generate novel derivatives and to assess the bioactivity of the newly generated compounds. Inactivation of the first biosynthetic step, formation of *O*-methyltyrosine **257** by the methyltransferase PyiA, abolished formation of pyrichalasin H **258**.[244] Feeding of diversely substituted tyrosines (compounds **259**-**265**) to the Δ*pyiA*-strain gave rise to a series of halogenated, *O*-alkyl, *O*-allyl, *O*-propargyl and 4′-azido analogs (**266**-**272**) in wild-type titers (Scheme 5.1).[244] Notably, introduction of azide groups allowed further derivatisation *via* classic synthetic methods (*e.g.* HUISGEN cycloaddition, 'click'-chemistry).[245]



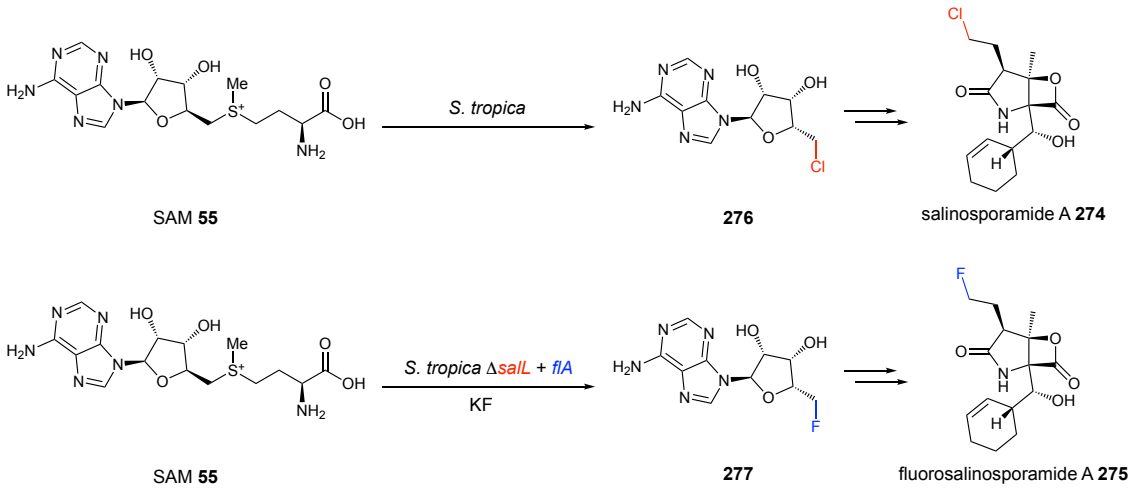

**Scheme 5.1** Mutasynthetic strategy to diversify the cytochalasin family of natural products: **A,** biosynthesis of pyrichalasin H **258** in *M. grisea* NI980; **B,** precursor-directed mutasynthesis in the *M. grisea* Δ*pyiA* strain.[244]

## B) Enzyme-Engineering

Engineering individual enzymes constitutes a more rational approach towards new natural product derivatives. Intensive efforts have been directed to modify the backbones of polyketides and non-ribosomal peptides. Given the modular assembly line fashion of the PKS and NRPS biosynthetic machinery, 'mixing-and-matching' modules to generate chimaeric genes and to produce rationally modified polyketides has been investigated in great detail. However, despite a series of proof-of-concept studies, current approaches are often impaired by drastically reduced production titers when compared to production levels of wild-type compounds.[21]

Recent advances in the area of non-ribosomal peptide synthethases by BODE *et al*. suggest that it should be possible to overcome these issues in the near future. Engineering of NPRS from bacterial strains belonging to the family of *Xenorhabdus/Photorhabdus* using a new interdomain fusion point resulted in the production of NRP-analogues and the *de novo* construction of artificial NRPSs in promising titers.[246]

More success has been achieved with strategies designed to modify the tailoring steps of natural product biosynthesis. For instance, MOORE *et al*. successfully modified the biosynthesis of salinosporamide A **274** from *Salinispora tropica*. Replacement of the native chlorinase encoding gene *salL* with the fluorinase encoding gene *flA* from *Streptomyces cattleya* gave rise to fluorosalinosporamide **275** when grown in the presence of potassium fluoride (Scheme 5.2).[247] Fluorinated natural products are extremely rare in nature, yet, up to 15% of pharmaceutical products on the market contain at least one fluorine atom.[247]




**Scheme 5.2:** Proposed biosynthesis of salinosporamide A **274** and successful strain engineering resulting in the biosynthesis of fluorosalinosporamide A **275**.

In similar fashion cryptic cytochrome P450´s have been successfully introduced in the cytochalasan pathway, giving rise to several derivatives with new functionality.[73]

### C) Pathway-level combinatorial biosynthesis

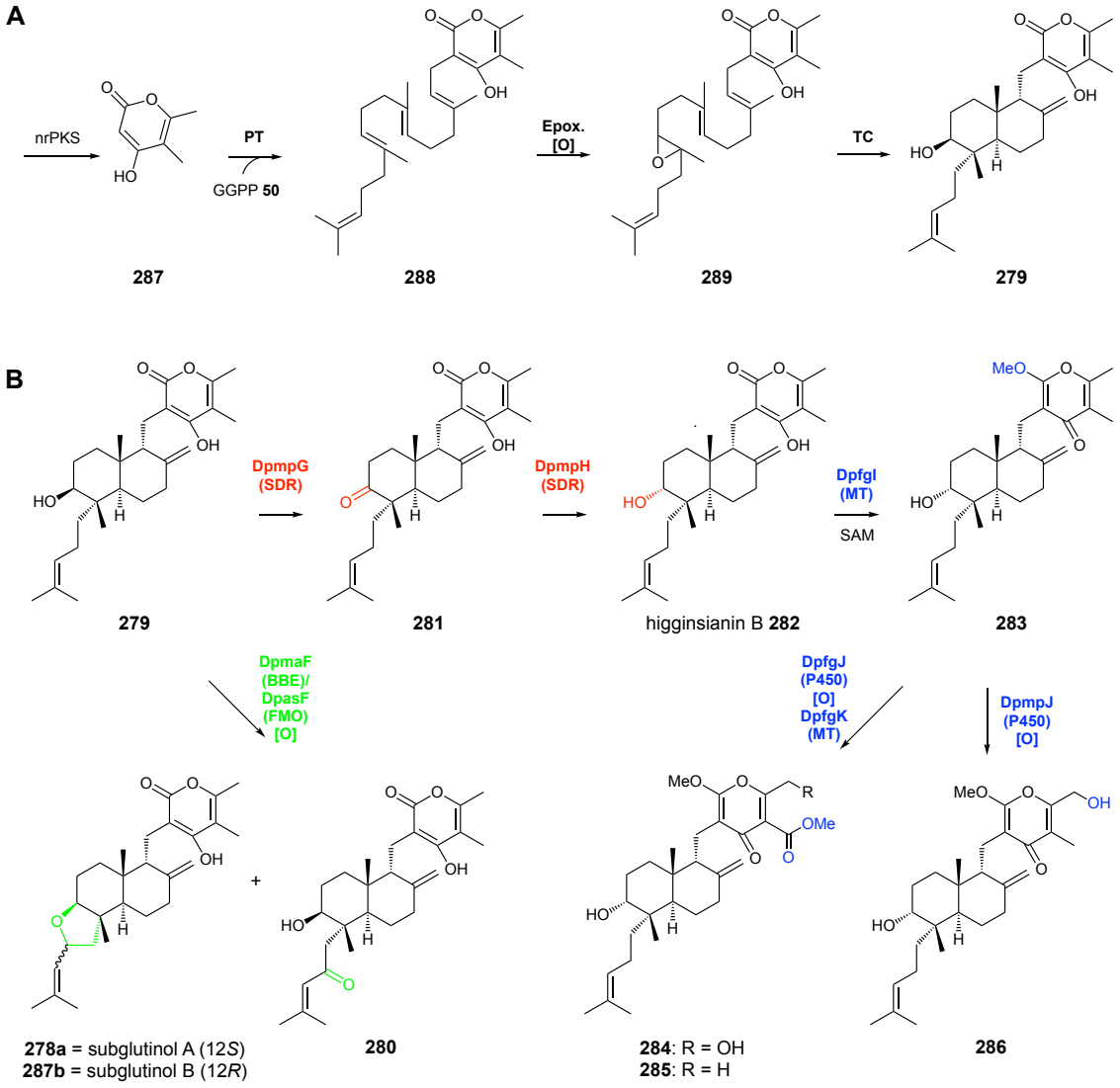
Pathway-level combinatorial biosynthesis is achieved by combining genes from multiple biosynthetic pathways in a heterologous host.[241] In a typical synthetic biology-based approach, the pathway to a specific natural product is initially reconstituted in the heterologous host. Genome-mined related biosynthetic genes are then introduced in the reconstituted pathway, either redirecting the original pathway to produce related natural products or by extending the original pathway *via* additional biosynthetic modifications.

ASAI *et al*. recently demonstrated the power of this approach to assess diversely functionalized decalin-containing diterpenoid pyrones (DDPs).[248] DDPs are natural products with diverse biological activities. Prominent examples such as metarhizin A or subglutinol A **278a** display potent antiproliferative and immunosuppressive activity.[249,250] Genome-mining identified homologous gene clusters in five fungal strains responsible for DDP-formation. Each cluster contained homologous copies of a central NR PKS, a geranylgeranyl diphosphate synthase, a prenyltransferase, an FAD-dependent epoxidase and a terpene cyclase.[248]

When expressed in *A. oryzae* NSAR1 these five genes gave rise to the common intermediate **279**, in excellent titers (87 $mg \cdot L^{-1}$; Scheme 5.3 A). The **279**-producing strain was used as the chassis for further combinatorial modifications (Scheme 5.3 B). *E.g.*, inclusion of FAD-dependent mononoxygenase DpasF or BBE domain-containing oxidoreductase DpmaF gave rise to the highly oxidized/tetrahydrofuran-containing DDPs **278ab** and **280** (Scheme 5.3 B). Alternatively, inclusion of the short-chain dehydrogenases DpmpG and DpmpH yielded the oxidized congeners **281** and higginsianin B **282**, respectively. Further inclusion of methyltransferases (DpfgI and DpfgK) and cytochrome P450 enzymes (DpmpJ or DpfgJ) gave rise to a series of oxidized and methylated unnatural DDPs, **283**-**286**. This combinatorial biosynthesis approach afforded 22 DDP´s, of which 15 were previously not reported in the literature. A variety of biological activities was determined for the unnatural DDP´s, including cytotoxicity against cancer cell lines, anti-HIV and prevention of amyloid β aggregation.[248]

The work by AISA *et al*. also demonstrates the remarkable power of heterologous expression in the rational generation of (new) compounds. Since the first heterologous reconstruction of a complete fungal BGC by LAZARUS *et al*. in 2010,[103,105] standard procedures and protocols for the handling, manipulation and cultivation of model organisms such as *A. oryzae* or *A. nidulans*

have been developed.[251] Nowadays, these experiments are routinely performed by the natural product scientific community. The use of standardized plasmids and transformation protocols grants access to many natural products and derivatives. Total biosynthesis of natural products has become a viable alternative to classic synthetic approaches. Notably, the yields are constantly improving: all DDP intermediates were isolated in multi-milligram amounts from the simple submerged fermentation of respective transformants.[248]




**Scheme 5.3** Combinatorial biosynthesis of DDP´s: **A,** conserved biosynthesis of the universal DPP precursor **279**; **B,** diversification and extension of DDP biosynthesis through heterologous expression of different gene combinations in *A. oryzae* NSAR1;[248] nrPKS = non-reducing PKS; PT = prenyltransferase; Epox. = FAD-dependent epoxidase; TC = terpene cyclase; FMO = FAD-dependent monooxygenase; SDR = short-chain dehydrogenase; MT = *O*-methyltransferase; P450 = cytochrome P450.

Likewise, clinically important compounds (or precursors) such as psilocybin (110 $mg \cdot L^{-1}$), arteminisic acid (25 $g \cdot L^{-1}$) or oxygenated taxanes (25 $mg \cdot L^{-1}$; the precursor of the anticancer drug paclitaxel) have been heterologously produced in excellent yields.[252–254]

## 5.2. Project Aims

The previous parts of this study were dedicated to the understanding of tropolone sesquiterpenoid biosynthesis. A combination of heterologous expression experiments and the *in vitro* characterisation of individual enzymes successfully delineated the function of key enzymes involved in the biosynthesis of xenovulene A **1**, pycnidione **9** and eupenifeldin **5**. Here, the obtained information is used to engineer new, unnatural, derivatives of **1**, **9** and **5**. In a systematic synthetic biology driven approach, a series of different gene combinations from the *asPKS1*, the *eup2* and the *pyc* BGC will be recombined in *A. oryzae* NSAR1. Through the rational design of reconstituted pathways control over formation of mono- *vs.* bistropolones, C-10-hydroxylation and oxidative ring-contractions will be attempted. New tropolone sesquiterpenoids will be isolated from the submerged fermentation of representative transformants and fully characterised. In cooperation with the research group of PROF. LEE-THEDIECK (Leibniz University Hannover) the bioactivity of selected compounds will be evaluated in regard to the capability to induce the production of erythropoietin in human cells.

## 5.3. Results – Combinatorial Biosynthesis

### 5.3.1. Construction of Expression Plasmids

In the course of this study a series of expression plasmids carrying genes from the *asPKS1* BGC, the *pyc* BGC and the *eup2* BGC was constructed. A combination of yeast homologous recombination and Gateway® technology was used to clone biosynthetic genes into the *A. oryzae* NSAR1 expression plasmids *pTYGS*-arg/ade/met.

For heterologous expression all genes are usually cloned from cDNA, to prevent problems arising from alternative intron splicing in the fungal host.[255] However, for cloning of genes from the pycnidione pathway (*pycR1*, *pycR6*) repeated attempts to isolate RNA were not successful and cDNA could not be produced. Since all genes required from this cluster were predicted to be intron-free, these genes were cloned from gDNA of *Leptobacillium* sp. CF-236968. For fungus *Phaeosphaeriaceae* sp. CF-150626 total RNA was isolated using the TRIzol™ Reagent (*Thermo Fischer Scientific*) and directly transcribed into cDNA using the High-Capacity RNA-to-cDNA™ kit (*Thermo Fischer Scientific*). The quality of the produced cDNA was assessed by amplification of the *eupL4* gene.

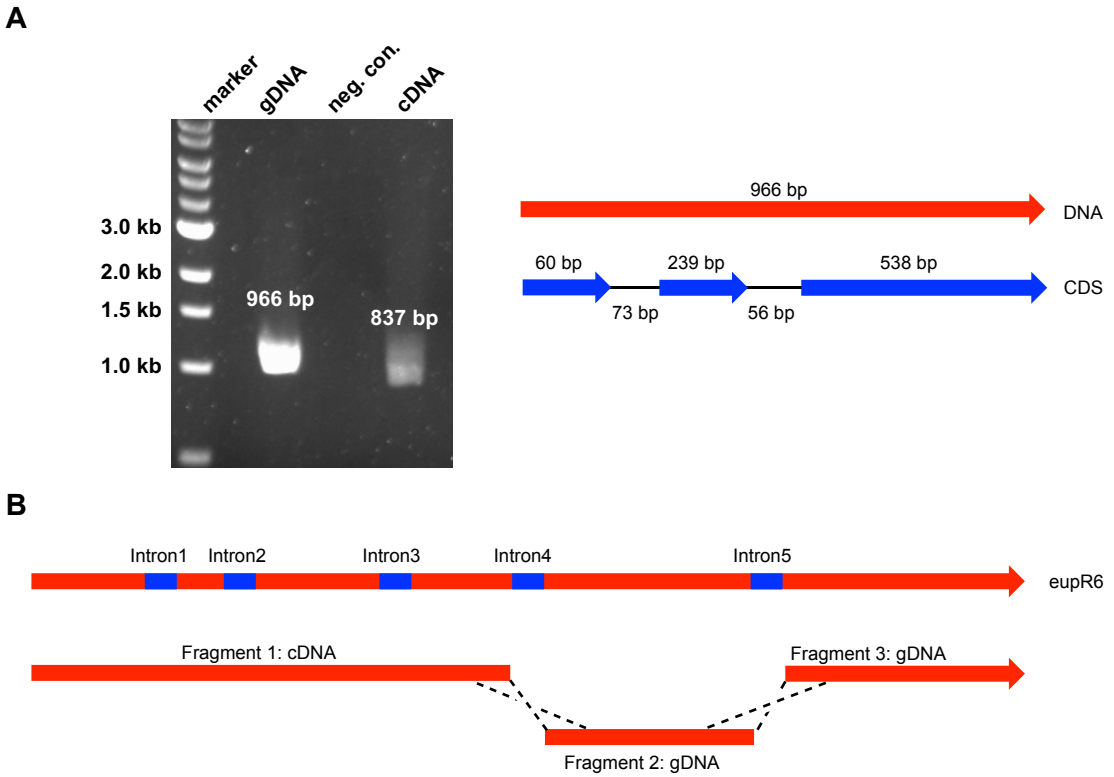



**Figure 5.1** Cloning of *eupR6* from the *eup2* BGC: **A,** analysis of *Phaeosphaeriaceae* sp. CF-150626 cDNA preparation; agarose gel electrophorese of the *eupL4* gene cassette when amplified from gDNA and cDNA; water was used as negative control (left); *eupL4* DNA sequence and predicted coding sequence (CDS; right); **B,** cloning strategy for the cytochrome P450 encoding gene *eupR6*.

Intron-prediction for *eupL4* by GeneMark[256] suggested the presence of two introns (combined size: 126 bp; Figure 5.1 A). Amplification of the *eupL4* gene cassette should afford the full-length gene product from gDNA (966 bp), and a shorter fragment when amplified from cDNA (837 bp). Indeed, a single DNA fragment was obtained upon amplification of *eupL4* from cDNA that was shorter than the fragment obtained by amplification from gDNA (Figure 5.1 A). Sequencing of the DNA fragment (*Eurofins*) confirmed excision of two introns. Notably, DNA sequencing revealed that the intron prediction by GeneMark was not accurate. While the number of predicted introns was correct, the beginning/end of both introns was wrongly predicted for several genes within the *eup2* BGC. Systematic sequencing of cDNA was used to determine the correct coding sequences for all used biosynthetic genes and the properly annotated sequences have been published in the course of this thesis.[122]

All biosynthetic genes from the *eup2* BGC were cloned from cDNA, except *eupR6*. Amplification of the cytochrome P450 encoding gene cassette *eupR6* did not result in amplification of the entire gene, as the 5´-reverse primer only bound to gDNA. A hybrid gDNA/cDNA *eupR6* gene was assembled from three DNA fragments in *pE-YA via* yeast homologous recombination (Figure 5.1 B). Fragment-1 was amplified from cDNA and excision of introns 1-3 was confirmed by DNA sequencing. Fragment-2 and fragment-3 were amplified from gDNA.

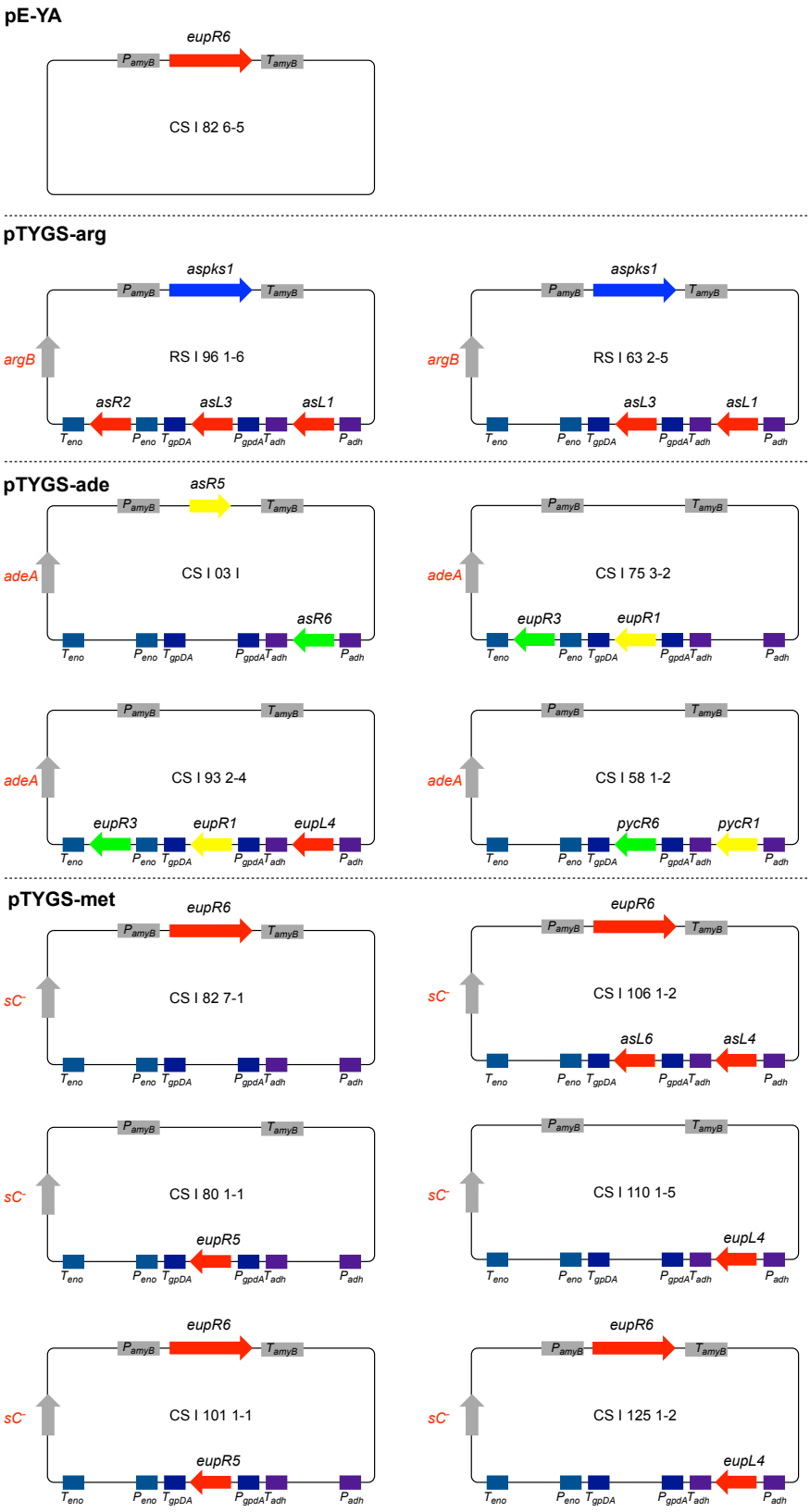



**Figure 5.2** Overview of expression plasmids used in this study: plasmids containing the prefix 'RS' were kindly provided by RAISSA SCHOR.[123]

A suitable primer design ensured that intron-4 and intron-5 were not amplified (see experimental section for details). Successful generation of *pE-YA_eupR6* was confirmed by DNA sequencing. An overview of all expression plasmids used in this study is displayed in Figure 5.2.

### 5.3.2. Strategy for the Combinatorial Biosynthesis of Unnatural Tropolone Sesquiterpenoids

Initial considerations were directed to a suitable starting point for engineering tropolone sesquiterpenoid biosynthesis. Rational engineering should start from a readily accessible, well-studied and highly bioavailable intermediate. During the investigations into the xenovulene biosynthetic pathway a key expression experiment led to formation of xenovulene B **131** (Section 2.3.5). Xenovulene B **131** is the only xenovulene with an intact tropolone nucleus. In terms of structure, it is thus more closely related to eupenifeldin **5** and pycnidione **9** than the other xenovulenes.

**Table 5.1** Overview of combinatorial expression experiments performed in this study. Each experiment represents a separate transformation into *A. oryzae* NSAR1, with the indicated combination of genes. Highlighted in red = genes from the *asPKS1* BGC; blue = genes from the *pyc* BGC; black = genes from the *eup2* BGC. PKS = polyketide synthase; FMO = FAD-dependent monooxygenase; NHI = non-heme iron dioxygenase; P450 = cytochrome P450; hDA = hetero DIELS-ALDERase; Hum = humulene synthase; RC = ring-contraction enzyme. #Xen. A = gene combination required for biosynthesis of xenovulene A **1**; #Xen. B = gene combination required for biosynthesis of xenovulene B **131**.

| #Experiment | | #Xen. A | #Xen. B | #1 | #2 | #3 | #4 | #5 | #6 | #7 | #8 | #9 | #10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ***gene*** | **Func.** | | | | | | | | | | | | |
| *asPKS* | PKS | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| *asL1* | FMO | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| *asL3* | NHI | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| *asR2* | P450 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | | | | | |
| *asR5* | hDA | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | | | |
| *asR6* | Hum | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ | | | | |
| *asL4* | RC | ✓ | | | ✓ | | ✓ | | | | | | |
| *asL6* | RC | ✓ | | | ✓ | | ✓ | | | | | | |
| *pycR1* | hDA | | | | | | | | | | | | ✓ |
| *pycR6* | Hum | | | | | | | | | | | | ✓ |
| *eupL4* | SDR | | | | | | | ✓ | ✓ | | ✓ | ✓ | |
| *eupR1* | hDA | | | | | | | | | ✓ | ✓ | ✓ | |
| *eupR3* | Hum | | | | | | ✓ | | | ✓ | ✓ | ✓ | |
| *eupR5* | RC | | | | | ✓ | | | | | | ✓ | |
| *eupR6* | P450 | | | ✓ | ✓ | | | | ✓ | | | ✓ | |

Initial experiments were designed to tailor the biosynthesis of xenovulene B **131** through inclusion of individual genes from the *eup2* BGC. Based on the results of these 'early' expression experiments more sophisticated experiments (*e.g.,* omissions/swaps of genes) were performed subsequently. In total, ten different gene combinations were heterologously expressed in *A. oryzae* (Table 5.1).

### 5.3.3. Transformation of *A. oryzae* NSAR1 and Confirmation of Transformants

For transformation into *A. oryzae* NSAR1 the previously deployed transformation protocol was used.[103] For the different experiments up to three plasmids were transformed in a single transformation step. Obtained transformants (typically ≥ 10) were selected on media lacking compounds corresponding to the auxotrophic markers present on the transformed plasmids. After three rounds of selection the transformants were screened by PCR for the presence of all transformed genes in the genomic DNA of the respective transformant (Figure 5.3). Transformants that tested positive for all genes (normally ca. 30-70%) were cultivated in liquid media (DPY, 7 d, 28 °C, 110 rpm), extracted with EtOAc and analysed by LCMS.

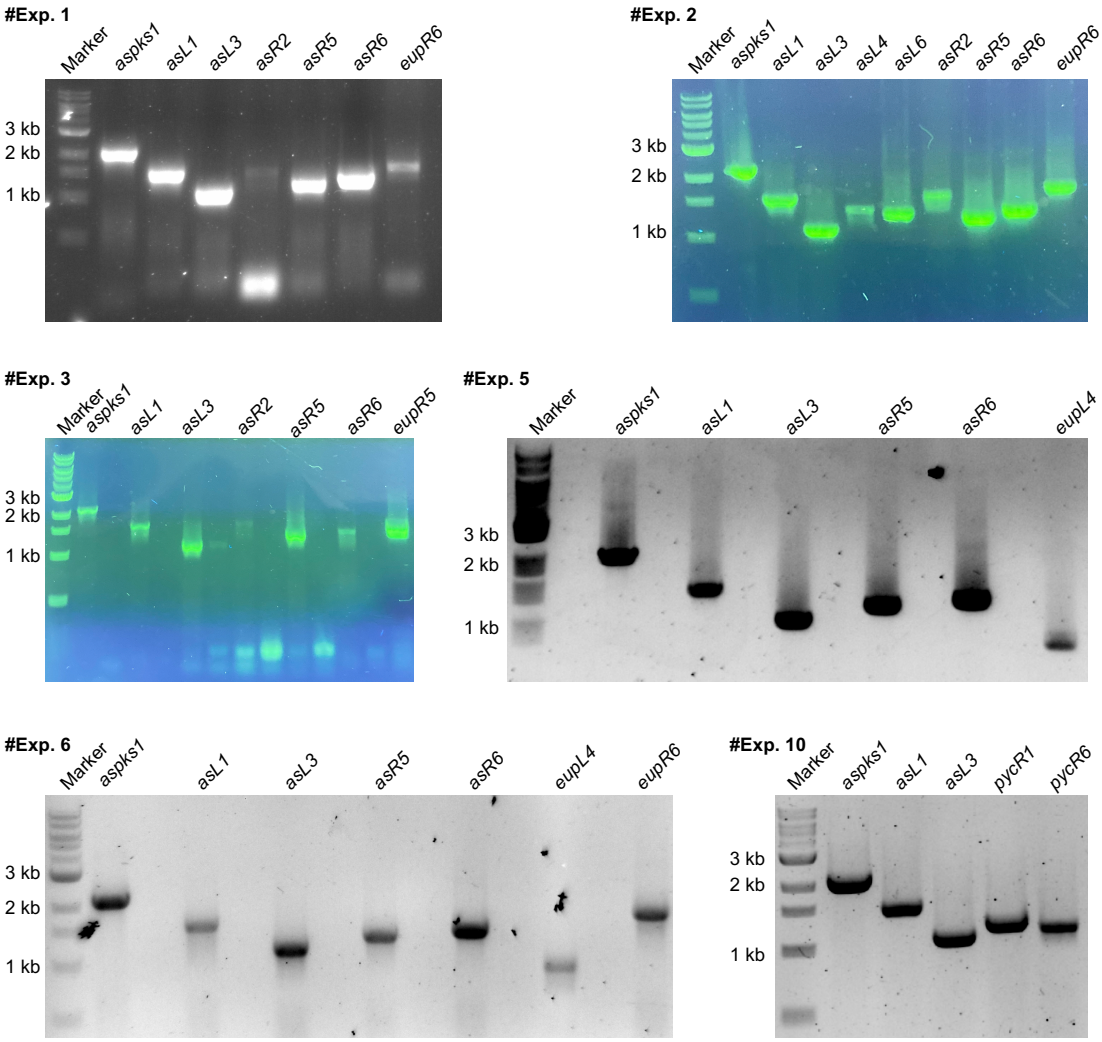



**Figure 5.3** PCR analysis of representative *A. oryzae* NSAR1 transformants: Gene-specific oligonucleotides were used for the amplification of each transformed gene cassette. Oligonucleotides used: *asPKS1* (83+322); *asL1* (421+423); *asL3* (424+426); *asL4* (645+706); *asL6* (650+709); *asR2* (427+428); *asR5* (761+762); *asR6* (759+760); *eupL4* (1500+1501); *eupR5* (1450+1451); *eupR6* (1458+1459); *pycR1* (1339+1340); *pycR6* (1341+1342). Data shown only for transformed gene sets that resulted in the production of novel meroterpenoids.

### 5.3.4. #Exp. 1: Expression of *eupR6*

*In silico* analysis of the *eup2* BGC suggested that the cytochrome P450 encoding gene *eupR6* is responsible for humulene hydroxylation at the C-10 position during the biosynthesis of eupenifeldin **5**. To probe the substrate promiscuity of EupR6, the *eupR6* gene was co-expressed with the genes required for xenovulene B **131** formation. Transformation of *A. oryzae* NSAR1 with seven genes (*asPKS1* + *asL1* + *asL3* + *asR2* + *asR5* + *asR6* + *eupR6*) yielded seven transformants (Figure 5.4 A).

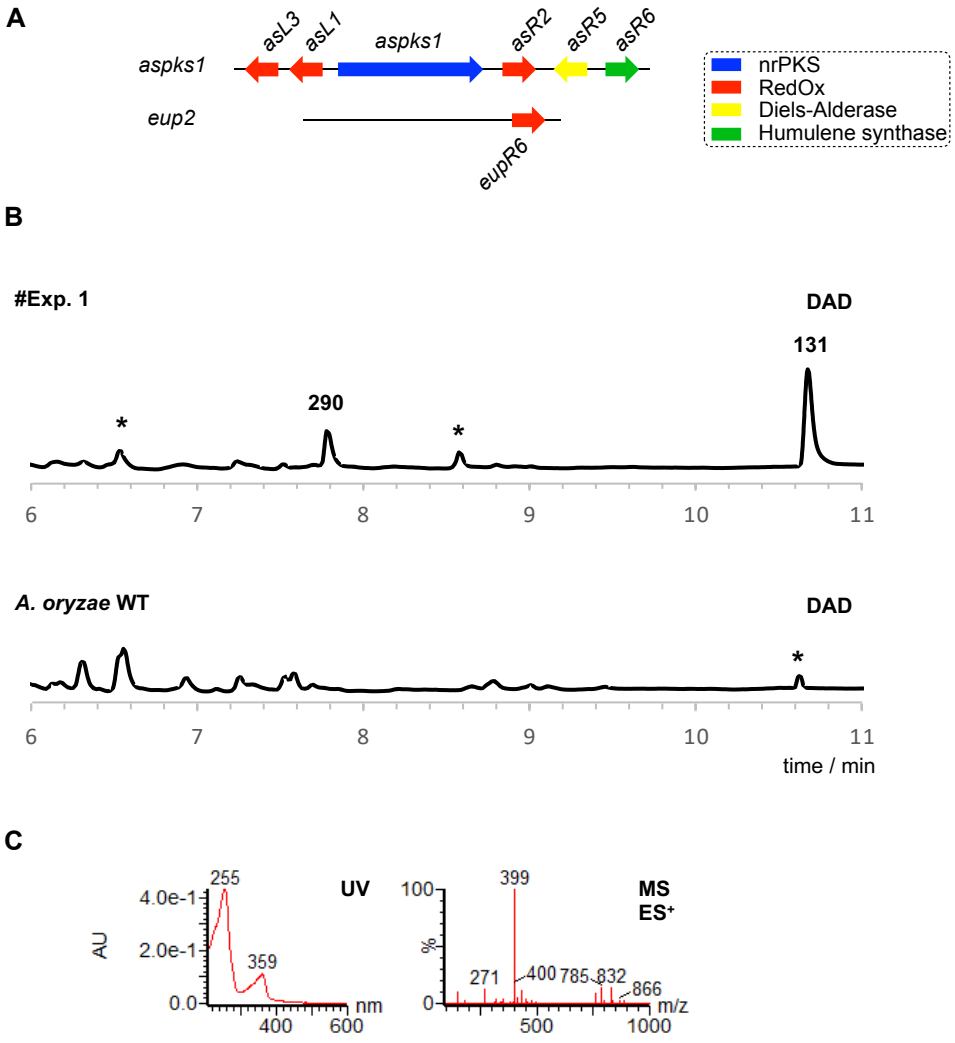



**Figure 5.4** Overview of #Exp. 1: **A,** co-expression of seven genes from the *asPKS1* and *eup2* BGC; **B,** LCMS DAD chromatograms of organic extracts of *A. oryzae* expression strain and wild-type control (arbitrary units); * = unrelated metabolite; the annotated *-metabolite in *A. oryzae* WT is different from compound **131**, despite identical retention times; **C,** chemical characterisation of **290** by UV and mass spectrum obtained from $ES^+$.

Crude extracts of representative transformants were analysed by LCMS for the production of tropolone sesquiterpenoids (Figure 5.4 B). Five extracts showed strong production of xenovulene B **131**, as identified by nominal mass (382), retention time ($t_R$ = 10.9 min) and characteristic UV absorption maxima ($\lambda$ = 250 nm and 360 nm). One major new peak was observed in four

chromatograms of the five extracts that showed production of xenovulene B **131**. The new compound **290** eluted at $t_R$ = 7.9 min, had a nominal mass of 398 and displayed UV absorption maxima at $\lambda$ = 255 nm and 360 nm (Figure 5.4 C). Purification of 5 mg from 1 L submerged fermentation in DPY medium enabled full NMR characterisation. The almost identical UV absorption maxima of **131** and **290**, the higher polarity of **290** as well as the 16 amu difference in nominal mass, suggested **290** to be a hydroxylated derivative of xenovulene B **131**. NMR characterisation of **290** confirmed the carbon skeleton of the known metabolite xenovulene B (Figure 5.5 A; Table 5.2). However, the C-10 methylene observed in **131** ($\delta_H$ = 2.12/2.26 ppm; $\delta_C$ = 38.0 ppm) was found to be replaced by a downfield shifted oxygenated carbon ($\delta_H$ = 4.36 ppm; $\delta_C$ = 77.1 ppm) in **290** (Table 5.2). Together with HRMS data ([M]H$^+$ calculated $C_{24}H_{31}O_4$ 399.2171, found 399.2174) this confirmed **290** as 10-hydroxyxenovulene B.

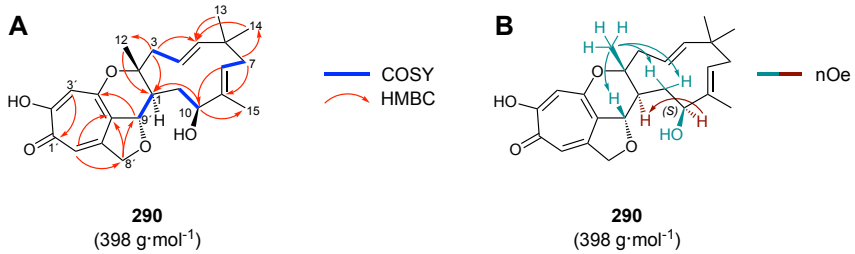
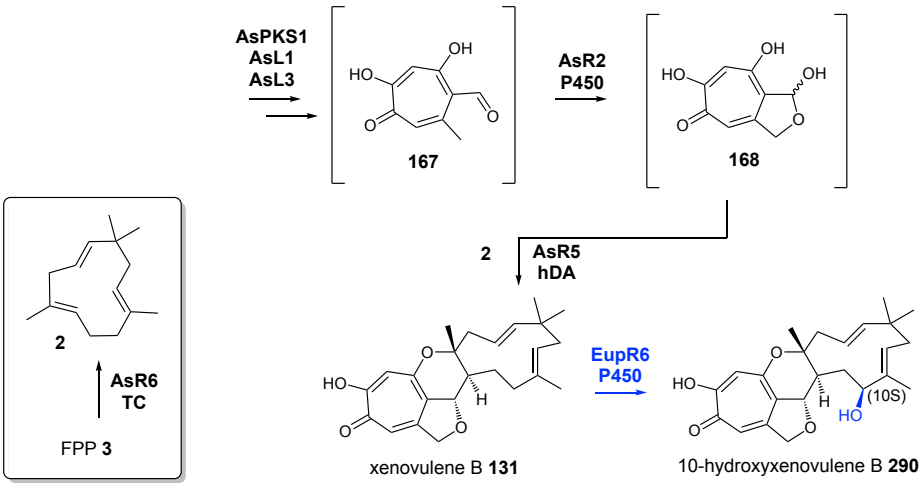



**Figure 5.5** Characterisation of 10-hydroxyxenovulene B **290**: **A,** structure elucidation by NMR spectroscopy (key COSY and HMBC correlations highlighted); **B,** analysis of the relative stereochemistry by NOESY-NMR.

**Table 5.2** NMR data for 10-hydroxyxenovulene B **290** in $CDCl_3$ (500 MHz) referenced to $CDCl_3$.

| Pos. | $\delta_H$ / ppm (mult, *J* in Hz) | $\delta_C$ / ppm | C-type | HMBC (H to C) | NOESY |
|---|---|---|---|---|---|
| **1** | 1.79 (m) overlap | 39.3 | CH | overlap | |
| **2** | - | 86.4 | C | - | - |
| **3a** | 2.29 (dd, 14.9, 10.5) | 43.1 | $CH_2$ | 4, 5 | 12 |
| **3b** | 2.64 (dt, 14.9, 2.3) | | | 1, 2, 4, 5, 12 | 12 |
| **4** | 4.88 (m) overlap | 119.5 | CH | overlap | - |
| **5** | 5.20 (m) overlap | 143.2 | CH | overlap | - |
| **6** | - | 38.8 | C | - | - |
| **7a** | 1.80 (m) overlap | 40.9 | $CH_2$ | overlap | - |
| **7b** | 2.23 (t, 12.3) | | | 6, 8, 9, 13, 14 | 13, 15 |
| **8** | 5.20 (m) overlap | 124.0 | CH | overlap | - |
| **9** | - | 138.2 | C | - | |
| **10a** | 4.36 (dd, 10.2, 1.3) | 77.2 | CH | 1, 8, 9, 11, 15 | 1, 11a, 11b |
| **10b** | | | | | |
| **11a** | 1.60 (m) | 36.8 | $CH_2$ | 1, 9, 10, 9′ | 10, 9′ |
| **11b** | 1.75 (m) overlap | | | overlap | 12 |
| **12** | 1.27 (s) | 22.8 | $CH_3$ | 1, 2, 3 | 3a, 3b, 11a, 11b, 9′ |
| **13** | 1.06 (s) | 30.0 | $CH_3$ | 5, 6, 7, 14 | 7b |
| **14** | 0.97 (s) | 24.1 | $CH_3$ | 5, 6, 7, 13 | - |
| **15** | 1.64 (s) | 10.6 | $CH_3$ | 8, 9, 10 | 7b |
| **1′** | - | 173.6 | C | - | - |
| **2′** | - | 166.1 | C | - | - |
| **3′** | 6.96 (s) | 110.9 | CH | 1′, 2′, 4′, 5′, 9′ | - |
| **4′** | - | 158.7 | C | - | - |
| **5′** | - | 121.0 | C | - | - |
| **6′** | - | 152.1 | C | - | - |
| **7′** | 6.92 (s) | 113.3 | CH | 1′, 2′, 5′, 8′ | - |
| **8′a** | 4.92 (m) overlap | 75.2 | $CH_2$ | overlap | |
| **8′b** | 4.99 (ddd, 13.6, 2.3, 1.2) | | | 5′, 6′, 7′, 9′ | 7′ |
| **9′** | 4.85 (m) | 83.8 | CH | 1, 11, 4′, 5′, 6′, 7′, 8′ | 11a, 12 |

The relative stereochemistry of **290** was assessed by NOESY-NMR. The absence of correlations between $H_3$-12 and H-1 established these protons to be *trans*, as previously observed for all other xenovulenes.[123] nOe correlations from $H_3$-12 to H-9′, $H_2$-3 and $H_2$-11, but not to H-10 places $H_3$-12 and OH-10 on the same face (Figure 5.5 B). This determines the relative stereochemistry at the C-10 position as 10*S*, identical with the observed stereochemistry at this position in eupenifeldin-type tropolone sesquiterpenoids.[81] The optical rotation of **290** was determined ($[\alpha]_D^{28}$ = + 100.0 [*c* = 0.090 g/100 ml, $CHCl_3$]). #Exp. 1 demonstrates that EupR6 is capable of hydroxylating non eupenifeldin-type tropolone sesquiterpenoids *in vivo*. The proposed biosynthetic pathway to **290** is displayed in Scheme 5.4.




**Scheme 5.4** Formation of 10-hydroxyxenovulene B **290** in *A. oryzae* NSAR1 upon expression of *asPKS1, asL1, asL3, asR2, asR5, asR6* and *eupR6*; compounds in brackets not observed; TC = humulene synthase; hDA = hetero DIELS-ALDERase; P450 = cytochrome P450.

### 5.3.5. #Exp. 2: Expression of *eupR6*, *asL4* and *asL6*

In a follow-up experiment *eupR6* was co-expressed with all genes required for formation of xenovulene A **1**, including the ring-contraction enzymes AsL4 and AsL6. Transformation of *A. oryzae* NSAR1 with nine genes (*asPKS1* + *asL1* + *asL3* + *asL4* +*asL6* + *asR2* + *asR5* + *asR6* + *eupR6*) afforded 19 transformants (Figure 5.6 A). Seven transformants produced xenovulene A **1**, as identified by low resolution mass (358), UV absorption maximum ($\lambda$ = 276 nm) and known retention time ($t_R$ = 7.8 min).[123] However, no new peaks corresponding to the expected hydroxylated derivative 10*S*-hydroxylxenovulene A **291a** was observed in DAD/ELSD chromatograms.

Extracted ion chromatograms were examined for trace production of hydroxylated xenovulene A **291** (expected nominal mass: 374) and the $ES^+/ES^-$ mass traces were searched for *m/z* $[M]H^+/[M-H]^-$ 375/373. Only one transformant showed traces of two such ions ($t_R$ = 5.6 and 6.1 min) in the extracted ion chromatogram (Figure 5.6 B).

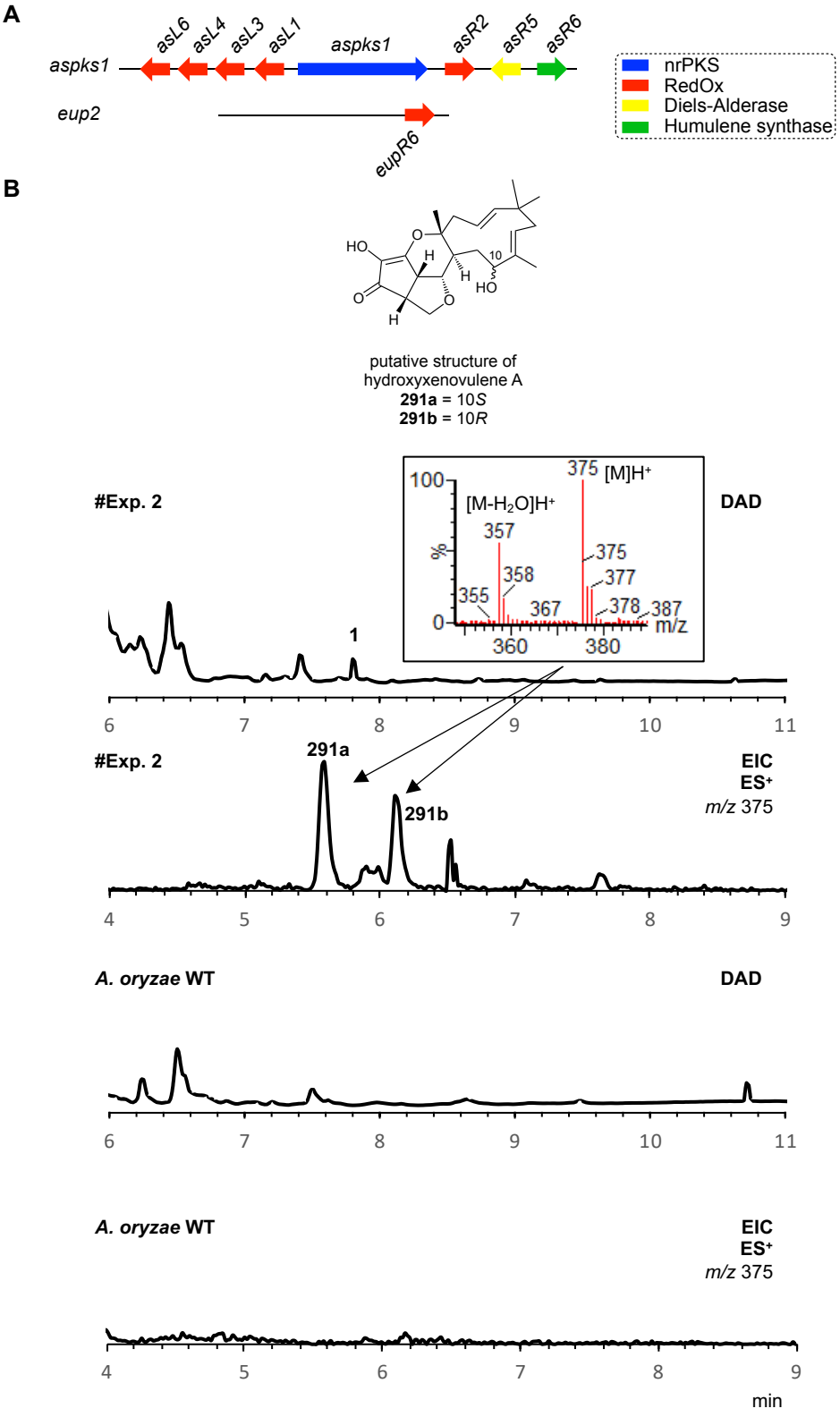



**Figure 5.6** Overview of #Exp 2: **A,** co-expression of nine genes from the *asPKS1* and *eup2* BGC; **B,** LCMS DAD/EIC chromatograms of organic extracts of *A. oryzae* expression strain and wild-type control (arbitrary units); EIC chromatograms searching for *m/z* = 375.

The nominal mass and more polar retention time in comparison with **1** was in agreement with these possibly being the 10*S*- and 10*R*-hydroxylated derivatives of xenovulenes A, **291ab**. Analysis of the mass fragmentation pattern provided additional evidence for the presence of **291a**

and **291b**, as key ions that corresponded to the loss of water $[M\text{-}H_2O]H^+$ were identified (Figure 5.6 B). However, due to the very low titers purification from large-scale fermentation of this transformant was not attempted.

### *5.3.6.* #Exp. 3: Expression of *eupR5*

Previous isotopic labelling studies in *Phaeosphaeriaceae* sp. CF-150626 had revealed that formation of the wildtype compound noreupenifeldin B **6** is the result of a ring-contraction (Section 3.3.3). A likely gene candidate to catalyse this transformation, *eupR5*, was previously identified in the *eup2* BGC. However, attempts to investigate the role of *eupR5 in vitro* were unsuccessful (Section 3.6). Expression in the fungal host *A. oryzae* might be more promising to yield correctly folded protein and help delineate the biosynthetic function of EupR5.

Transformation of *A. oryzae* NSAR1 with seven genes (*asPKS1* + *asL1* + *asL3* + *asR2* + *asR5* + *asR6* + *eupR5*) afforded ten transformants. Five out of ten transformants produced xenovulene B **131**. Two new peaks were observed in DAD chromatograms of these five transformants that were not produced in transformants lacking *eupR5* (Figure 5.7 B). Both compounds (**145** and **146**) displayed identical weak UV absorption and had identical low-resolution masses (370; Figure 5.7 C+D). The 12 amu difference in comparison to xenovulene B **131** (382) was consistent with a ring-contraction and excision of one carbon atom. Further analysis by HRMS ($[M]H^+$ calculated $C_{23}H_{31}O_4$ 371.2222, found 371.2218 for **145** and 371.2227 for **146**) corroborated this hypothesis.

Purification of both compounds by mass-directed LCMS was attempted from extracts of 1 L expression cultures in DPY medium but was not successful in separating the two compounds. Unfortunately, **145** and **146** showed only marginal differences in retention time at preparative scale ($t_R$ = 8.6 min for **145** and $t_R$ = 8.8 min for **146**). Consequently, a mixture of **145** and **146** (3.4 mg) was analysed by full NMR spectroscopy (Figure 5.8; Table 5.3).

Careful comparison to literature data identified **145** and **146** as the previously characterised phenolic xenovulenes,[123] that are the products of the ring-contraction enzymes AsL4 and AsL6 (previously characterised by RAISSA SCHOR).[123] **145** and **146** are regioisomers in which one of the hydroxyl groups in the phenol is either linked to the C-2′- or C-6′-carbon. The NMR data of **145** and **146** are largely identical – however, differences in the chemical shift of proton H-6′ in **145** ($\delta_H$ = 6.43 ppm; $\delta_C$ = 100.7 ppm) and H-2′ in **146** ($\delta_H$ = 6.25 ppm; $\delta_C$ = 100.9 ppm) were clearly discernible and matched the literature data sets. The proposed biosynthetic pathway leading to **145** and **146** is displayed in Scheme 5.5.

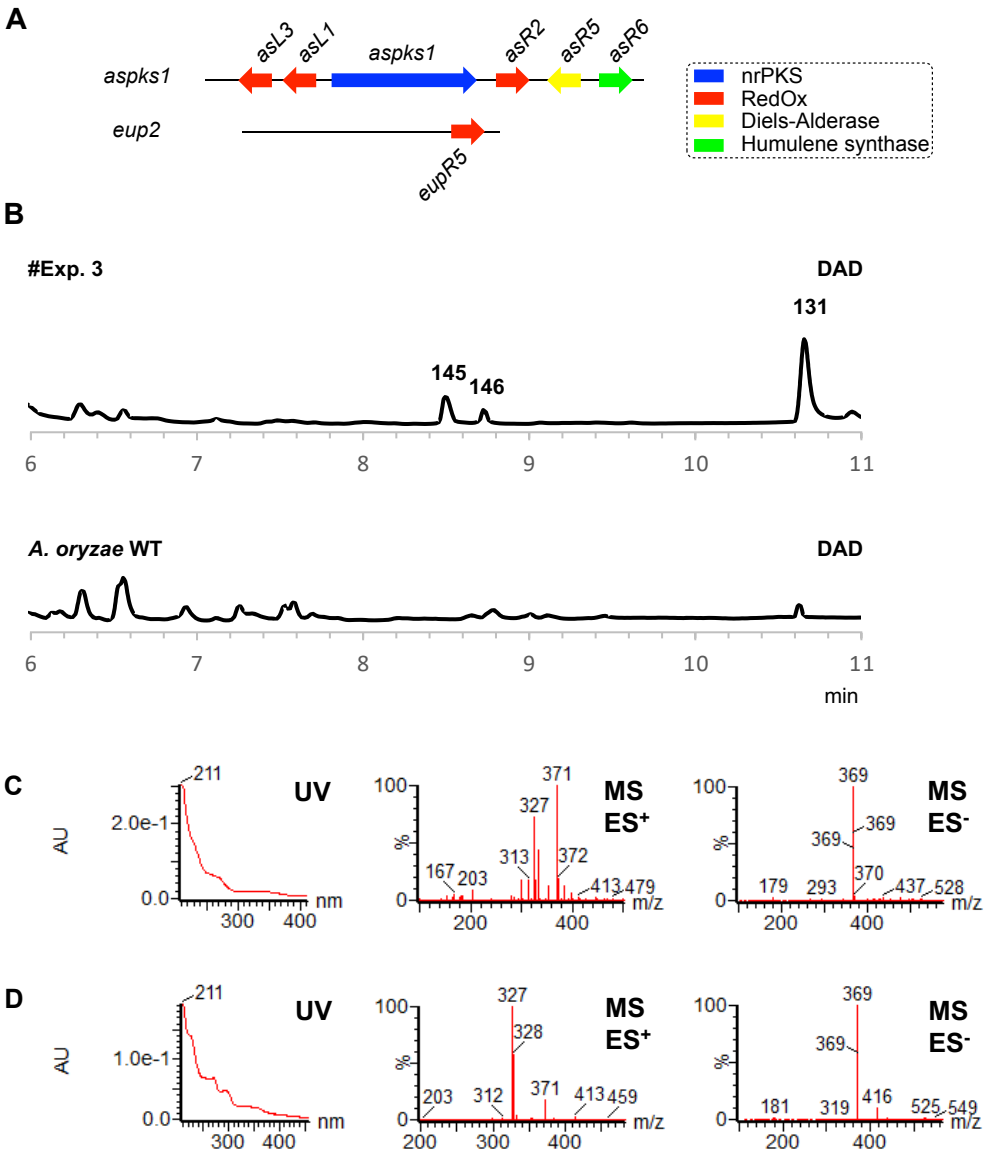
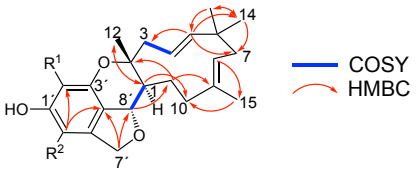



**Figure 5.7** Overview of #Exp. 3: **A,** co-expression of 7 genes from the *asPKS1* and *eup2* BGC; **B,** LCMS DAD chromatograms of organic extracts of *A. oryzae* expression strain and wild-type control (arbitrary units); **C,** chemical characterisation of **145** by UV and mass spectra obtained from $ES^+$ and $ES^-$; **D,** chemical characterisation of **146** by UV and mass spectra obtained from $ES^+$ and $ES^-$.

EupR5 catalyses the same reaction as previously observed for AsL4 and AsL6. Notably, AsL4 and AsL6 were more regioselective and only **145** (AsL4) or **146** (AsL6) were ever observed as the respective reaction product. EupR5 does not display the preference for one product and constitutes a less selective ring-contracting homologue.

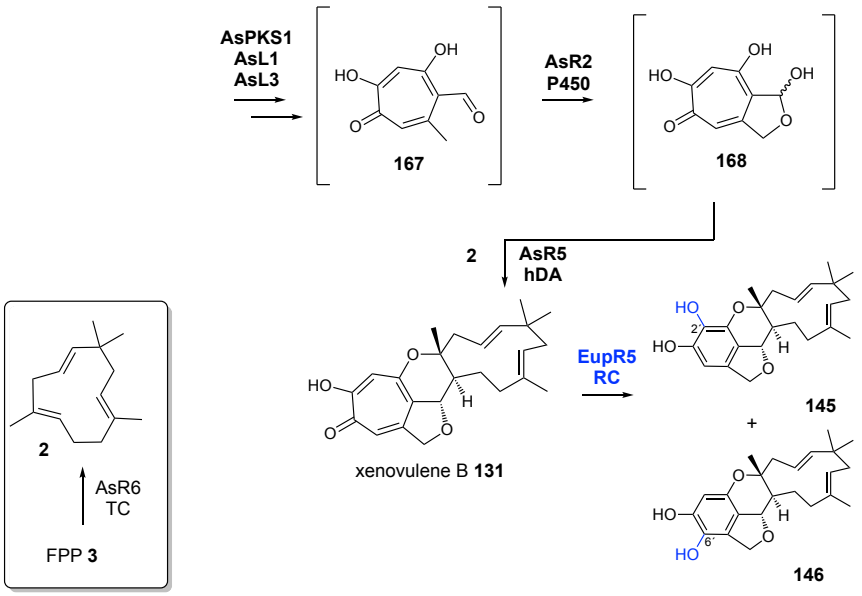



**145**: $R^1$ = OH, $R^2$ = H
**146**: $R^1$ = H, $R^2$ = OH
(370 g·mol$^{-1}$)

**Figure 5.8:** Characterization of **145** and **146** by NMR spectroscopy (key COSY and HMBC correlations highlighted).

**Table 5.3** NMR data for **145** and **146** in $CDCl_3$ (500 MHz) referenced to $CDCl_3$. Black colour = NMR data for **145**; Red colour = NMR data for **146**.

| Pos. | $\delta_H$ / ppm (mult, *J* in Hz) | $\delta_C$ / ppm | C-Type | HMBC (H to C) | COSY |
|---|---|---|---|---|---|
| **1** | 1.72 (m) | 43.2 | CH | 10, 11, 12 | 8´ |
| **2** | - | 86.8 | C | - | |
| **2** | | 84.8 | C | | |
| **3a** | 2.34 (m) | 43.2 | $CH_2$ | | |
| **3b** | 2.60 (d, 14.8) | | | 2, 4 | 4 |
| **3a** | 2.28 (m) | 43.4 | $CH_2$ | | |
| **3b** | 2.52 (d, 14.1) | | | 2, 4, 5 | |
| **4** | 4.95 (m) | 119.8 | CH | 5, 6 | |
| **4** | 4.97 (m) | 120.3 | CH | | |
| **5** | 5.14 (m) | 142.6 | CH | 3, 4, 7, 14 | 3 |
| **5** | 5.11 (m) | 143.2 | | | |
| **6** | - | 38.4 | C | - | |
| **7a** | 1.73 (m) | 41.7 | $CH_2$ | | |
| **7b** | 2.14 (m) | | | | |
| **8** | 5.01 (m) | 123.3 | CH | 15 | - |
| **8** | | 123.9 | CH | | |
| **9** | - | 136.5 | C | - | |
| **10** | 2.12 (m) | 38.0 | $CH_2$ | 8, 9 | |
| | 2.24 (m) | | | | |
| **11** | 1.31 (m) | 29.0 | $CH_2$ | 2 | |
| | 1.64 (m) | | | | |
| **12** | 1.25 (s) | 22.2 | $CH_3$ | 1, 2 | |
| **13** | 1.02 (s) | 30.3 | $CH_3$ | 5, 6, 7, 14 | - |
| **14** | 0.95 (s) | 24.3 | $CH_3$ | 5, 6, 7, 13 | |
| **15** | 1.59 (s) | 17.2 | $CH_3$ | 8, 9, 10 | |
| **1´** | - | 145.3 | C | - | |
| **2´** | - | 128.8 | C | - | - |
| **2´** | 6.25 (s) | 100.9 | CH | | |
| **3´** | | | | | |
| **4´** | - | 118.8 | C | - | |
| **5´** | - | 131.7 | C | - | |
| **6´** | 6.43 (s) | 100.7 | CH | 2´, 4´ | |
| **7´a** | 4.80 (m) | 73.4 | $CH_2$ | 4´, 5´, 8´ | |
| **7´b** | 5.01 (m) | | | - | |
| **7´a** | 4.95-5.00 (m) | 75.3 | C | | |
| **7´b** | 5.06 (m) | | | | |
| **8´** | 4.77 (m) | 82.2 | CH | 11 | 1 |
| | 4.75 (m) | 82.7 | CH | | |




**Scheme 5.5** Formation of phenolic xenovulenes **145** and **146** in *A. oryzae* NSAR1 upon expression of *asPKS1, asL1, asL3, asR2, asR5, asR6* and *eupR5*; compounds in brackets not observed; TC = humulene synthase; hDA = hetero DIELS-ALDERase; P450 = cytochrome P450; RC = ring-contraction enzyme.

### 5.3.7. #Exp. 4: Expression of *eupR3*

Co-expression of *eupR5* and *eupR6* from the *eup2* BGC successfully modified the xenovulene B **131** biosynthetic pathway. However, both changes represent minor modifications of the xenovulene pathway as they occur after the biosynthesis of xenovulene B **131** is completed. In the next experiments more fundamental changes were introduced into the expression system. Exchange of the α-humulene synthase encoding gene *asR6* for the corresponding 8*Z*-humulene synthase encoding gene *eupR3* might give rise to xenovulenes with the previously unprecedented 8*Z*-humulene backbone.

Transformation of *A. oryzae* NSAR1 with 8 genes (*asPKS1* + *asL1* + *asL3* + *asL4* + *asL6* + *asR2* + *asR5* + *eupR3*) afforded 19 transformants that were analysed for the production of compounds with the nominal mass of xenovulene A **1** (358; Figure 5.9 A). An analysis of DAD/ELSD traces did not show production of any tropolone sesquiterpenoids. However, analysis of extracted ion chromatograms searching for *m/z* $[M]H^+$ = 359 identified a corresponding peak in 10/19 transformants that was not observed in a wild-type control (Figure 5.9 B). Encouragingly, the retention time ($t_R$ = 8.4 min) was different from the retention time of xenovulene A **1** ($t_R$ = 8.2 min). Further analysis by HRMS ($[M]H^+$ calculated $C_{22}H_{31}O_4$ 359.2222, found 359.2224) supported trace production of the xenovulene A-homologue **292** with 8*Z*-configured alkene (Scheme 5.6). The isolation and purification of the putative new humulene derivative was attempted from 2L expression of a representative transformant in DPY medium but did not provide sufficient material for NMR analysis.

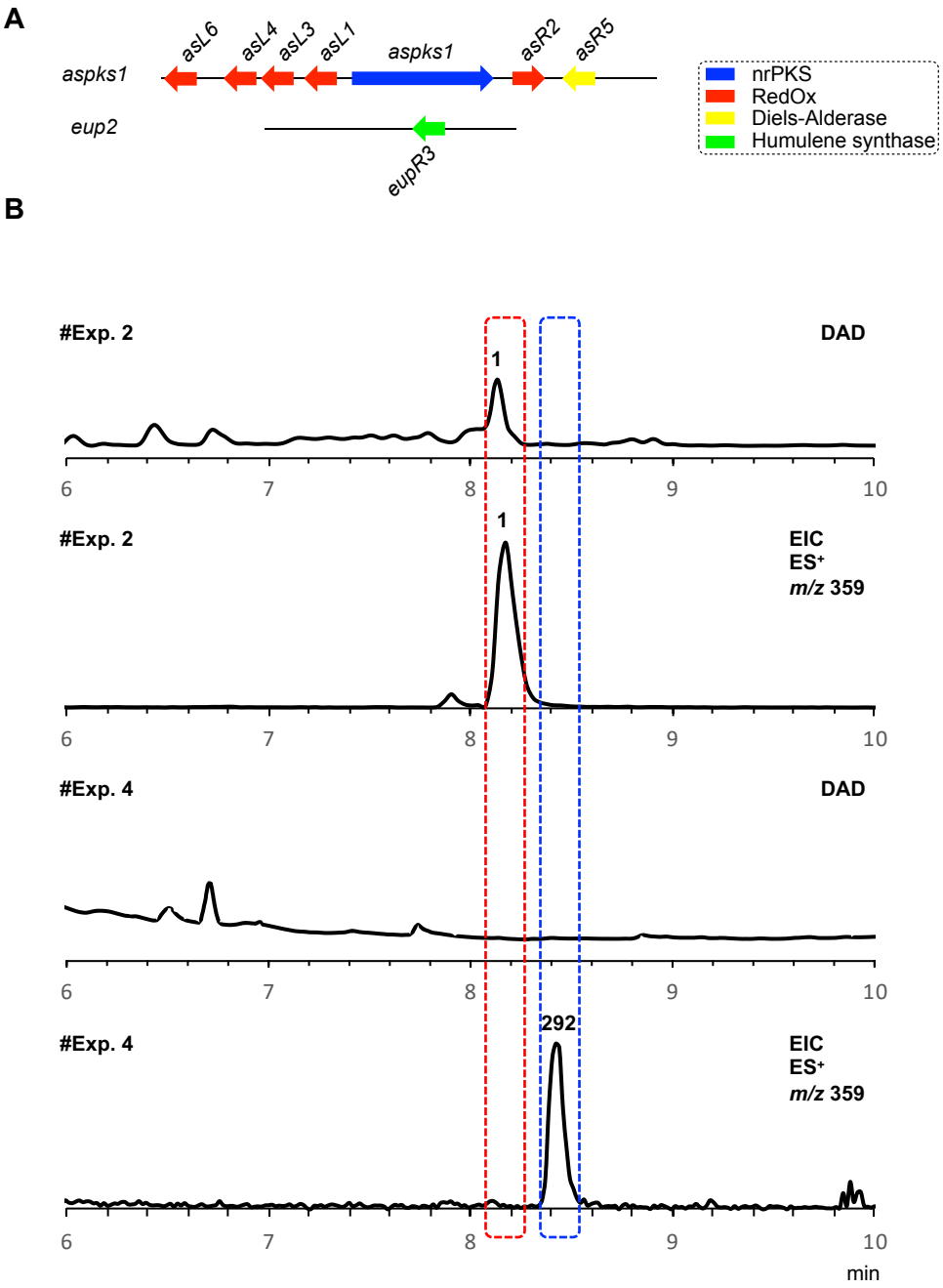
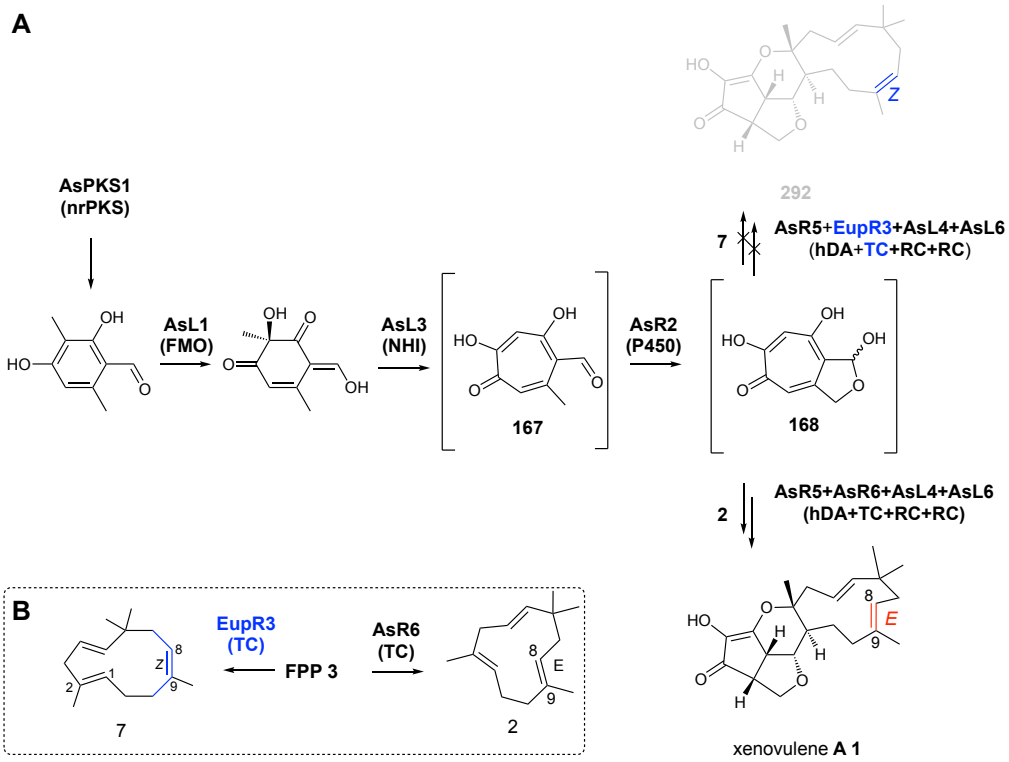



**Figure 5.9** Overview of #Exp. 4: **A,** co-expression of 8 genes from the *asPKS1* and *eup2* BGC; **B,** LCMS DAD/EIC chromatograms of organic extracts of *A. oryzae* expression strain (#Exp. 4) in comparison with previous data obtained for #Exp. 2 (in #Exp. 2 *asR6* was expressed instead of *eupR3* [arbitrary units]).




**Scheme 5.6** Putative formation of xenovulene-derivative **292** with an 8*Z*-humulene backbone: **A,** production of trace amounts of **292** in *A. oryzae* NSAR1 upon expression of *asPKS1, asL1, asL3, asL4, asL6, asR2, asR5,* and *eupR3*; **B,** humulene formation by AsR6 and EupR3; compounds in brackets not observed; nrPKS = non-reducing PKS; FMO = FAD-dependent monooxygenase; NHI = non-heme iron dioxygenase; TC = humulene synthase; hDA = hetero DIELS-ALDERase; P450 = cytochrome P450; RC = ring-contraction enzyme.

## 5.3.8. #Exp. 5: Expression of *eupL4*

The biosynthesis of xenovulene-type and eupenifeldin-type tropolone sesquiterpenoids diverges after formation of stipitaldehyde **167**. During the biosynthesis of xenovulenes the cytochrome P450 AsR2 *oxidizes* **167** to the hemiacetal **167**, giving rise to the unique tetrahydrofuran ring only observed in the xenovulenes. In the biosynthesis of eupenifeldin-type tropolone sesquiterpenoids the short-chain dehydrogenase EupL4 *reduces* **167** to the corresponding alcohol stipitol **172**. As described above, omission of *asR2* from the xenovulene expression system abolishes xenovulene formation (Section 2.3.4). Introduction of short-chain dehydrogenase encoding gene *eupL4* might direct the biosynthesis of xenovulenes into the production of eupenifeldin-like compounds that lack the characteristic tetrahydrofuran moiety. Transformation of *A. oryzae* NSAR1 with six genes (*asPKS1* + *asL1* + *asL3* + *eupL4* + *asR5* + *asR6*) afforded eight transformants (Figure 5.10 A + B).

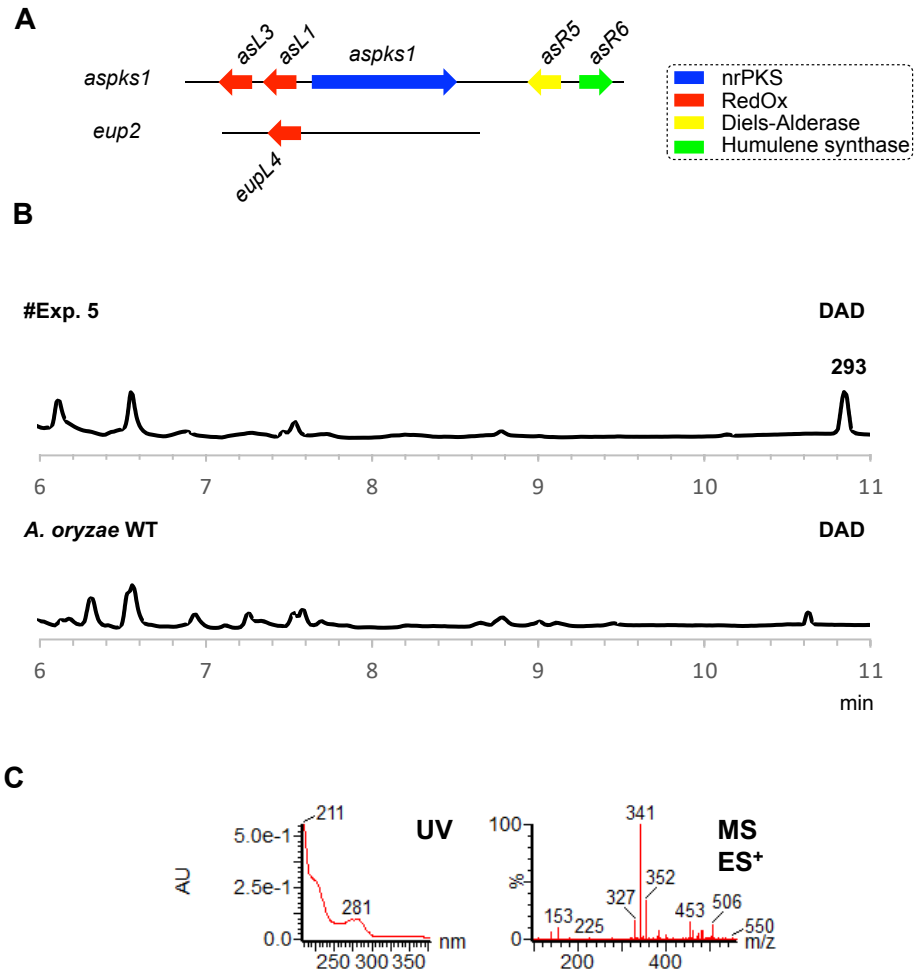
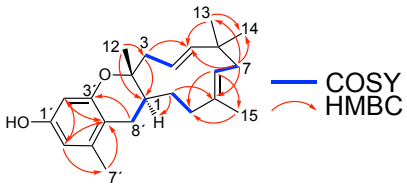



**Figure 5.10** Overview of #Exp. 5: **A,** co-expression of 6 genes from the *asPKS1* and *eup2* BGC; **B,** LCMS DAD chromatograms of organic extracts of *A. oryzae* expression strain and wild-type control (arbitrary units); **C,** chemical characterisation of **293** by UV and mass spectrum obtained from $ES^+$.

Two transformants produced a new non-polar compound **293**, eluting at $t_R$ = 11.1 min. The compound showed only weak UV activity ($\lambda$ = 211 nm and 281 nm; Figure 5.10 C). Surprisingly, the nominal mass of 340 was too small to correspond to either a mono- or a bistropolone. Purification of 3.4 mg from 0.8 L expression culture was successful. Full NMR characterization (Figure 5.11; Table 5.4) showed that instead of the characteristic tropolone signals (*e.g.,* two aromatic protons at $\delta_H$ = 6.9 - 7.2 ppm; aromatic methyl group singlet at $\delta_{H3}$ = ~2.4 ppm; carbonyl group at $\delta_C$ = ~170 ppm) aromatic protons were in place at $\delta_H$ = 6.18 ppm (H-2′) and $\delta_H$ = 6.25 ppm (H-6′). Together with an aromatic methyl group singlet at $\delta_H$ = 2.19 ppm ($H_3$-7′) this suggested replacement of the tropolone nucleus by a hydroxylated benzene. Indeed, full COSY and HMBC analysis identified **293** as a monobenzopyranyl-humulene. Key HMBC correlations from $H_2$-8′ to C-11, C-3′ and C-5′ place the monobenzopyranyl at the western face of humulene. Combined 1D- and 2D nOe data confirmed the relative stereochemistry at the humulene/dihydropyran ring junction. The absence of correlations between $H_3$-12 and H-1 suggest these to be *anti*-standing, in agreement with biosynthetic considerations and the

employment of AsR5 as the DIELS-ALDERase. The optical rotation of **293** was subsequently determined ($[\alpha]_D^{31}$ = - 10.5 [*c* = 0.095 g/100 ml, $CH_2Cl_2$:MeOH 1:1]).



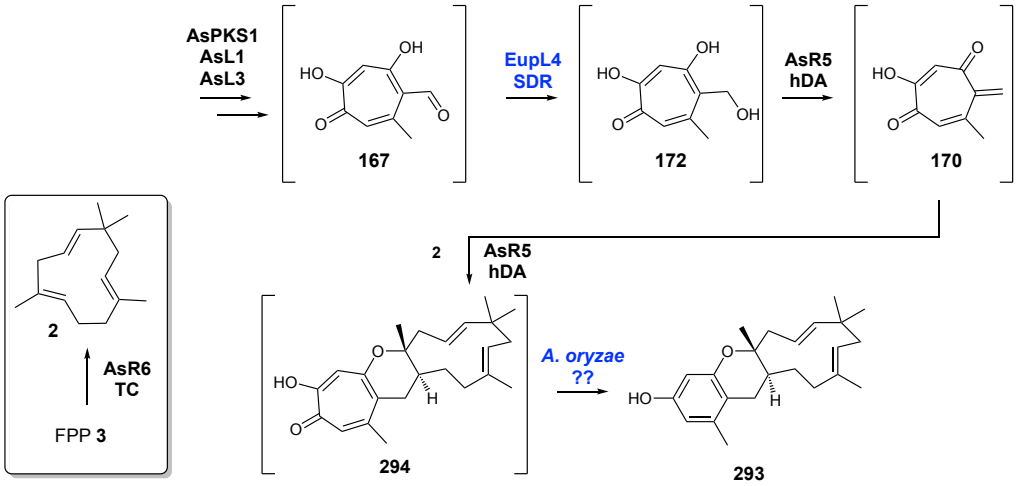

**293**
(340 g·mol$^{-1}$)

**Figure 5.11** Characterization of **293** by NMR spectroscopy (key COSY and HMBC correlations highlighted).

**Table 5.4** NMR data for **293** in $CDCl_3$ (500 MHz) referenced to $CDCl_3$.

| Pos. | $\delta_H$ / ppm (mult, *J* in Hz) | $\delta_C$ / ppm | C-type | HMBC (H to C) | 1D-NOESY | 2D-NOESY |
|---|---|---|---|---|---|---|
| **1** | 1.69 (td, 6.4, 3.6) | 34.3 | CH | 11, 12 | 8´b, 10 | 10 |
| **2** | - | 79.6 | C | - | | |
| **3a** | 2.23 (m) overlap | 43.0 | $CH_2$ | Overlap | | |
| **3b** | 2.51 (d, 14.4) | | | 1, 2, 4, 5, 12 | | |
| **4** | 5.17 (m) overlap | 121.5 | CH | Overlap | | - |
| **5** | 5.17 (m) overlap | 141.3 | CH | Overlap | | - |
| **6** | - | 38.8 | C | - | | |
| **7a** | 1.77 (m) overlap | 40.7 | $CH_2$ | Overlap | | |
| **7b** | 2.23 (m) overlap | | | Overlap | | |
| **8** | 5.20 (dd, 16.1, 1.7) | 123.6 | CH | 10, 15 | | 10, 14 |
| **9** | - | 138.7 | C | - | | |
| **10a** | 3.99 (d, 9.8) | 78.4 | CH | 1, 8, 11, 15 | 1, 8, 8´b | 1, 8´b, 11a |
| **10b** | | | | | | |
| **11a** | 1.12 (m) | 39.7 | $CH_2$ | 1, 9, 10 | | 10 |
| **11b** | 1.77 (m) overlap | | | Overlap | | |
| **12** | 1.06 (s) overlap | 20.0 | $CH_3$ | 1, 2, 3 | | |
| **13** | 1.06 (s) overlap | 30.5 | $CH_3$ | 5, 6, 7, 14 | | |
| **14** | 1.01 (s) | 24.3 | $CH_3$ | 5, 6, 7, 13 | | |
| **15** | 1.65 (s) | 10.8 | $CH_3$ | 8, 9, 10 | | |
| **1´** | - | 154.5 | C | - | | |
| **2´** | 6.18 (d, 2.5) | 101.5 | CH | 3´, 4´, 6´ | | |
| **3´** | - | 154.7 | C | - | | |
| **4´** | - | 112.8 | C | - | | |
| **5´** | - | 138.0 | C | - | | |
| **6´** | 6.26 (d, 2.5) | 108.8 | CH | 1´, 2´, 4´. 7´ | | 7´ |
| **7´** | 2.19 (s) | 19.4 | $CH_3$ | 4´, 5´, 6´ | | 6´ |
| **8´a** | 2.23 (m) overlap | 27.5 | $CH_2$ | Overlap | | |
| **8´b** | 2.89 (dd, 16.5, 5.6) | | | 1, 2, 3´, 4´, 5´ | 1, 10 | 1, 10 |

Intriguingly, the occurrence of the monobenzopyranyl moiety could not be linked to any of the transformed genes. The expected pathway product was the mono-tropolone **294** (Scheme 5.7). In the absence of any putative gene candidate to catalyze formation of/ring-contraction to the six-membered ring in **293** it seemed likely that *A. oryzae* NSAR1 catalyzed an unprecedented ring-contraction by itself.




**Scheme 5.7** Formation of monobenzopyranyl-humulene **294** in *A. oryzae* NSAR1 upon expression of *asPKS1, asL1, asL3, eupL4, asR5,* and *asR6*; compounds in brackets not observed; TC = humulene synthase; hDA = hetero Diels-ALDERase; SDR = short-chain dehydrogenase.

## 5.3.9. #Exp. 6: Expression of *eupL4* and *eupR6*

In a follow-up experiment, the gene set leading to formation of unexpected monobenzopyranyl-humulene **293** was co-expressed with the known humulene hydroxylase encoding gene *eupR6*. Transformation of *A. oryzae* NSAR1 with seven genes (*asPKS1* + *asL1* + *asL3* + *eupL4* + *eupR6* + *asR5* + *asR6*) afforded five transformants, four of which produced the formerly characterised compound **293** (Figure 5.12 A and B). All four transformants produced three additional peaks (**295**, **296**, **297**) that were not observed in absence of *eupR6*.

A new compound **295** eluted at $t_R$ = 7.8 min. The new compound had a nominal mass of 356 (HRMS $[M]H^+$ calculated $C_{23}H_{33}O_3$ 357.2430, found 357.2428; Figure 5.12 C). The 16 amu difference in regard to **293** was consistent with **295** being a hydroxylated derivative of **293**. Purification to homogeneity afforded 2.8 mg of **295** from 1L submerged fermentation. NMR analysis quickly confirmed the carbon skeleton of a monobenzopyranyl-humulene, identical with **293** (Figure 5. 13 A; Table 5.5). However, the C-10 methylene group in **293** ($\delta_H$ = 1.83 ppm and 2.10 ppm; $\delta_C$ = 37.9 ppm) was replaced by a downfield shifted oxygenated carbon in **295** ($\delta_H$ = 3.99 ppm; $\delta_H$ = 78.4 ppm), in agreement with **295** being a C-10 hydroxylated monobenzopyranyl-humulene. The relative stereochemistry of **295** was investigated using NOESY-NMR (Figure 5.13 B). Key nOe correlations from H-10 to H-1 and H-8´b and *vice versa* confirmed that all protons were on the same face. Further absence of nOe correlations between $H_3$-12 and H-1, H-10 and H-8´b confirmed the expected *trans* fusion at the humulene/dihydropyran ring junction. Accordingly, the stereochemistry at the C-10 position was determined to be *S*, in agreement with all previously observed hydroxylations in eupenifeldin-type tropolone sesquiterpenoids. The optical rotation of **295** was determined ($[\alpha]_D^{31}$ = - 33.3 [*c* = 0.050 g/100 ml, $CH_2Cl_2$:MeOH 1:1]).

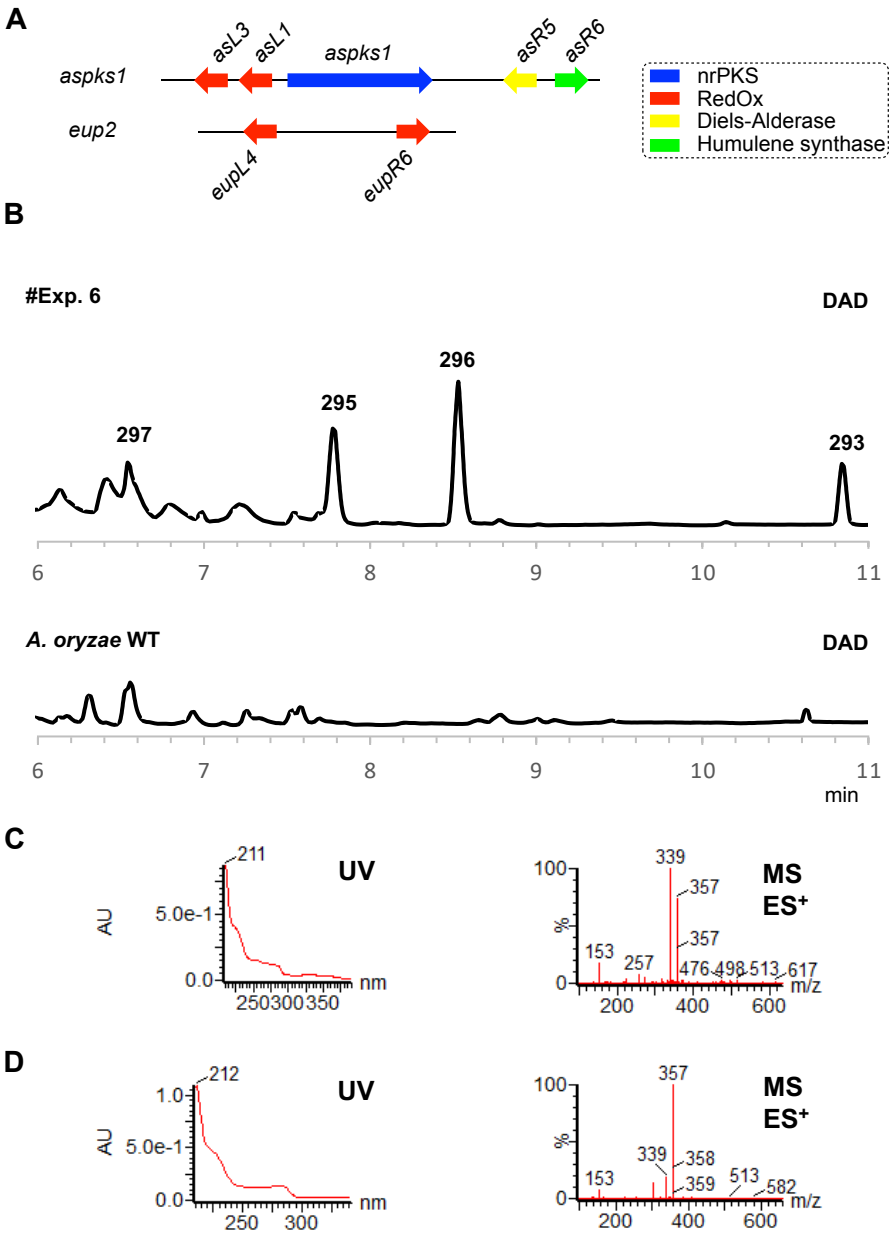



**Figure 5.12** Overview of #Exp. 6: **A,** co-expression of seven genes from the *asPKS1* and *eup2* BGC; **B,** LCMS DAD chromatograms of organic extracts of *A. oryzae* expression strain and wild-type control (arbitrary units); **C,** chemical characterisation of **295** by UV and mass spectrum obtained from $ES^+$; **D,** chemical characterisation of **296** by UV and mass spectrum obtained from $ES^+$.

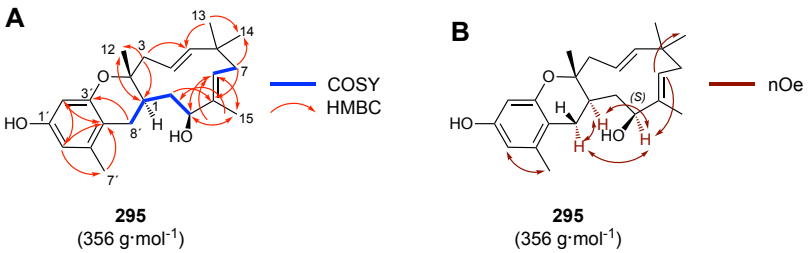



**Figure 5.13: A,** characterization of **295** by NMR spectroscopy (key COSY and HMBC correlations highlighted); **B,** analysis of the relative stereochemistry by NOESY-NMR.

**Table 5.5** NMR data for **295** in $CDCl_3$ (500 MHz) referenced to $CDCl_3$.

| Pos. | $\delta_H$ / ppm (mult, *J* in Hz) | $\delta_C$ / ppm | C-type | HMBC (H to C) | 1D-NOESY | 2D-NOESY |
|---|---|---|---|---|---|---|
| **1** | 1.69 (td, 6.4, 3.6) | 34.3 | CH | 11, 12 | 8´b, 10 | 10 |
| **2** | - | 79.6 | C | - | | |
| **3a** | 2.23 (m) overlap | 43.0 | $CH_2$ | Overlap | | |
| **3b** | 2.51 (d, 14.4) | | | 1, 2, 4, 5, 12 | | |
| **4** | 5.17 (m) overlap | 121.5 | CH | Overlap | | - |
| **5** | 5.17 (m) overlap | 141.3 | CH | Overlap | | - |
| **6** | - | 38.8 | C | - | | |
| **7a** | 1.77 (m) overlap | 40.7 | $CH_2$ | Overlap | | |
| **7b** | 2.23 (m) overlap | | | Overlap | | |
| **8** | 5.20 (dd, 16.1, 1.7) | 123.6 | CH | 10, 15 | | 10, 14 |
| **9** | - | 138.7 | C | - | | |
| **10a** | 3.99 (d, 9.8) | 78.4 | CH | 1, 8, 11, 15 | 1, 8, 8´b | 1, 8´b, 11a |
| **10b** | | | | | | |
| **11a** | 1.12 (m) | 39.7 | $CH_2$ | 1, 9, 10 | | 10 |
| **11b** | 1.77 (m) overlap | | | Overlap | | |
| **12** | 1.06 (s) overlap | 20.0 | $CH_3$ | 1, 2, 3 | | |
| **13** | 1.06 (s) overlap | 30.5 | $CH_3$ | 5, 6, 7, 14 | | |
| **14** | 1.01 (s) | 24.3 | $CH_3$ | 5, 6, 7, 13 | | |
| **15** | 1.65 (s) | 10.8 | $CH_3$ | 8, 9, 10 | | |
| **1´** | - | 154. | C | - | | |
| **2´** | 6.18 (d, 2.5) | 101.5 | CH | 3´, 4´, 6´ | | |
| **3´** | - | 154.7 | C | - | | |
| **4´** | - | 112.8 | C | - | | |
| **5´** | - | 138.0 | C | - | | |
| **6´** | 6.26 (d, 2.5) | 108.8 | CH | 1´, 2´, 4´. 7´ | | 7´ |
| **7´** | 2.19 (s) | 19.4 | $CH_3$ | 4´, 5´, 6´ | | 6´ |
| **8´a** | 2.23 (m) overlap | 27.5 | $CH_2$ | Overlap | | |
| **8´b** | 2.89 (dd, 16.5, 5.6) | | | 1, 2, 3´, 4´, 5´ | 1, 10 | 1, 10 |

Purification of the second new compound **296** afforded 4.4 mg from 1L submerged fermentation (Figure 5.12 D). Like **295**, **296** had a nominal mass of 356 (HRMS $[M]H^+$ calculated $C_{23}H_{33}O_3$ 357.2430, found 357.2431) and displayed only weak UV-absorption. Similar retention time ($t_R$ = 8.6 min), nominal mass and UV absorption suggested **296** to be a stereoisomer of **295**. Indeed, full NMR characterisation (Figure 5.14 A; Table 5.6) revealed that the carbon skeleton of **296** was identical to the skeleton of **295**. Again, the C-10 methylene group in **293** was replaced by a downfield shifted oxygenated carbon in **296** ($\delta_H$ = 3.99 ppm; $\delta_H$ = 78.4 ppm; Table 5.6). Notably, compared to **295**, **296** displayed significant differences in $^1$H-NMR shifts: both H-10 ($\delta_{H\text{-}10}$ = 4.35 ppm in **296**; $\delta_{H\text{-}10}$ = 3.99 ppm in **295**) and H-1 ($\delta_{H\text{-}1}$ = 2.05 ppm in **296**; $\delta_{H\text{-}1}$ = 1.69 ppm in **295**) were found to be shifted downfield in **296** by 0.36 ppm.

NOESY NMR was used to determine the relative stereochemistry of **296** (Figure 5.14 B). Key nOe correlations from H-1 to H-8´b and *vice versa* but not from $H_3$-12 to H-1, H-10 and H-8´b again confirmed *trans* fusion at the dihydropyran-humulene ring-junction. Accordingly, the stereochemistry at the C-10 position was tentatively assigned to be *R*, opposite to the stereochemistry observed in **295**. **296** is the first tropolone sesquiterpenoid identified so far with (10*R*)-stereochemistry. The optical rotation of **296** was determined ($[\alpha]_D^{31}$ = - 22.2 [*c* = 0.045 g/100 ml, $CH_2Cl_2$:MeOH 1:1]).

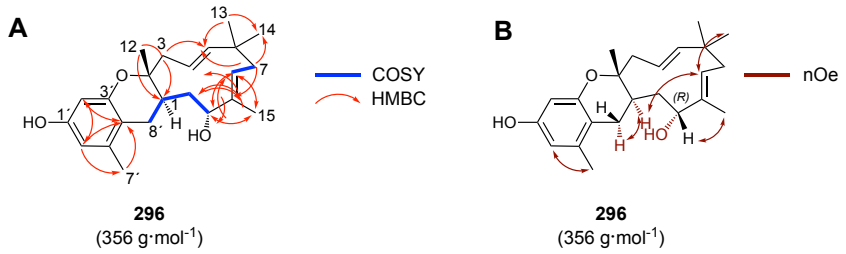



**Figure 5.14: A,** characterization of **296** by NMR spectroscopy (key COSY and HMBC correlations highlighted); **B,** analysis of the relative stereochemistry by NOESY-NMR.

**Table 5.6** NMR data for **296** in $CDCl_3$ (500 MHz) referenced to $CDCl_3$.

| Pos. | $\delta_H$ / ppm (mult, *J* in Hz) | $\delta_C$ / ppm | C-type | HMBC (H to C) | 1D-NOESY | 2D-NOESY |
|---|---|---|---|---|---|---|
| **1** | 2.05 (m) | 29.8 | C | 2, 10, 11 | 8, 8´b | 11a/11b, 8´b |
| **2** | - | 80.0 | C | - | - | - |
| **3a** | 2.18 (m) overlap | 42.8 | $CH_2$ | - | - | - |
| **3b** | 2.52 (dt, 15.1, 2.1) | | | 1, 2, 4, 5, 6, 7, 12 | 12 | 12 |
| **4** | 5.15 (m) overlap | 121.9 | CH | - | - | - |
| **5** | 5.15 (m) overlap | 140.8 | CH | - | - | - |
| **6** | - | 38.9 | C | - | - | - |
| **7a** | 1.82 (dd, 12.5, 4.6) | 40.9 | $CH_2$ | 5, 6, 8, 9, 14 | 8 | 12 |
| **7b** | 2.18 (m) overlap | | | - | - | - |
| **8** | 5.40 (dd, 12.0, 5.1) | 120.4 | CH | 7, 10, 15 | 1, 7a | 1, 7a, 14 |
| **9** | - | 139.7 | C | - | - | - |
| **10** | 4.35 (d, 6.2) | 73.9 | CH | 1, 8, 9, 11, 15 | 15 | 15 |
| **11a** | 1.55 (m) | 35.4 | $CH_2$ | 2, 9, 10, 8´ | - | - |
| **11b** | 1.55 (m) | | | | | |
| **12** | 1.06 (s) | 20.2 | $CH_3$ | 1, 2, 3 | 3b | 3b |
| **13** | 1.08 (s) | 30.3 | $CH_3$ | 5, 6, 7, 14 | - | 7a |
| **14** | 1.04 (s) | 24.2 | $CH_3$ | 5, 6, 7, 13 | | 7a, 8 |
| **15** | 1.55 (s) | 16.7 | $CH_3$ | 8, 9, 10 | 10 | 10 |
| **1´** | - | 154.2 | C | - | - | - |
| **2´** | 6.17 (d, 2.5) | 101.3 | CH | 1´, 3´, 4´, 6´ | - | - |
| **3´** | - | 154.6 | C | - | - | - |
| **4´** | - | 113.9 | C | - | - | - |
| **5´** | - | 137.8 | C | - | - | - |
| **6´** | 6.24 (d, 2.5) | 108.6 | CH | 1´, 2´, 4´, 7´ | | 7´ |
| **7´** | 2.17 (s) | 19.4 | $CH_3$ | 4´, 5´, 6´ | | 6ß |
| **8´a** | 2.18 (m) overlap | 30.7 | $CH_2$ | - | - | |
| **8´b** | 3.11 (dd, 16.5, 5.5) | | | 1, 2, 3´, 4´, 5´ | 1 | 1 |

Compound **297** was observed as a minor co-metabolite. Low-resolution mass spectrometry suggested a nominal mass of 218 (HRMS $[M]H^+$ calculated $C_{15}H_{23}O$ 219.1749, found 219.1749). Metabolite **297** eluted at 6.5 min and had no significant UV absorption. The nominal mass was too small to correspond to a tropolone sesquiterpenoid. Nominal mass and the absence of UV absorption suggested **297** to be an oxidized humulene congener (Figure 5.15). Purification of

4.0 mg from 1L submerged fermentation was followed by full NMR characterisation (Figure 5.15; Table 5.7). Interpretation of 2D-NMR data quickly confirmed that **297** was a humulene derivative.

However, careful analysis of $^{1}H$ and $^{13}C$ NMR spectra identified 15 carbon and 24 proton signals that contradicted the obtained HRMS data. A key proton signal at $\delta_{H\text{-}10}$ = 4.15 ppm in **297** was characteristic of a hydroxyl group at this position. Further carbon signals at $\delta_{C\text{-}1}$ = 59.3 ppm and $\delta_{C\text{-}2}$ = 63.1 ppm suggested the presence of an epoxide between C-1 and C-2 that was further supported by C-2 being a quaternary carbon and H-1 ($\delta_{H\text{-}1}$ = 2.37 ppm) shifted upfield in comparison to H-1 in α-humulene **2** ($\delta_{H\text{-}1}$ = 4.96 ppm). NMR analysis thus suggested **297** to be the 1,2-epoxy-10-hydroxy derivative of humulene **2** (nominal mass 236). Re-evaluation of low-resolution mass data indeed identified the corresponding ion $[M]H^+$ (*m/z* = 237) as a minor ion (Figure 5.15 B).

The previously assumed nominal mass of 218 corresponds to the $[M\text{-}H_2O]H^+$ ion (*m/z* = 219). Further identification of the $[M\text{-}H_2O\text{-}H_2O]H^+$ ion (*m/z* = 201) further supports the nature of **297**. The NMR data of **297** was highly similar to NMR data previously reported for phomanoxide **298**, with **298** displaying an additional epoxide group at C-4/C-5 (Figure 5.15 A). **297** was named phomanoxide B. The optical rotation of **297** was determined ($[\alpha]_D^{24}$ = + 309.1 [*c* = 0.055 g/100 ml, $CH_2Cl_2$:MeOH 1:1]). The relative stereochemistry of **297** was not further investigated. The proposed biosynthetic pathway leading to formation of **295**, **296** and **297** in *A. oryzae* NSAR1 is displayed in Scheme 5.8. The results of #Exp. 6 further validate the high substrate promiscuity of EupR6 as a broad-range humulene hydroxylase.

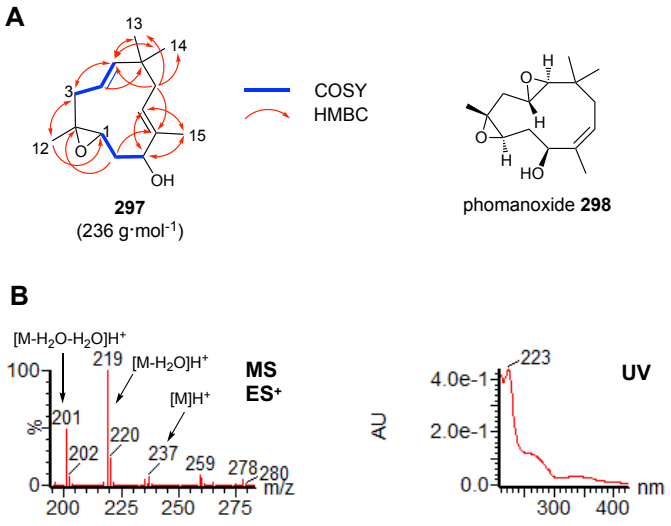
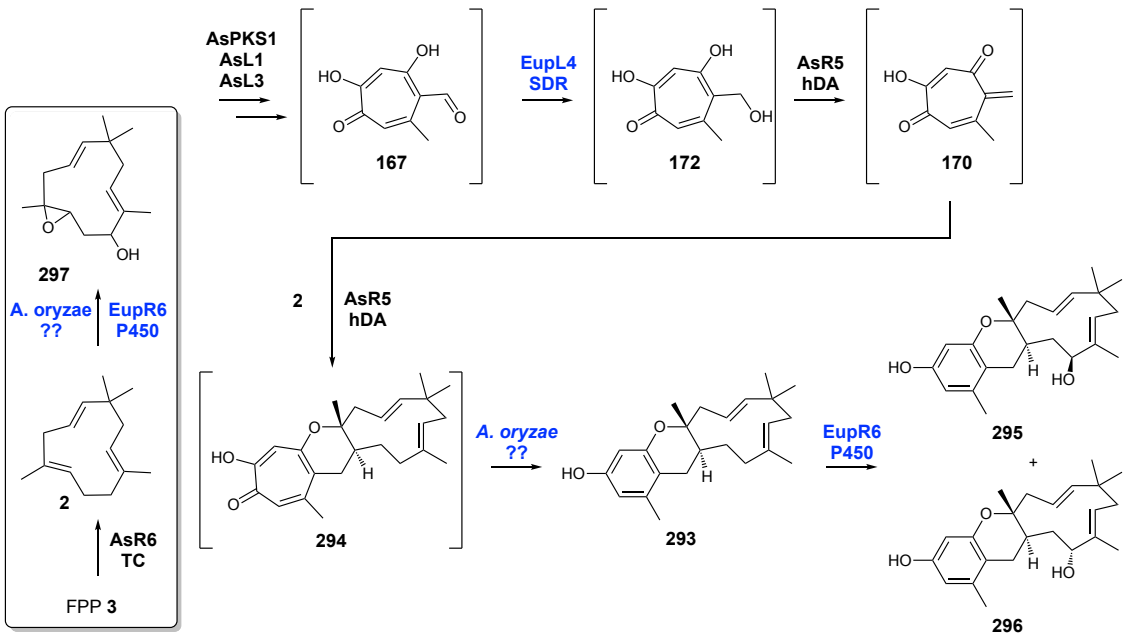



**Figure 5.15** Characterisation of **297: A,** characterization of **297** by NMR spectroscopy (key COSY and HMBC correlations highlighted); **B,** chemical characterisation of **291** by UV and mass spectrum obtained from $ES^+$.

**Table 5.7** NMR data for **297** in $CDCl_3$ (500 MHz) referenced to $CDCl_3$.

| Pos. | $\delta_H$ / ppm (mult, *J* in Hz) | $\delta_C$ / ppm | C-type | COSY | HMBC (H to C) |
|---|---|---|---|---|---|
| 1 | 2.37 (dd, 9.9, 5.1) | 59.3 | CH | 11 | 11 |
| 2 | - | 63.1 | C | - | - |
| 3a | 1.54 (t, 11.8) | 43.1 | $CH_2$ | 4 | 2, 4, 5, 12 |
| 3b | 2.58 (dd, 11.8, 4.8) | | | 4 | 1, 2, 4, 5 12 |
| 4 | 5.26 (ddd, 15.7, 10.7, 4.8) | 122.2 | CH | 5 | 3, 6 |
| 5 | 5.08 (d, 15.7) | 143.4 | CH | 4 | 3, 6, 13, 14 |
| 6 | - | 37.0 | C | - | - |
| 7a | 1.84 (dd, 13.8, 1.9) | 39.4 | $CH_2$ | 8 | 5, 6, 8, 9, 14 |
| 7b | 2.11 (dd, 13.8, 11.6) | | | 8 | 6, 8, 9, 13, 14 |
| 8 | 5.18 (dd, 11.6, 1.9) | 128.0 | CH | 7 | 10, 15 |
| 9 | - | 134.5 | C | - | - |
| 10 | 4.15 (dd, 9.9, 6.5) | 76.2 | CH | 11 | 8, 11, 15 |
| 11a | 1.67 (ddd, 13.4, 9.9, 6.5) | 33.5 | $CH_2$ | 1, 10 | 1, 9, 10 |
| 11b | 2.25 (ddd, 13.4, 9.9, 5.1) | | | 1, 10 | 1, 2, 10 |
| 12 | 1.32 (s) | 17.1 | $CH_3$ | - | 1, 2, 3 |
| 13 | 1.09 (s) | 30.3 | $CH_3$ | - | 5, 14 |
| 14 | 1.12 (s) | 23.8 | $CH_3$ | - | 5, 13 |
| 15 | 1.62 (s) | 10.8 | $CH_3$ | - | 8, 9, 10 |




**Scheme 5.8** Formation of monobenzopyranyl-humulenes **295** and **296** in *A. oryzae* NSAR1 upon expression of *asPKS1, asL1, asL3, eupL4, asR5, asR6* and *eupR6*; compounds in brackets not observed; TC = humulene synthase; hDA = hetero DIELS-ALDERase; SDR = short-chain dehydrogenase; P450 = cytochrome P450.

### 5.3.10. #Exp. 7: Expression of *eupR1* and *eupR3*

Previous expression experiments (#Exp. 1-6) successfully generated a series of non-natural xenovulene-type tropolone sesquiterpenoids. However, the obtained new compounds were restricted in terms of monosubstitution of the humulene macrocycle. This suggests that the DIELS-ALDERase from the xenovulene pathway, AsR5, is limited in the number of DIELS-ALDER reactions it can catalyse. In general, formation of bistropolones has not yet been achieved either *in vitro* or *in vivo*. HU *et al*. previously proposed that a second enzyme is involved in bistropolone formation that catalyses a second DIELS-ALDER reaction.[175]

However, a detailed bioinformatic analysis (Section 3.4.1) of all identified tropolone sesquiterpenoid BGC had not revealed a likely candidate. In the absence of such a candidate, we decided to probe the potential of EupR1, the DIELS-ALDERase from the eupenifeldin pathway, to catalyse formation of bistropolones. Co-transformation with the 8*Z*-humulene synthase encoding gene *eupR3*, instead of *asR6*, should prevent problems arising from the different alkene configurations.

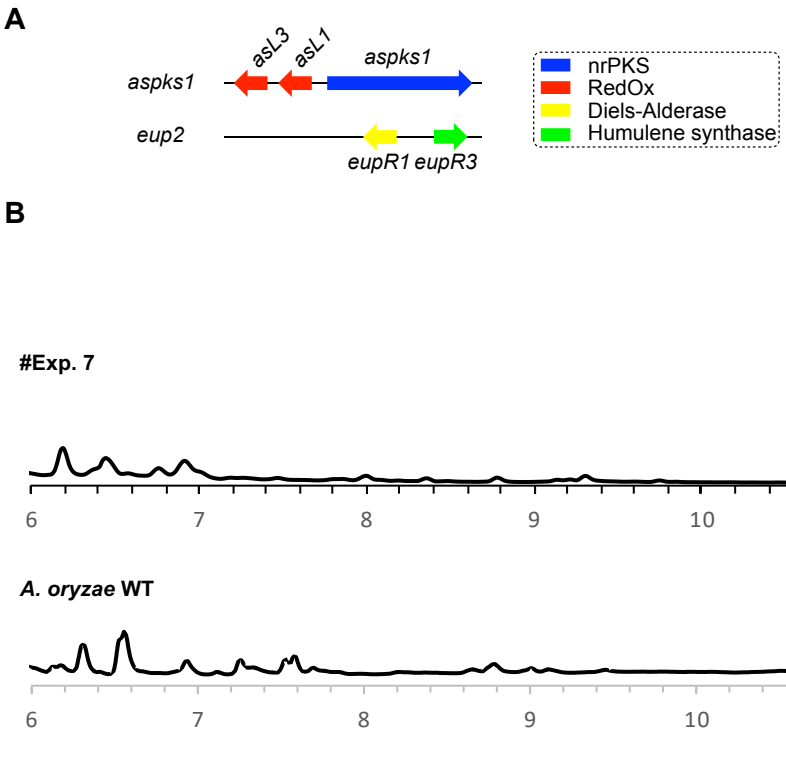



**Figure 5.16** Overview of #Exp. 7: **A,** co-expression of five genes from the *asPKS1* and *eup2* BGC; **B,** LCMS DAD chromatograms of organic extracts of *A. oryzae* expression strain and wild-type control (arbitrary units).

Transformation of *A. oryzae* NSAR1 with five genes (*asPKS1* + *asL1* + *asL3* + *eupR1* + *eupR3*) afforded 22 transformants. Submerged fermentation in DPY medium was followed by LCMS analysis. The early tropolone producing steps were active, as confirmed by identification of the

early tropolone intermediate **166** based on retention time ($t_R$ = 2.8 min), low resolution mass (182) and UV absorption ($\lambda$ = 301 nm; data not shown). However, no tropolone sesquiterpenoids were identified in the crude extracts of any transformants (Figure 5.16).

### 5.3.11. #Exp. 8: Expression of *eupR1, eupR3* and *eupL4*

Notably, #Exp. 7 did not include the short-chain dehydrogenase gene required for the reduction of stipitaldehyde **167** to the corresponding alcohol, stipitol **172**. In a follow-up experiment the short-chain dehydrogenase encoding gene *eupL4* was included in the heterologous expression experiment. Transformation of *A. oryzae* NSAR1 with six genes (*asPKS1* + *asL1* + *asL3* + *eupL4* + *eupR1* + *eupR3*) afforded 10 transformants. As observed for #Exp. 7, early tropolone intermediates were identified. However, no tropolone sesquiterpenoids were identified in any of the 10 transformants (Figure 5.17).

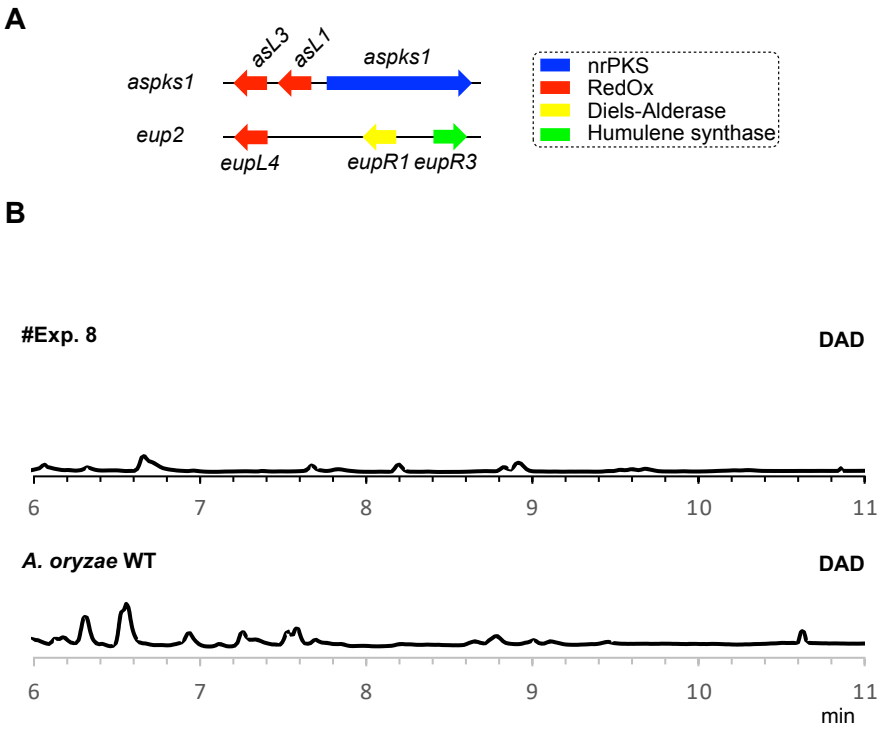



**Figure 5.17** Overview of #Exp. 8: **A,** co-expression of six genes from the *asPKS1* and *eup2* BGC; **B,** LCMS DAD chromatograms of organic extracts of *A. oryzae* expression strain and wild-type control (arbitrary units).

### 5.3.12. #Exp. 9: Expression of *eupR1*, *eupR3*, *eupR5*, *eupR6*, *eupL4*

With expression experiments #Exp. 7 and #Exp. 8 both being unsuccessful in the generation of tropolone sesquiterpenoids, further genes from the *eup2* BGC were included in the expression system. Notably, HU *et al*. had demonstrated *in vitro* that the DIELS-ALDERase from the

eupenifeldin pathway accepts hydroxyhumulene as a substrate, suggesting that hydroxylation might be required for recognition by the DIELS-ALDER enzyme EupR1.[175]

Accordingly, in #Exp. 9 the known humulene hydroxylase encoding gene *eupR6* was included. Furthermore, the FAD-dependent monooxygenase encoding gene *eupR5* was included as well. Transformation of *A. oryzae* NSAR1 with eight genes afforded 15 transformants. Submerged fermentation in DPY medium for 7 d was followed by the analysis of crude extracts by LCMS. Again, the well-known tropolone precursor **166** was readily identified (data not shown) but no tropolone sesquiterpenoids were observed (Figure 5.18).

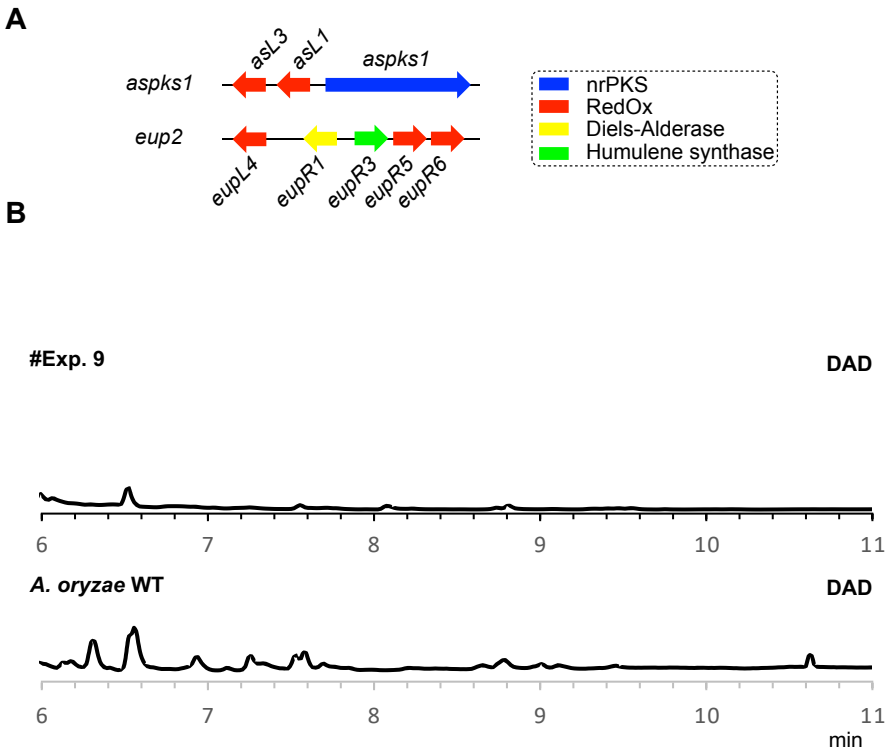



**Figure 5.18** Overview of #Exp. 9: **A,** co-expression of eight genes from the *asPKS1* and *eup2* BGC; **B,** LCMS DAD chromatograms of organic extracts of *A. oryzae* expression strain and wild-type control (arbitrary units).

### 5.3.13. #Exp. 10: Expression of *pycR1* and *pycR6*

Previous experiments #Exp. 7-9 were not successful in the generation of tropolone sesquiterpenoids. Most likely one of the transformed genes from the *eup2* BGC is inactive in *A. oryzae* NSAR1. Alternatively, additional enzymes might be required for the production of eupenifeldin-type natural products. In the former case, problems arising from individual genes could be circumvented by transforming homologous genes from the *pyc* BGC. *pycR1* and *pycR6* encode the homologous DIELS-ALDERase and humulene synthase from the *pyc* BGC.

Transformation of *A. oryzae* NSAR1 with *asPKS1, asL1, asL3, pycR1* and *pycR6* afforded 34 transformants. It is noteworthy that in this experiment the short-chain dehydrogenase encoding gene *eupL4*, required for reduction of key intermediate stipitaldehyde **167** to the corresponding alcohol **172**, was not co-transformed. Analysis of all transformants revealed the production of two new metabolites **299** and **300** (Figure 5.19 A + B).

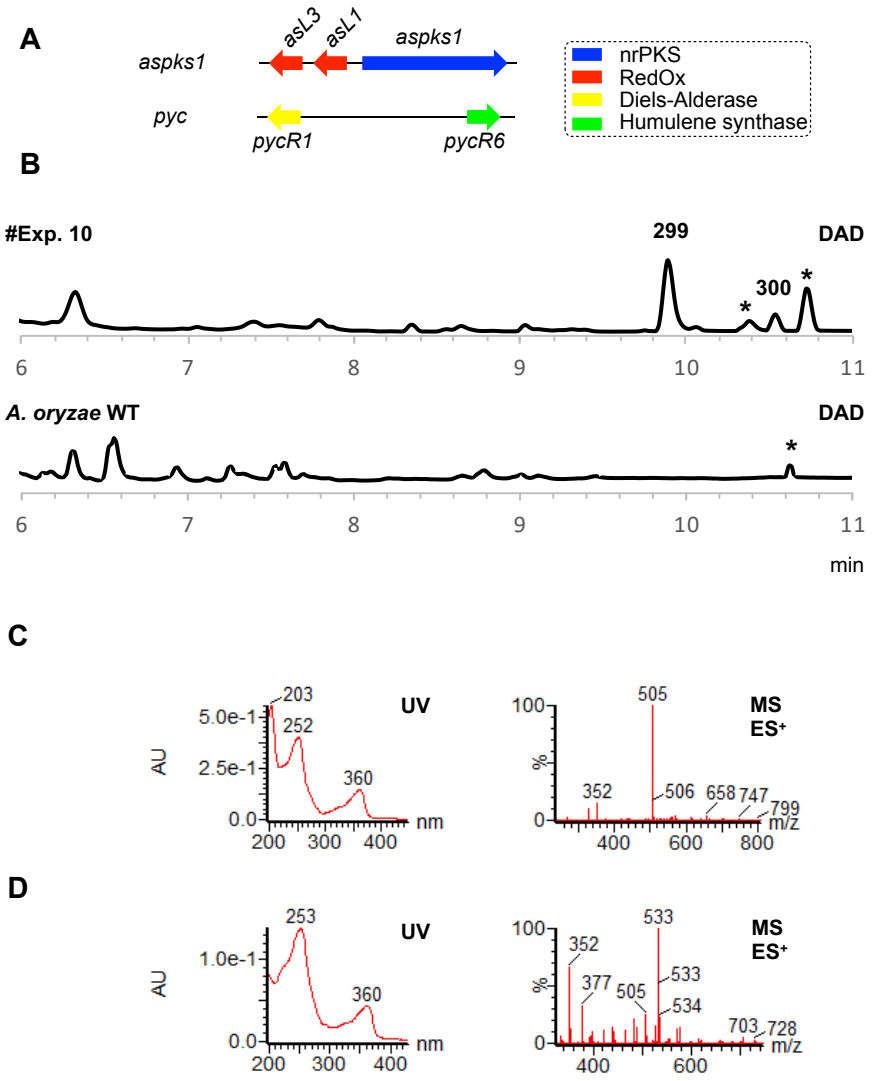



**Figure 5.19** Overview of #Exp. 10: **A,** co-expression of five genes from the *asPKS1* and *pyc* BGC; **B,** LCMS DAD chromatograms of organic extracts of *A. oryzae* expression strain and wild-type control (arbitrary units); * = unrelated metabolites; **C,** chemical characterisation of **299** by UV and mass spectrum obtained from $ES^+$; **D,** chemical characterisation of **300** by UV and mass spectrum obtained from $ES^+$.

The dominant new metabolite **299** had distinctive UV absorption maxima ($\lambda$ = 203 nm, 252 nm, 360 nm; Figure 5.19 C). HRMS analysis revealed the chemical formula $C_{32}H_{40}O_5$ ($[M]H^+$ $C_{32}H_{41}O_5$ calculated 505.2954, found 505.2941). Isolation of 4.0 mg from 1L submerged fermentation and full NMR characterization (Figure 5.20; Table 5.8) revealed key NMR signals (*e.g.*, aromatic protons at $\delta_H$ = 7.08 ppm and $\delta_H$ = 7.22 ppm; methyl singlet at $\delta_H$ = 2.41 ppm) that indicated the presence of one tropolone nucleus. HMBC correlations from $H_2$-9′ to C-7, C-

9, C-4′ and C-6′ confirmed that the tropolone ring was connected to the eastern side of humulene. Further aromatic protons at $\delta_H = 6.14$ ppm and $\delta_H = 6.25$ ppm and a characteristic methyl singlet at $\delta_H = 2.21$ ppm suggested the presence of a hydroxybenzyl instead of a second tropolone nucleus. Key HMBC correlations from $H_2$-8′′ to C-2, C-11, C-3′′ and C-5′′ confirmed the structure and placed the benzene ring at the western face of humulene. The relative stereochemistry was analysed by 2D NOESY NMR.

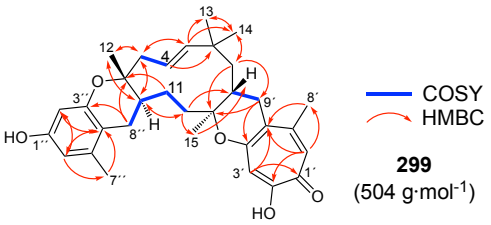



**Figure 5.20** Characterization of **299** by NMR spectroscopy (key COSY and HMBC correlations highlighted).

**Table 5.8** NMR data for **299** in $CDCl_3$ (500 MHz) referenced to $CDCl_3$.

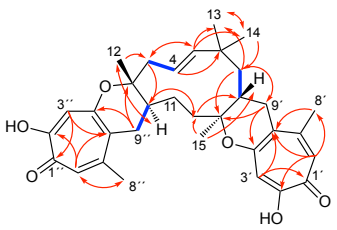
| Pos. | $\delta_H$ / ppm (mult, *J* in Hz) | $\delta_C$ / ppm | C-type | HMBC (H to C) | NOESY |
|---|---|---|---|---|---|
| **1** | 1.52 (m) | 41.1 | CH | 2, 3, 10, 11, 12, 8′′ | |
| **2** | - | 79.9 | C | - | |
| **3a** | 2.59 (dd, 13.5, 4.0) | 48.7 | $CH_2$ | 1, 2, 4, 5, 12 | 12 |
| **3b** | 2.39 (m) overlap | | | | |
| **4** | 5.54 (m) | 122.1 | CH | 3, 6 | |
| **5** | 5.55 (m) | 142.9 | CH | 3, 4, 6, 13, 14 | |
| **6** | - | 36.6 | C | - | |
| **7a** | 1.56 (m) | 47.3 | $CH_2$ | 6, 9, 14, 9′ | |
| **7b** | 1.11 (m) overlap | | | - | |
| **8** | 1.89 (m) | 30.3 | CH | - | 14, 9′ |
| **9** | - | 81.4 | C | - | |
| **10a** | 2.00 (m) | 38.2 | $CH_2$ | 8, 9, 11, 15 | 10b, 15 |
| **10b** | 1.66 (ddd, 15.0, 12.2, 2.5) | | | 1, 11 | |
| **11a** | 1.86 (m) | 25.1 | $CH_2$ | 1, 2, 8′′ | 11b |
| **11b** | 1.09 (m) overlap | | | - | |
| **12** | 1.16 (s) | 18.5 | $CH_3$ | 1, 2, 3 | 3a, 11a, 8′′b |
| **13** | 1.07 (s) | 31.0 | $CH_3$ | 5, 6, 7, 14 | 7a |
| **14** | 1.12 (s) | 23.1 | $CH_3$ | 5, 6, 7, 13 | 8 |
| **15** | 1.13 (s) | 21.3 | $CH_3$ | 8, 9, 10 | 10a |
| **1′** | - | 171.8 | C | - | |
| **2′** | - | 163.1 | C | - | |
| **3′** | 7.08 (s) | 113.8 | CH | 1′, 2′, 4′, 5′ | |
| **4′** | - | 161.6 | C | - | |
| **5′** | - | 121.5 | C | - | |
| **6′** | - | 150.6 | C | - | |
| **7′** | 7.22 (s) | 124.6 | CH | 1′, 2′, 5′, 6′, 8′ | |
| **8′** | 2.41 (s) | 27.4 | $CH_3$ | 5′, 6′, 7′ | |
| **9′a** | 2.69 (dd, 18.7, 5.6) | 34.1 | $CH_2$ | 8, 9, 4′, 5′, 6′ | |
| **9′b** | 2.39 (m) overlap | | | - | 1 |
| **1′′** | - | 154.5 | C | - | |
| **2′′** | 6.14 (d, 2.1) | 101.3 | CH | 1′′, 3′′, 4′′, 6′′ | |
| **3′′** | - | 154.1 | C | - | |
| **4′′** | - | 113.1 | C | - | |
| **5′′** | - | 137.9 | C | - | |
| **6′′** | 6.25 (d, 2.1) | 108.9 | CH | 2′′, 4′′, 7′′ | 7′′ |
| **7′′** | 2.21 (s) | 19.4 | $CH_3$ | 4′′, 5′′, 6′′ | |
| **8′′a** | 2.72 (dd, 16.6, 5.6) | 28.8 | $CH_2$ | 1, 2, 3′′, 4′′, 5′′ | 1 |
| **8′′b** | 2.18 (m) | | | 1, 11, 3′′, 4′′, 5′′ | 1, 12 |

Absence of correlations between H-1 and $H_3$-12 and between H-8 and $H_3$-15 established the stereochemistry at the two dihydropyran-humulene ring junctions as *anti*, in agreement with the humulene backbone being derived from all-*trans* α-humulene **2**.[8] Compound **299** displayed an optical rotation of $[\alpha]_D^{30} = +355.6$ ($c$ = 0.1125 g/100 mL, $CHCl_3$).

A minor co-metabolite was sporadically observed in extracts of representative transformants. Compound **300** had characteristic UV absorption maxima ($\lambda$ = 253 nm, 360 nm; Figure 5.19 D) that were identical to other compounds containing intact tropolone nuclei. The nominal mass of 532 was consistent with **300** being a dehydroxy analogue of a bistropolone.

Isolation of 1.4 mg from 1L submerged fermentation was sufficient for NMR analysis (Figure 5.21; Table 5.9). Four characteristic aromatic proton signals at $\delta_H$ = 7.33 ppm, $\delta_H$ = 7.27 ppm, $\delta_H$ = 7.17 ppm and $\delta_H$ = 7.15 ppm together with two aromatic methyl singlets at $\delta_H$ = 2.50 ppm and $\delta_H$ = 2.43 ppm suggested the presence of two tropolone nuclei. COSY and HMBC correlations validated the structure of **300**. The relative stereochemistry was not investigated as the quality of 1D- and 2D-NOESY data was poor; however, biosynthetic considerations and very similar chemical shifts suggested the same relative stereochemistry as observed in pycnidione **9**.

The proposed biosynthetic pathway leading to formation of **299** and **300** in *A. oryzae* NSAR1 is displayed in Scheme 5.9. In the absence of a co-transformed short-chain dehydrogenase activation of stipitaldehyde **167** must be catalysed by *A. oryzae* NSAR1. PycR1 – unlike AsR5 – catalyses two rounds of hetero DIELS-ALDER reactions, giving rise to **300**. No additional DIELS-ALDER enzyme is required for production of bistropolones. Observation of **299** suggests that *A. oryzae* NSAR1 itself catalyses a ring-contraction, as previously observed in #Exp. 5.



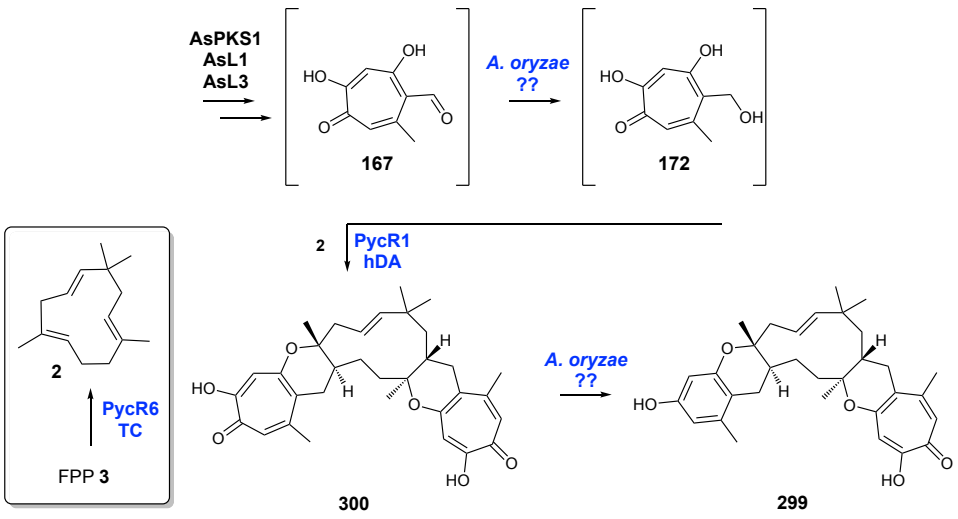

**300**
(532 g·mol$^{-1}$)

**Figure 5.21:** Characterization of **300** by NMR spectroscopy (key COSY and HMBC correlations highlighted).

[8] LUCA CODUTTI is thanked for processing/phasing of 2D NOESY data obtained for compound **299**.

**Table 5.9** NMR data for **300** in $CDCl_3$ (500 MHz) referenced to $CDCl_3$.

| Pos. | $\delta_H$ / ppm (mult, *J* in Hz) | $\delta_C$ / ppm | C-type | HMBC (H to C) | COSY |
|---|---|---|---|---|---|
| **1** | 1.50 (m) | 40.9 | CH | 2, 10, 11, 12 | - |
| **2** | - | 82.4 | C | - | - |
| **3a** | 2.70 (m) | 48.5 | $CH_2$ | 1, 5 | |
| **3b** | 2.33 (m) overlap | | | | |
| **4** | 5.50 (m) | 121.3 | CH | 3, 6 | 3 |
| **5** | 5.52 (m) | 143.8 | CH | 3, 4, 6, 13, 14 | - |
| **6** | - | 36.9 | C | - | - |
| **7a** | 1.56 (m) | 46.8 | $CH_2$ | 6, 9, 13, 9′′ | 8 |
| **7b** | 1.12 (m) | | | - | - |
| **8** | 1.85 (m) | 30.3 | CH | 9 | - |
| **9** | - | 81.1 | C | | - |
| **10a** | 2.02 (m) | 37.9 | $CH_2$ | 8, 9 | - |
| **10b** | 1.66 (m) | | | - | - |
| **11a** | 1.85 (m) | 24.8 | $CH_2$ | | |
| **11b** | 1.64 (m) | | | | |
| **12** | 1.25 (s) | 19.0 | $CH_3$ | 1, 2, 3 | - |
| **13** | 1.12 (s) | 22.5 | $CH_3$ | 5, 6, 7, 14 | - |
| **14** | 1.08 (s) | 31.4 | $CH_3$ | 5, 6, 7, 13 | - |
| **15** | 1.15 (s) | 21.3 | $CH_3$ | 8, 9, 10 | - |
| **1′** | - | 171.0 | C | - | - |
| **2′** | - | 163.9 | C | - | - |
| **3′** | 7.17 (s) | 114.6 | CH | 1′, 2′, 5′ | - |
| **4′** | | 161.9 | C | | |
| **5′** | - | 123.2 | C | - | - |
| **6′** | - | 150.4 | C | - | - |
| **7′** | 7.33 (s) | 124.5 | C | 1′, 2′, 5′, 8′ | |
| **8′** | 2.50 (s) | 27.5 | $CH_3$ | 6′, 7′ | - |
| **9′a** | 2.82 (m) | 32.6 | $CH_2$ | 1, 2, 4′, 5′, 6′ | 1 |
| **9′b** | 2.34 (m) overlap | | | | - |
| **1′′** | - | 171.7 | C | - | - |
| **2′′** | - | 163.2 | C | - | - |
| **3′′** | 7.15 (s) | 113.8 | CH | 1′′, 2′′, 5′′ | - |
| **4′′** | | 161.7 | C | | |
| **5′′** | - | 121.8 | C | | |
| **6′′** | - | 150.9 | C | - | - |
| **7′′** | 7.27 (s) | 124.7 | CH | 1′′, 2′′, 5′′, 8′′ | |
| **8′′** | 2.43 (s) | 27.4 | $CH_3$ | 5′′, 6′′, 7′′ | - |
| **9′′a** | 2.68 (m) | 34.1 | $CH_2$ | 8, 9, 4′′, 5′′, 6′′ | - |
| **9′′b** | 2.38 (m) overlap | | | | |



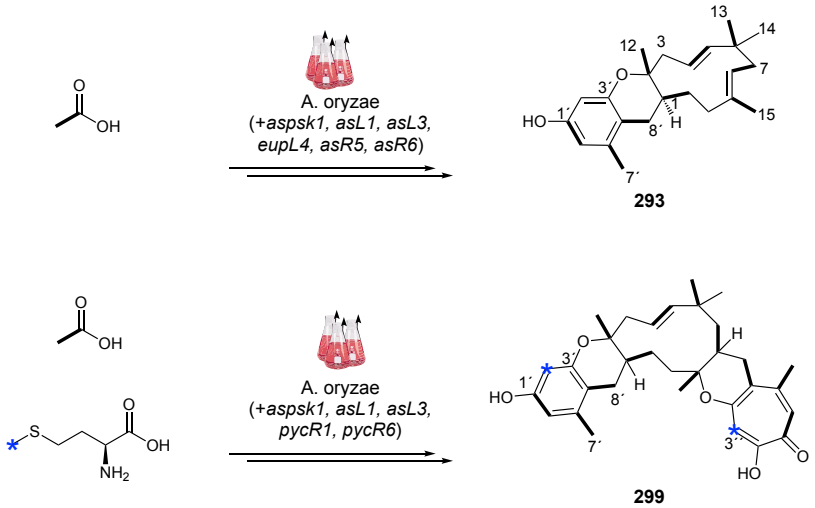

**Scheme 5.9** Formation of **299** and **300** in *A. oryzae* NSAR1 upon expression of *asPKS1, asL1, asL3, pycR1* and *pycR6*; compounds in brackets not observed.;TC = humulene synthase; hDA = hetero DIELS-ALDERase.

## 5.4. Isotopic Labelling of Key Tropolone Sesquiterpenoids

Heterologous expression experiments #Exp. 5 and #Exp. 10 afforded tropolone sesquiterpenoids with benzene rings in place of the expected tropolone nucleus. As no gene candidate was co-transformed that was likely to catalyze a ring-contraction it is reasonable that *A. oryzae* NSAR1 itself harbors the potential to catalyze this reaction. Alternatively, as for the wild-type strain *Phaeosphaeriaceae* sp. CF-150626, there is the possibility that an endogenously produced benzene substrate is activated and transformed by the DIELS-ALDERase AsR5/PycR1.

To investigate the two possibilities labelling experiments were performed using the stable isotopes [1,2-$^{13}C_2$]-sodium acetate and [*methyl*-$^{13}$C]-methionine. In a first experiment the **293**-producing strain was fed with [1,2-$^{13}C_2$]-acetate. For this, a 0.5 L culture of the best producing transformant was fed with [1,2-$^{13}C_2$]-acetate to a final concentration of 12 mM across five days. Mass-directed purification of [1,2-$^{13}C_2$]-labelled **293** afforded ~6 mg that was analyzed by $^{13}$C-NMR (Figure 5.23).

As carbon atoms derived from an intact acetate unit show distinct coupling constants the incorporation of intact acetate units was rapidly identified. Nine intact acetate units were integrated into the carbon skeleton of **293** (Table 5.10; Figure 5.22 +5.23). As expected, six out of nine acetate units were incorporated into the humulene macrocycle (Scheme 5.10). Three intact acetate units were incorporated into the tetraketide portion, with intact labelling between C-7′/C-5′, C-6′/C-1′ and C-4′/C-8′. Neither carbon C-2′ or C-3′ derive from intact acetates, ruling out the possibility that AsR5 activated an orsellinic-acid derived six-membered precursor.




**Scheme 5.10** Isotopic labelling of **293** and **299** with [1,2-$^{13}C_2$]-acetate and [*methyl*-$^{13}$C]-labelled methionine.

In parallel work the **299**-producing strain was also fed with [1,2-$^{13}C_2$]-acetate. Isolation and purification afforded 5.5 mg labelled **299**. $^{13}$C-NMR (Table 5.10; Figure 5.22 + 5.24) identified

the same labelling pattern of the hydroxybenzene as already observed for **293**, with incorporation of only three intact acetate units (Scheme 5.10). C-2′ and C-3′ were likewise found not to be an intact acetate unit. Feeding of the **299**-producing strain with [*methyl*-$^{13}$C]-labelled methionine afforded 7 mg of labelled **299**. $^{13}$C-NMR analysis revealed two carbon signals to be enhanced, corresponding to C-3′′ within the tropolone ring and C-2′ within the hydroxybenzene ring (Figure 5.25; Scheme 5.10). Together the data conclusively demonstrate that the six-membered ring must be derived by a ring-contraction of a tropolone precursor.

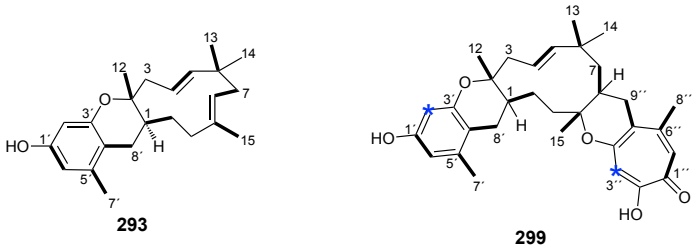



**Figure 5.22**: [1,2-$^{13}C_2$]-labelled **293** and **299**.

**Table 5.10** NMR data for labelled **293** and **299** in $CDCl_3$ (500 MHz) referenced to $CDCl_3$

| | Compound **293** | | Compound **299** | |
|---|---|---|---|---|
| **Pos.** | **$\delta_C$ type** | **Coupling constant *J* / Hz** | **$\delta_C$ type** | **Coupling constant *J* / Hz** |
| 1 | 35.9 | 18.1 Hz | 41.1 | 35.8 |
| 2 | 79.9 | 19.6 Hz | 79.9 | 39.5 |
| 3 | 43.1 | - | 48.7 | - |
| 4 | 121.1 | 36.9 Hz | 122.1 | 72.2 |
| 5 | 142.1 | 36.9 Hz | 142.9 | 72.2 |
| 6 | 38.4 | 18.4 Hz | 36.6 | 35.7 |
| 7 | 41.4 | 22.3 Hz | 47.3 | 36.7 |
| 8 | 123.2 | 22.3 Hz | 30.3 | 36.7 |
| 9 | 136.7 | 21.6 Hz | 81.4 | 39.8 |
| 10 | 38.1 | - | 38.2 | - |
| 11 | 30.8 | 18.1 Hz | 25.1 | 35.3 |
| 12 | 20.0 | 19.6 Hz | 18.5 | 39.4 |
| 13 | 30.5 | - | 31.0 | 35.3 |
| 14 | 24.5 | 18.4 Hz | 23.1 | - |
| 15 | 17.3 | 21.8 Hz | 21.3 | 39.8 |
| 1‘ | 154.4 | 32.9 Hz | 154.5 | 66.0 |
| 2‘ | 101.5 | - | 101.3 | - |
| 3‘ | 154.9 | 32.1 Hz | 154.1 | 64.5 |
| 4‘ | 113.2 | 21.6 Hz | 113.1 | 43.0 |
| 5‘ | 137.9 | 22.0 Hz | 137.9 | 43.9 |
| 6‘ | 108.6 | 32.9 Hz | 108.9 | 66.0 |
| 7‘ | 19.4 | 22.0 Hz | 19.4 | 43.9 |
| 8‘ | 27.1 | 21.5 Hz | 28.8 | 43.0 |
| 1′′ | | | 171.8 | 66.5 |
| 2′′ | | | 163.1 | 9.0 |
| 3′′ | | | 113.8 | - |
| 4′′ | | | 161.6 | 9.1 |
| 5′′ | | | 121.5 | 41.4 |
| 6′′ | | | 150.6 | 41.8 |
| 7′′ | | | 124.6 | 66.4 |
| 8′′ | | | 27.4 | 41.8 |
| 9′′ | | | 34.1 | 41.5 |

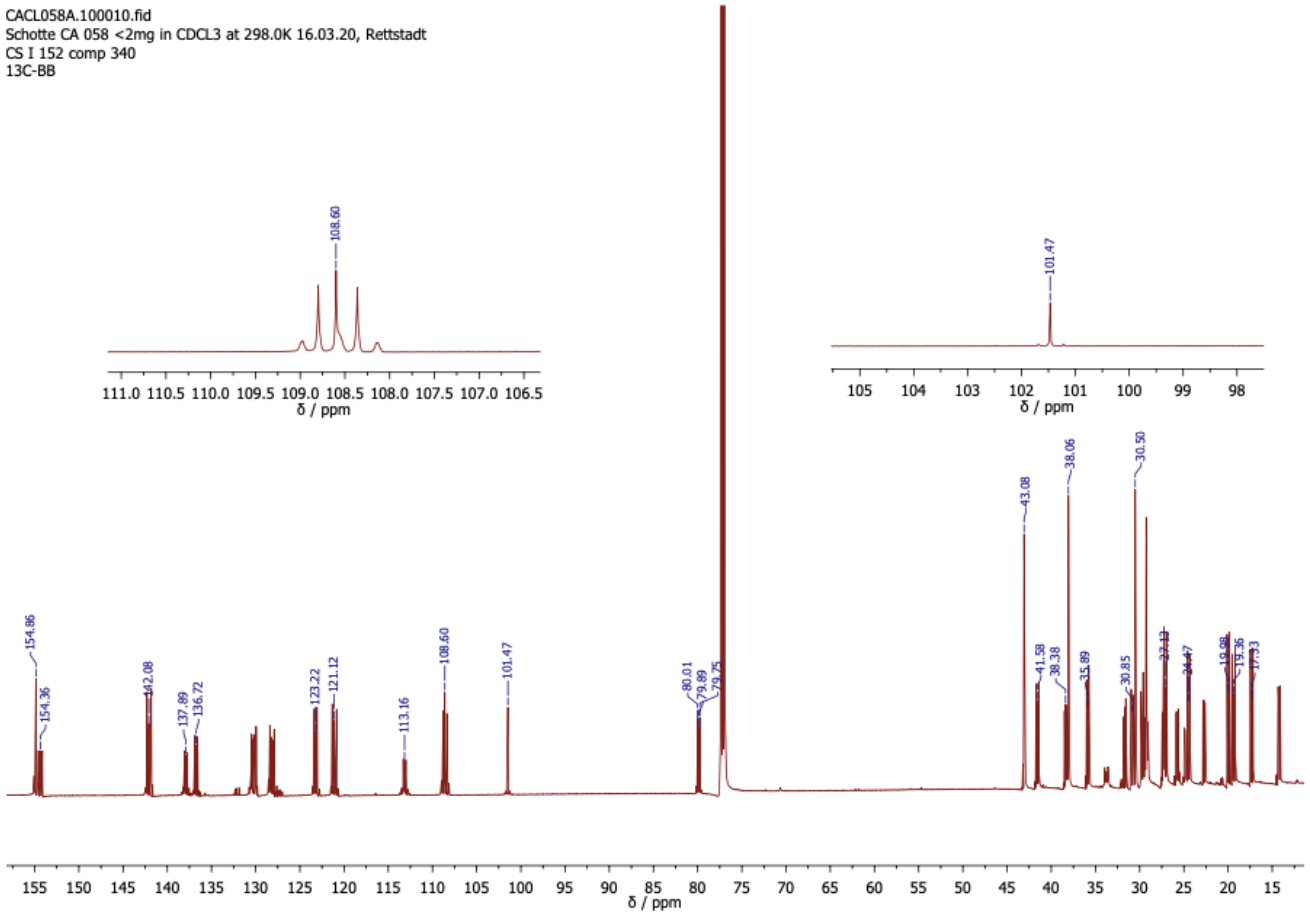



**Figure 5.23** $^{13}$C-NMR spectrum for [1,2-$^{13}C_2$]-labelled **293** recorded at 600 MHz in $CDCl_3$.

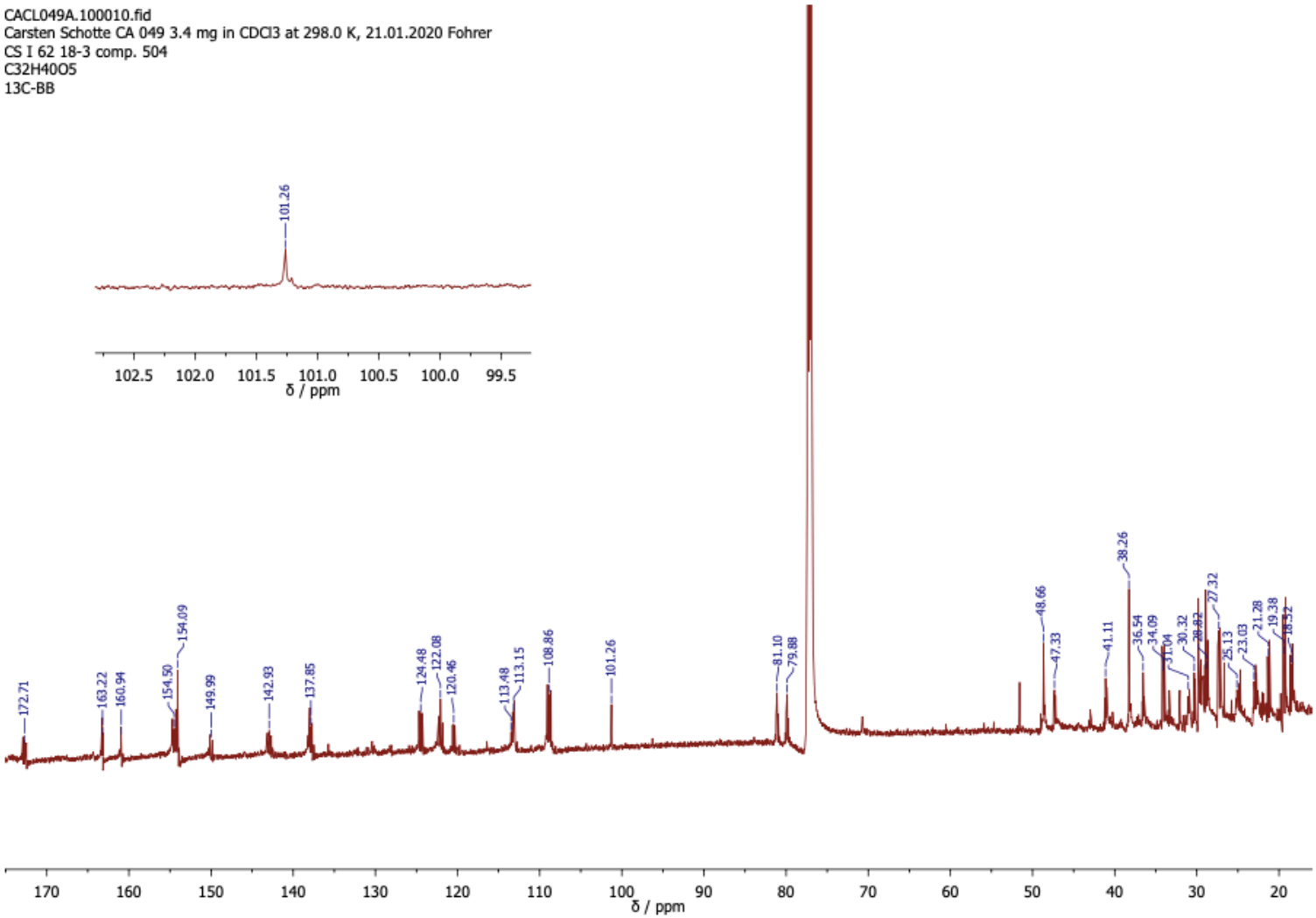



**Figure 5.24** $^{13}$C-NMR spectrum for **299** labelled with [1,2-$^{13}C_2$]-acetate recorded at 600 MHz in CDCl3.

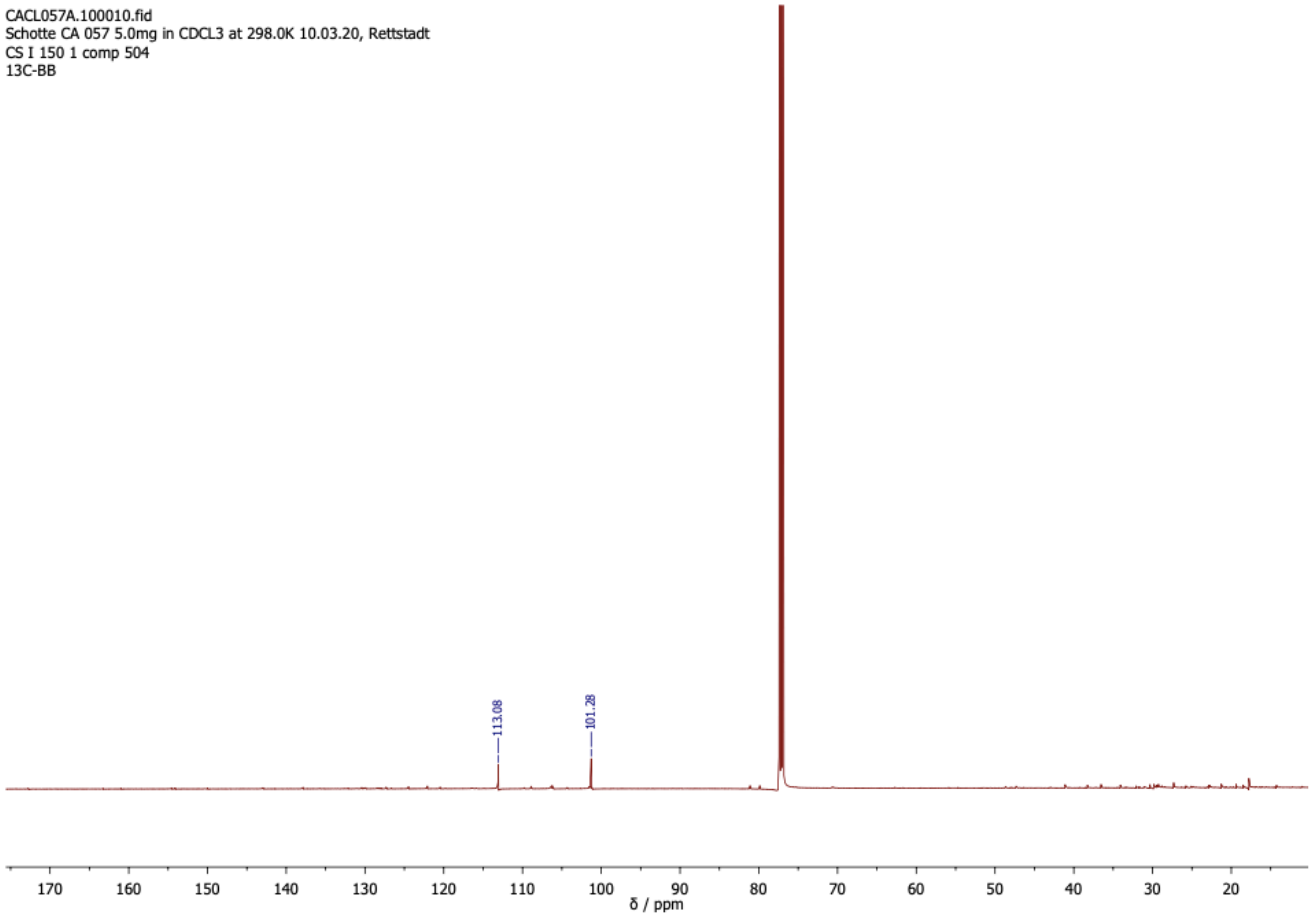



**Figure 5.25** $^{13}$C-NMR spectrum for **299** labelled with [*methyl*-$^{13}$C]-methionine recorded at 600 MHz in $CDCl_3$.

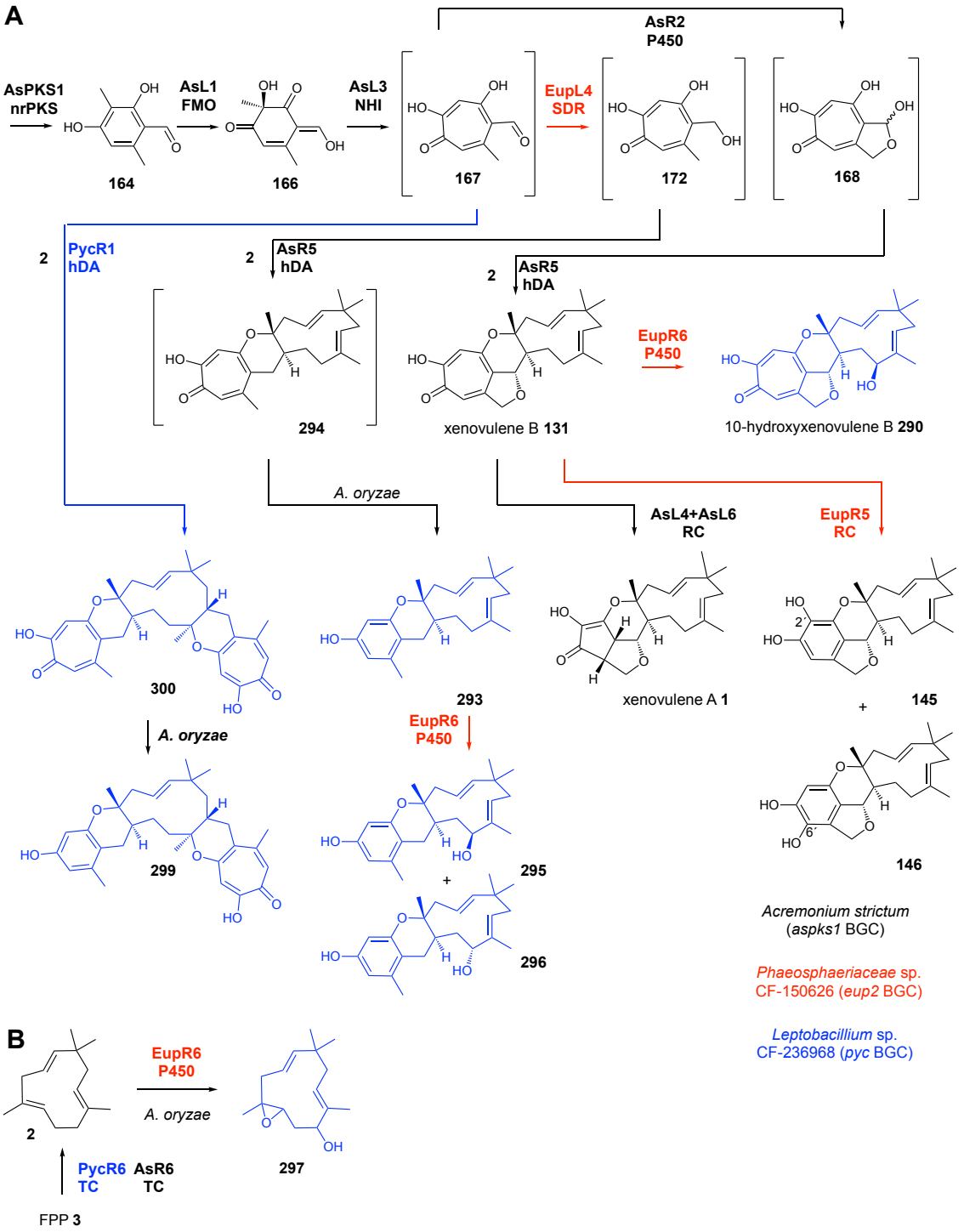
## 5.5. Conclusion, Outlook and Discussion

The aim of this part of the study was to systematically expand the chemical space around tropolone sesquiterpenoids and to generate new, unnatural derivatives. In the course of this study ten different expression experiments in the heterologous host *A. oryzae* NSAR1 were performed. A total of seven new derivatives of xenovulene A **1** and pycnidione **9** were isolated, purified to homogeneity and fully characterised. Additionally, two previously reported phenolic xenovulenes, **145** and **146**, were obtained by deploying different biosynthetic enzymes (EupR5 instead of AsL4/AsL6). It should be noted that there are very few chemical syntheses of this class of metabolites and the present study thus significantly expands the chemical space of tropolone sesquiterpenoids. Only after this work was complete the SARLAH lab published a synthetic modular cycloaddition-based approach to a number of mono- and bistropolones from the pycnidione series.[257] New-To-Nature derivatives belonging to the eupenifeldin series of tropolone sesquiterpenoids, however, have not yet been (bio)synthesized and remain an attractive target for further studies.

This is the first time that disubstituted tropolone sesquiterpenoids (**299** and **300**) and benzopyranyl-humulenes (**293**, **295** and **296**) were successfully biosynthesized and engineered *in vivo*. A summary of the engineered pathways is displayed below (Scheme 5.11). The deployed synthetic biology-based approach successfully exploits the common biosynthetic origin of all tropolone sesquiterpenoids, stipitaldehyde **167**. Aldehyde **167** is the key intermediate from which all subsequent reconstituted pathways diverge: Rational inclusion of either AsR2 or EupL4 determines the overall pathway direction. Inclusion of AsR2 *oxidatively* directs the pathways towards the xenovulenes. Inclusion of EupL4 *reductively* directs the pathway towards the benzopyranyl-humulenes. The subsequent choice of DIELS-ALDERases (AsR5 *vs* PycR1) allows to differentiate between mono- and disubstituted DIELS-ALDER adducts.

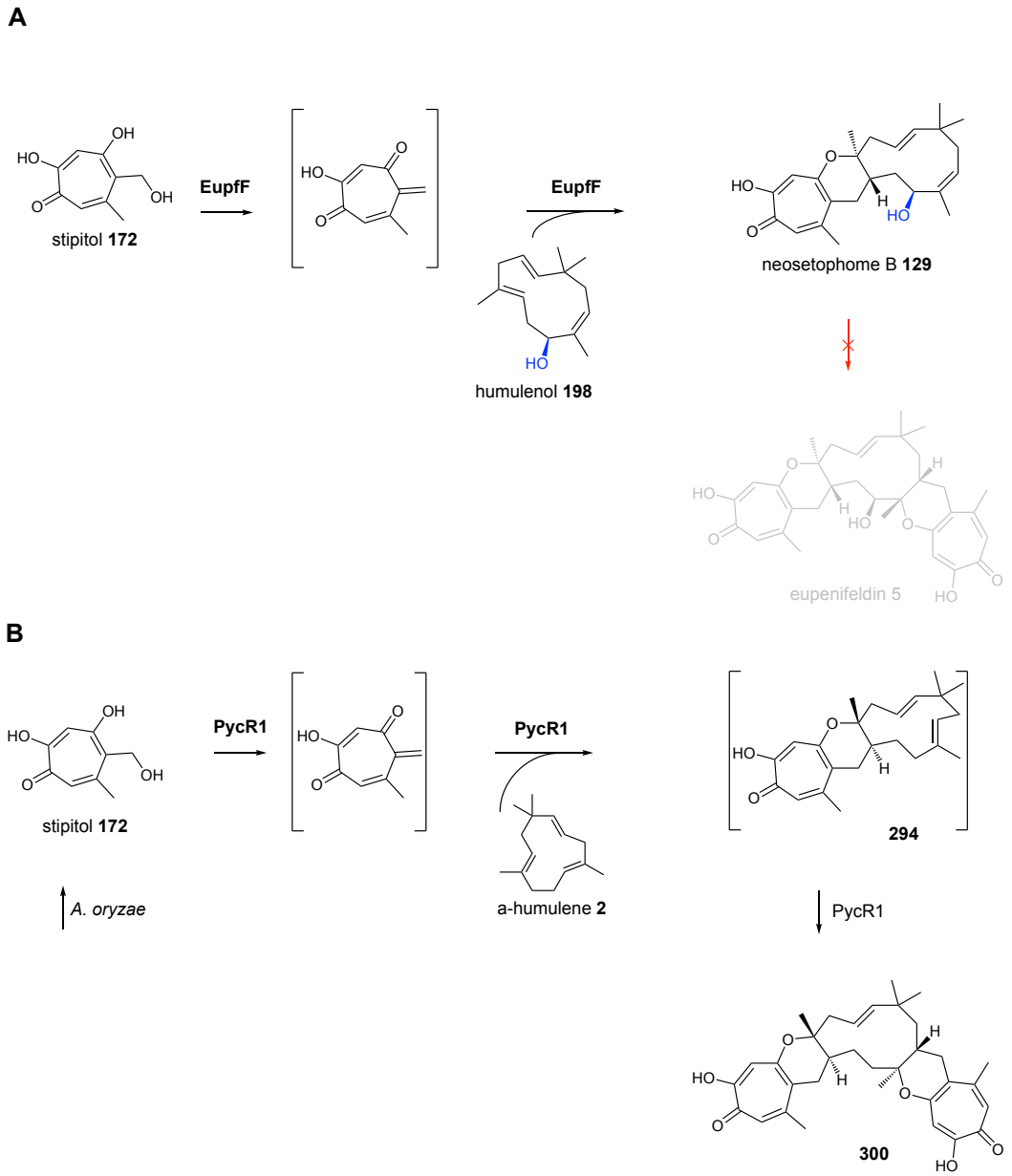
### 5.5.1. DIELS-ALDER Chemistry in Bistropolone Formation

One of the most intriguing questions of bistropolone biosynthesis is the number of enzymes required for the two successive hetero DIELS-ALDER reactions that first produce monotropolones such as neosetophome B **129** and then bistropolones such as eupenifeldin **5**. HU *et al*. proposed that two different enzymes catalyse the two reactions.[175] This proposal was based on extensive *in vitro* and *in vivo* studies with the hetero DIELS-ALDERase EupfF from the *eupf* BGC: *In vitro* assays with recombinant EupfF in the presence of stipitol **172** and humulenol **198** solely afforded the monotropolone neosetophome B **129** (Scheme 5.12A).[175]




**Scheme 5.11** Overview of engineered biosynthetic pathways: **A,** isolated tropolone sesquiterpenoids upon heterologous expression; each indicated path represents a separate expression experiment in *A. oryzae* NSAR1; compounds in blue are newly generated in this study and characterised for the first time; compounds in brackets were not directly observed; **B,** humulene formation and formation of phomanoxide B **297**.

Feeding of humulenol **198** to *A. nidulans* transformed with *eupfF* likewise resulted only in the formation of the monotropolone **129**.[175]




**Scheme 5.12** Possible implication of humulene-hydroxylation in bistropolone formation: **A,** *in vitro* reconstitution of recombinant EupfF affords monotropolone **129**;[175] **B,** proposed pathway leading to bistropolone **300** upon expression of *pycR1* in *A. oryzae* NSAR1; compounds in brackets/grey not observed.

The observation of the bistropolone **300** and its ring-contracted congener **299** upon transformation of the single DIELS-ALDERase PycR1 in this study, however, suggests that a single enzyme catalyses both DIELS-ALDER reactions during bistropolone formation in *A. oryzae* (Scheme 5.12 B). In addition to deploying different enzymes in this study (from the pycnidione pathway) formation of bistropolones was observed in the absence of humulene hydroxylase encoding gene *eupR6*. It is noteworthy to emphasize that HU *et al.* did not probe for the conversion of the non-hydroxylated terpene 8*Z*-humulene **7**, as this could not be isolated.[175]

As the exact timing of humulene-hydroxylation in tropolone sesquiterpenoid formation remains enigmatic, it is possible that humulene hydroxylation affects the second DIELS-ALDER reaction. Absence of bistropolones in the experiments performed by HU *et al.*, combined with our data, suggests that hydroxylation of the humulene macrocycle inhibits the subsequent second DIELS-ALDER reaction. The isolation of non-hydroxylated bistropolones from wild-type organisms, both in this study (*e.g.* dehydroxyeupenifeldin **207**) and from *Neosetophoma* sp. (*e.g.* dehydroxyeupenifeldin **207**) and *Neosetophoma* sp. RKD0834 (*e.g.* dehydroxypycnidione), support the hypothesis in this form and suggest the humulene hydroxylation occurs late in the pathway.[81,258]

### 5.5.2. Tailoring Genes in Tropolone Sesquiterpenoid Formation

In the course of the conducted experiments genes from three different fungal BGC were successfully used to engineer New-To-Nature tropolone sesquiterpenoids. It is remarkable that tailoring enzymes such as the humulene hydroxylase EupR6 and the ring-contracting enzyme EupR5 apparently act on the xenovulene `base´ structure and thus display significant substrate promiscuity. However, the conducted experiments also revealed the limitations of possible engineering attempts. Introduction of the 8*Z*-humulene **7** producing terpene cyclase EupR3 into the xenovulene pathway (instead of AsR6) did not result in any xenovulenes with a C-8/C-9 *Z*-configured alkene. This finding suggests that the hetero DIELS-ALDERase AsR5 is `primed´ for the native α-humulene **2** substrate and does not accept the geometric isomer **7**. Likewise, Hu *et al.* concluded that the DIELS-ALDERase EupfF from the eupenifeldin biosynthetic pathway does not recognize the unnatural α-humulene **2** scaffold.[175] In future engineering attempts it is thus strongly recommended to transform matching pairs of humulene synthase and DIELS-ALDERase from the same biosynthetic pathway.

With EupR6 a promiscuous humulene hydroxylase was described in this study. Apart from the observed hydroxylations in the wild-type strains it was demonstrated to hydroxylate a range of xenovulenes and benzopyranyl-humulenes. With compound **295** and **296** the first tropolone sesquiterpenoids were isolated where the hydroxylation afforded both the 10*R* and the 10*S* stereoisomer (#Exp. 6). Notably, in the separate experiment #Exp. 1, solely the 10*S*-hydroxyxenovulene B **290** was obtained. Since the DIELS-ALDER enzyme AsR5 was transformed in both experiments this finding further supports a mechanism for EupR6, where the hydroxylation of humulene occurs late after the DIELS-ALDER reaction. Depending on the nature of the tropolone sesquiterpenoid intermediate different hydroxylations can occur.

In addition to the generation of new tropolone sesquiterpenoids the combinatorial biosynthesis approach successfully delineated a function for EupR5. The FAD-dependent monooxygenase was

shown to catalyse the oxidative ring-contraction from xenovulene B **131** to the benzylic derivatives **145** and **146**. In contrast to AsL4 and AsL6, EupR5 is less specific as it generates both **145** and **146** instead of just one of the ring-contracted xenovulenes. Furthermore, unlike AsL4/AsL6 EupR5 is not capable of catalysing the second ring-contraction from **145**/**146** to the 5-membered ring of xenovulene A **1**.

While EupR5 *can* catalyse oxidative ring-contractions this does necessarily lead to the conclusion that EupR5 catalyses this reaction in its native host organism. As emphasized before, the ring contraction in *Phaeosphaeriaceae* sp. must differ in terms of mechanism from the ring-contractions observed during xenovulene biosynthesis. In the xenovulenes the ring-contractions correlate with the retention of oxygen on the benzylic products **145** (hydroxy group at the C-2′) and **146** (hydroxy group at the C-6′; Scheme 5.11), whereas in noreupenifeldin B **6** only one hydroxylgroup is present (Scheme 5.13). To confirm whether EupR5 can also catalyse the ring-contraction from eupenifeldin **5** to noreupenifeldin B **6** this has to be investigated *in vitro*. In future work efforts should be directed to obtain recombinant EupR5. As heterologous expression of *eupR5* in *E. coli* was not successful and attempts to produce recombinant AsL4 and AsL6 in yeast failed, more sophisticated approaches are required to obtain recombinant EupR5.

The presence of membrane-spanning domains in AsL4, AsL6 and EupR5 suggests that these enzymes might be attached to a cell membrane after production in the host. HERTWECK *et al.* and WATANABE *et al*. both successfully isolated membrane associated enzymes from the microsomal fraction of yeast expression systems.[259,260] Alternatively, with EupR5 being active *in vivo,* expression in *A. oryzae* might be another possibility. While *A. oryzae* has not yet been extensively used as an expression host for recombinant protein production, the feasibility of this approach has been reported. For example, KASHIWAGI *et al.* successfully produced recombinant aspartyl aminopeptidase (DAP) in *A. oryzae* under control of the inducible *AmyB* promotor. A *C*-terminal $His_6$-tag allowed isolation of DAP *via* conventional IMAC technology.[261]

### 5.5.3. Influence of the Heterologous Host

*A. oryzae* NSAR was successfully deployed to reconstruct a series of artificial biosynthetic pathways, underlining again its importance as a heterologous host for the total mycosynthesis of natural products. All characterized tropolone sesquiterpenoids were isolated in multi-milligram yields that allowed for structural characterization and biotesting of the pure compounds. However, *A. oryzae* NSAR1 also adds another layer of complexity to the combinatorial biosynthesis process, by catalysing unexpected oxidative ring-contractions and reductive events.

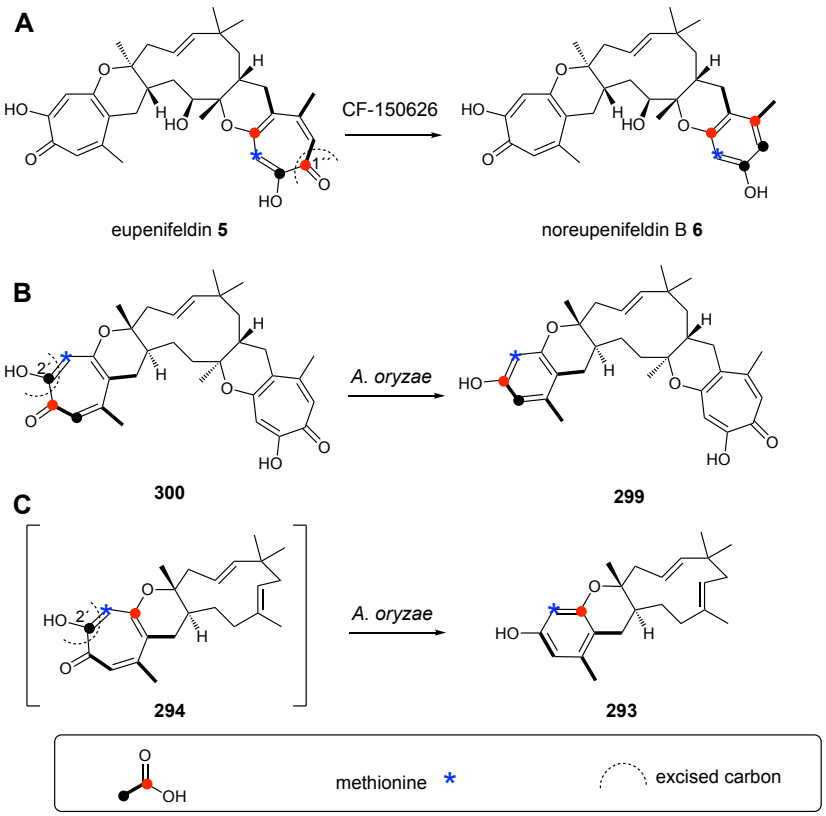
Both in #Exp. 5 and in #Exp. 10 the dominant products were benzopyranyl-humulenes (**293**, **299**). Feeding experiments demonstrated that the benzene rings were labelled consistent with the operation of tropolone ring-contractions. As no genes encoding for ring-contracting enzymes have been transformed in the respective experiments these transformations have to be catalysed by the host itself. Co-isolation of the disubstituted benzopyranyl-humulene **299** with the bistropolone **300** suggests **300** to be its precursor (Scheme 5.13 B). Likewise, monobenzopyran **293** is most likely directly derived from the unobserved intermediate **294** (Scheme 5.13 C). Most notably, the obtained benzopyranyl moiety is identical with the moieties observed in the wild-type compound noreupenifeldin B **6** and epolone C **197** – suggesting a common biosynthetic origin.

However, the labelling pattern of **293** and **299** produced in *A. oryzae* differs from the labelling pattern in noreupenifeldin B **6** produced in *Phaeosphaeriaceae* sp. CF-150626. Ring contraction in *A. oryzae* proceeds *via* excision of the C-2′ carbon in **294** and **300** (Scheme 5.13 B and C). In eupenifeldin **5** the ring-contraction proceeds *via* excision of the C-1′′ carbon (Scheme 5.13 A). An endogenous enzyme within *A. oryzae* NSAR1 must be responsible for this conversion but remains unknown at the moment. Surprisingly, xenovulene B **131** is stable in *A. oryzae* NSAR1, albeit also containing an intact tropolone ring. The only structural difference between **294**/**300** and **131** is the characteristic tetrahydrofuran ring in **131** that must prevent the ring-contraction in *A. oryzae*. Further complexity is added to the system by the fact that the observed endogenous ring-contractions are highly regioselective. In *A. oryzae* only tropolone rings on the western face of humulene are contracted. Contrasting this is the observation that in *Phaeosphaeriaceae* sp. the ring-contraction only occurs on the eastern face of humulene.

The production of the bistropolones **299** and **300** is unexpected in itself. The transformed gene set (*asPKS1* + *asL1* + *asL3* + *pycR1* + *pycR6*) should give rise to stipitaldehyde **167**. As no short-chain dehydrogenase encoding gene was co-transformed, the key reduction of **167** to stipitol **172** cannot occur. The necessary reductive step must therefore also be catalysed by the *A. oryzae* host. Similar reductions of aldehydes have been encountered in *A. oryzae* during the expression of genes involved in the biosynthesis of cytochalasins and solanapyrone.[262–264] However, this reaction is not always observed. For example, in all previous studies by RAISSA SCHOR stipitol **172** was never observed and intact stipitaldehyde **167** was isolated in decent titres from crude extracts of *A. oryzae* transformed with the tropolone core genes.[123]

The occurrence of host-catalysed ring-contractions and reductive events therefore cannot be rationalised at the moment. While the host chemistry positively contributes to the observed

diversity of tropolone sesquiterpenoids future work should be directed to identify (and control) the endogenous *A. oryzae* enzymes to allow for even more rational engineering.




**Scheme 5.13** Proposed ring-contraction in the biosynthesis of noreupenifeldin B **6**, **293** and **299** based on observed labelling patterns: **A,** ring-contraction in *Phaeosphaeriaceae* sp. CF-150626 proceeding *via* excision of carbon C-1'' in **5**; **B,** ring-contraction in *A. oryzae* proceeding *via* excision of carbon C-2' in **294** and **300** respectively.

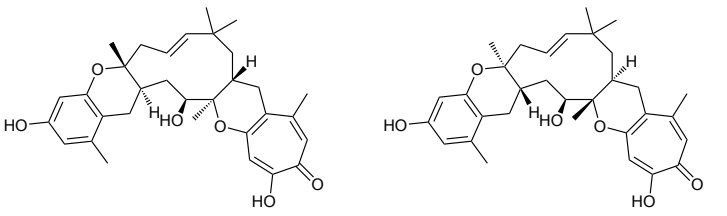
### 5.5.4. Outlook

In subsequent work the bioactivity of the newly derived compounds should be systematically assessed. Notable bioactivities of tropolone sesquiterpenoids comprise antitumor activity and the ability to induce the production of erythropoietin (EPO) in human cells.[78,85] It is very likely that the newly derived compounds display similar activity, but it remains to be seen whether any improved bioactivities can be determined.

Investigations into the ability to induce the production of EPO are currently ongoing in cooperation with the research group of PROF. LEE-THEDIECK. In initial work her research group successfully established a renal EPO-producing cells line (fibroblastoid atypical interstitial kidney [FAIK] cells). FAIK cells are responsive to oxygen concentration and under hypoxic conditions the production of EPO is induced.[265] Production of EPO can then be easily monitored using a sensitive sandwich ELISA assay. In ongoing work, the effect of the unnatural tropolone sesquiterpenoids on EPO-production will be assessed.

Future work should also be directed to further expand the created library of compounds. For example, co-expression of *pycR1* and *pycR6* already resulted in the production of bistropolones **299** and **300**. Co-expression of the cytochrome P450 encoding gene *eupR6* should principally afford the C-10 hydroxylated derivatives of **299** and **300**.

Finally, it should be noted that the relative stereochemistry of all isolated tropolone sesquiterpenoids was elucidated using extensive NOESY NMR spectroscopy and biosynthetic considerations. Eupenifeldin **5** and pycnidione **9**, the best described members of this family, were used as reference metabolites to establish the relative stereochemistry. However, in April 2021 synthetic studies by the SARLAH group led to a structural revision of pycnidione **9**: The authors proposed the alternative structure **9´** (Figure 5.26).[257] Since for most of the reported members of pycnidione-type natural products the relative stereochemistry was assigned in analogy to pycnidione **9**, the authors suggest that the relative stereochemistry of all compounds within this subclass of metabolites has to be reassigned.[257] In this thesis the original reported and proposed stereochemistry of **9** was used throughout. Future studies should systematically re-evaluate the stereochemistry of all tropolone sesquiterpenoids reported herein and in the relevant literature to ameliorate existing inconsistencies.




structure of pycnidione **9** used in this study

revised structure of pycnidione **9´** (Sarlah *et al., JACS* **2021**, *143*, 6006-6017)

**Figure 5.26**: Revised structure of pycnidione **9´** as proposed by SARLAH *et al*.[257]

# 6. Overall Conclusion and Future Experiments

In this study the biosynthesis of tropolone sesquiterpenoids was investigated using a combination of genetic, chemical, biochemical and microbiological methods. A particular focus was directed to engineer the terpene cyclase AsR6 and entire biosynthetic pathways with the aim of generating new-to-nature chemical entities.

Through the use of heterologous gene expression in the fungal host *A. oryzae* and in the bacterial host *E. coli* BL21 the minimal number and order of enzymes involved in the biosynthesis of xenovulene A **1** was delineated. Eight enzymes are required for **1**-formation. Among these, three unprecedented classes of enzymes involved in α-humulene **2** formation (AsR6), DIELS-ALDER chemistry (AsR5) and oxidative ring-contraction (AsL4, AsL6) were validated.

Based on the insights obtained from xenovulene A biosynthesis homologous gene clusters were identified in the filamentous fungi *Phaeosphaeriaceae* sp. CF-150626 (the *eup2* BGC) and *Leptobacillium* sp. 236968 (the *pyc* BGC). Heterologous expression of the humulene synthases encoded within both clusters and *in vitro* reconstitution of their bioactivity linked the clusters to the production of eupenifeldin **5**/noreupenifeldin B **6** (*eup2*) and pycnidione **9** (*pyc*). The humulene synthase EupR3 from the *eup2* BGC was shown to catalyse formation of 2*Z*-humulene **7**, a geometric isomer of α-humulene **2**.

With AsR6 and EupR3 a new type of class I terpene cyclase was then characterised on a structural and molecular level. The enzymes were found to use an unusual pyrophosphate binding motif that is unprecedented in this type of enzyme and that consists of a binuclear magnesium cluster and a lysine residue usually not involved in $PP_i$-coordination. Site-directed mutagenesis validated this motif and identified a single amino acid (L285) that controls the stereoselective formation of either α-humulene **2** or its geometric isomer *2Z*-humulene **7**. Possible mechanisms of humulene formation were investigated using different isoprenoid substrates (FPP **3**, (*rac*)-NPP **4** and (3*S*)-NPP **4b**). It was shown that AsR6 can – most unusually – convert both enantiomers of (*rac*)-NPP **4** into α-humulene **2**.

The investigative studies cumulated in the successful engineering of artificial tropolone sesquiterpenoid biosynthetic pathways in *A. oryzae*. Seven new-to-nature metabolites were produced and characterized, differing in the fused polyketide moiety, the degree of hydroxylation and the humulene substitution pattern. The usage of *A. oryzae* as a versatile tool in pathway elucidation and total biosynthesis was thus successfully demonstrated.

Selective labelling experiments with stable isotopes in wild-type *Phaeosphaeriaceae* sp. CF-150626 and in *A. oryzae* expression strains then demonstrated that noreupenifeldin B **6** and the metabolites **293** and **300** (obtained from heterologous expression) are derived from new types of (probably) oxidative ring-contractions. In *A. oryzae* this transformation must be catalysed by a native enzyme, as none of the transformed genes was linked to this kind of chemistry.

Several aspects of TS biosynthesis, however, remain unclear and require further investigation. These include the mechanisms of the terpene cyclase AsR6 (in particular how the enzyme converts both enantiomers of NPP **4**) and of the ring-contraction enzymes AsL4, AsL6 and EupR5. In the case of AsR6, labelling studies using isotopically traced isoprenoid substrates should be considered to confirm the mechanism of humulene formation. In the case of AsL4/AsL6 and EupR5 a focus of future experiments has to be on establishing suitable expression conditions to obtain recombinant enzymes. Fungal expression systems or insect cells are candidate systems that are promising to achieve this aim. Finally, the constructed library of tropolone sesquiterpenoids should be tested for potential bioactivity. The ongoing collaboration with PROF. LEE-THEDIECK has already returned promising initial data on several tropolone sesquiterpenoids involved in cell cycle arrest and induction of erythropoietin gene expression and these studies should be systematically continued.

# 7. Experimental

All chemicals, reagents and solvents were purchased in analytical grade (or higher) from *Sigma Aldrich*, *Carl Roth*, *Thermo Fischer Scientific, Acros Organics, Grüssing GmbH* or *Honeywell*. (*S*)-nerolidol **238b** was obtained from *Sigma Aldrich* and (*rac*)-nerolidol **238** from *ChemScene*. Media, buffers and solutions were prepared with deionised water that was further purified with a Barnstead GenPure Pro System from *Thermo Fischer Scientific* (in the following referred to as $ddH_2O$). The media and solutions used were autoclaved before use at 121 °C for 15 min or sterilised by disposable syringe filters (pore size: 0.2 – 0.45 µm, *Carl Roth*).

## 7.1. Antibiotics, Enzymes, Buffers, Solutions and Growth Media

### 7.1.1. Antibiotics and Enzymes

Antibiotics were prepared as 1000x concentrated stock solutions that were stored at - 20 °C for up to one year. Antibiotics were dissolved in the indicated solvents (Table 7.1) and sterilized by disposable syringe filters. All enzymes used in this study were obtained from *New England Biolabs* or *Thermo Fischer Scientific* and used according to the manufacturer´s instructions.

**Table 7.1** Antibiotic stocks, preparation and working concentrations.

| antibiotic | solvent | stock concentration | working concentration |
|---|---|---|---|
| carbenicillin ($Carb^{50}$) | $ddH_2O$ | 50 $mg \cdot mL^{-1}$ | 50 $\mu g \cdot mL^{-1}$ |
| chloramphenicol ($CA^{30}$) | ethanol | 30 $mg \cdot mL^{-1}$ | 30 $\mu g \cdot mL^{-1}$ |
| kanamycin ($Kan^{50}$) | $ddH_2O$ | 50 $mg \cdot mL^{-1}$ | 50 $\mu g \cdot mL^{-1}$ |

### 7.1.2. Buffers and Solutions

Buffers used in this study were prepared with $ddH_2O$ and sterilized by disposable Fisherbrand™ PES Bottle Top Filters (pore size: 0.2 µm). The pH was adjusted with 2 M HCl or 2 M NaOH and controlled *via* an FiveEasy Standard pH Meter Line (*Mettler Toledo*).

**Table 7.2** Buffers used in this study for protein purification, SDS-PAGE and agarose gel electrophoresis.

| buffer | composition |
|---|---|
| *Protein purification & SDS-PAGE* | |
| lysis buffer | 50 mM Tris-HCl, pH 8, 150 mM NaCl, 10 mM imidazole, 10 % (v/v) glycerole |
| elution buffer | 50 mM Tris-HCl, pH 8, 150 mM NaCl, 500 mM imidazole, 10 % (v/v) glycerole |
| assay buffer | 50 mM Tris-HCl, pH 8, 150 mM NaCl, 1 mM $MgCl_2$, 10 % (v/v) glycerole |
| coomassie dye | 25 % (v/v) acetic acid, 10 % (v/v) isopropanol, 0,1 % (w/v) Coomassie |
| coomassie bleach | 25 % (v/v) acetic acid, 10 % (v/v) isopropanol |

| buffer | composition |
|---|---|
| 4x Lämmli | 10 % (v/v) β-mercaptoethanol, 0.25 % (w/v) bromophenol blue, 30 % glycerol, 0.25 % xylene cyanol |
| 10x SDS running buffer | 25 mM Tris-HCl, 192 mM glycine, 0.1 % (w/v) SDS, pH 8.3 |
| *Agarose gel electrophoresis* | |
| 50x TAE | 2 M Tris acetate, 0.05 M EDTA, pH 8.3 |

**Table 7.3** Solutions used in this study.

| solutions | composition |
|---|---|
| *preparation of media* | |
| trace element solution | 0.1 % (w/v) $FeSO4 \cdot 7H_2O$, 0.057 % (w/v) $MnSO_4 \cdot 4H_2O$, 0.0025 % (w/v) $CuCl_2 \cdot 2H_2O$, 0.01 % (w/v) $CaCl_2 \cdot 2H_2O$, 0.0056 % (w/v) $H_3BO_3$, 0.0019 % (w/v) $(NH_4)_6NaMo_7O_{24} \cdot 4H_2O$, 0.2 % (w/v) $ZnSO4 \cdot 7H_2O$* |
| *fungal transformation* | |
| solution 1 | 0.8 M NaCl, 10 mM $CaCl_2$, 50 mM Tris-HCl, pH 7.5 |
| solution 2 | 60 % PEG 3350, 0.8 M NaCl, 10 mM $CaCl_2$, 50 mM Tris-HCl, pH 7.5 |

*dissolved dropwise in 0.6 M HCl.

### 7.1.3. Growth Media

Growth media used in this study was kindly prepared by the technical staff of the BMWZ and its composition is listed in Table 7.4.

**Table 7.4** Growth media used in this study.

| growth media | composition |
|---|---|
| *E.coli* | |
| LB agar | 0.5 % (w/v) NaCl, 1 % (w/v) tryptone, 0.5 % (w/v) yeast extract, 1.5 % agar |
| LB | 0.5 % (w/v) NaCl, 1 % (w/v) tryptone, 0.5 % (w/v) yeast extract |
| 2TY | 1 % (w/v) yeast extract, 1.6 % (w/v) tryptone, 0.5 % (w/v) NaCl |
| | |
| *S. cerevisiae* | |
| YPAD agar | 1 % (w/v) yeast extract, 2 % (w/v) tryptone, 2 % (w/v) D(+)-glucose monohydrate, 0.03 % (w/v) adenine, 1.5 % (w/v) agar |
| YPAD | 1 % (w/v) yeast extract, 2 % (w/v) tryptone, 2 % (w/v) D(+)-glucose monohydrate, 0.03 % (w/v) adenine |
| SM-URA agar | 0.17 % (w/v) yeast nitrogen base, 0.5 % (w/v) $(NH_4)_2SO_4$, 2 % (w/v) D(+)-glucose monohydrate, 0.077 % complete supplement mixture minus uracil, 1.5 % (w/v) agar |
| *A. oryzae* NSAR1 | |
| CZD/S agar | 3.5 % (w/v) Czapek Dox broth, 18.22 % (w/v) D-sorbitol, 0.1 % (w/v) $(NH_4)_2SO_4$, 0.15 % (w/v) L-methionine, 0.05 % (w/v) adenine, 1.5 % (w/v) agar |
| CZD/S soft agar | 3.5 % (w/v) Czapek Dox broth, 18.22 % (w/v) D-sorbitol, 0.1 % (w/v) $(NH_4)_2SO_4$, 0.15 % (w/v) L-methionine, 0.05 % (w/v) adenine, 0.8 % (w/v) agar |
| CZD/S1 agar | 3.5 % (w/v) Czapek Dox broth, 18.22 % (w/v) D-sorbitol, 0.1 % (w/v) $(NH_4)_2SO_4$, 0.15 % (w/v) L-methionine, 1.5 % (w/v) agar |
| CZD/S1 soft agar | 3.5 % (w/v) Czapek Dox broth, 18.22 % (w/v) D-sorbitol, 0.1 % (w/v) $(NH_4)_2SO_4$, 0.15 % (w/v) L-methionine, 0.8 % (w/v) agar |
| CZD/S1 agar - methionine | 3.5 % (w/v) Czapek Dox broth, 18.22 % (w/v) D-sorbitol, 0.1 % (w/v) $(NH_4)_2SO_4$, 1.5 % (w/v) agar |
| CZS/S1 soft agar - methionine | 3.5 % (w/v) Czapek Dox broth, 18.22 % (w/v) D-sorbitol, 0.1 % (w/v) $(NH_4)_2SO_4$, 0.8 % (w/v) agar |
| DPY agar | 2 % (w/v) dextrin from potato starch, 1 % (w/v) polypeptone, 0.5 % (w/v) yeast extract, 0.5 % (w/v) $KH_2PO_4$, 0.05 % (w/v) $MgSO_4 \cdot H_2O$, 2.5 % (w/v) agar |

| growth media | composition |
|---|---|
| DPY | 2 % (w/v) dextrin from potato starch, 1 % (w/v) polypeptone, 0.5 % (w/v) yeast extract, 0.5 % (w/v) $KH_2PO_4$, 0.05 % (w/v) $MgSO_4 \cdot H_2O$ |
| GN | 2 % (w/v) D(+)-glucose monohydrate, 3 % (w/v) nutrient broth nr.2 from oxoid |
| *Phaeosphaeriaceae sp. CF-150626* | |
| PD agar | 2.4 % (w/v) PDB powder, 1.5 % (w/v) agar |
| TOM medium | 1.0 % (w/v) oat meal*, 4 % (w/v) tomato paste, 0.5 % (w/v) corn steep liquor, 1 % (w/v) D(+)-glucose monohydrate, 1 % (w/v) trace element solution** |
| PM medium | 5 % (w/v) yellow cornmeal, 0.1 % (w/v) yeast extract, 8 % (w/v) sucrose |

*boiled with 2/3 of liquid for 1 h and filtered through a sieve (pore size: 0.5 mm).
**adjust pH to 6.8.

## 7.2. Microbiology Methods

All bacterial strains, yeast and *A.oryzae* NSAR1 were obtained from the BMWZ inhouse strain collection. All other fungal strains were obtained from the Fundación MEDINA (Granada, Spain).

### 7.2.1. Overview Strains

**Table 7.5** Bacterial, yeast and fungal strains used in this work.

| strain | genotype |
|---|---|
| *E. coli* One Shot® BL21(DE3) | $F^-$ *ompT hsdS*$_B$ ($r_B^- m_B^-$) *gal dcm* (DE3) |
| *E. coli* OneShot® Top10 | $F^-$ *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*) 7697 *gal*U *gal*K *rps*L ($Str^R$) *end*A1 *nup*G λ- |
| *E. coli* OneShot ccdB survival 2T1R | $F^-$*mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*Δ139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L ($Str^R$) *end*A1 *nup*G *fhu*A::*IS*2 |
| *Saccharomyces cervisiae* CEN.PK2 | *MAT*a/α *ura*3-52/*ura*3-52 *trp*1-289/*trp*1-289 *leu*2-3_112/leu2-3_112 *his*3 Δ1/*his*3 Δ1 *MAL*2-8C/*MAL*2-8C *SUC*2/*SUC*2 |
| *Aspergillus oryzae* NSAR1 | Δ*argB sC*$^-$ Δ*adeA niaD*$^-$ |
| *Phaeosphaeriaceae* sp. CF-150626 | wild-type |
| *Leptobacillium* sp. CF-236968 | wild-type |

### *7.2.2. Escherichia coli*

*Growth and Maintenance*

*E. coli* strains were cultured on solid LB agar supplemented with an appropriate selection marker at 37 °C for up to 18 h. Solid cultures were stored for short term usage at 4 °C for up to two weeks. Liquid cultures were grown in LB medium supplemented with an appropriate selection marker at 37 °C and 225 rpm for up to 18 h. 25 % (v/v) glycerol stocks were prepared for long term storage and kept indefinitely at - 80 °C.

*Heat-shock Transformation of Chemically Competent E. coli Cells*

50 µL aliquots of chemically competent *E. coli* strains (kindly provided by the BMWZ technical staff) were thawed on ice, mixed with 60 - 100 ng plasmid DNA or alternatively 5 - 10 µL of a

ligation mixture and incubated on ice for 25 min. A heat-shock was then performed at 42 °C for 45 s. After cooling the cells on ice for 90 s 250 µL of LB medium was added to the cells and followed by incubation at 37 °C and 225 rpm for 45 - 60 min. After that the cells were spun down by centrifugation (max. speed), resuspended in 50 µL LB medium and streaked on LB agar plates supplemented with an appropriate selection marker. The plates were incubated at 37 °C for up to 18 h.

### 7.2.3. *Saccharomyces cerevisiae*

*Growth and Maintenance*

*S. cerevisiae* CEN.PK2 was cultured on YPAD agar at 30 °C for 2 - 3 d. For liquid cultures a single colony was inoculated into YPAD medium and cultivated at 30 °C and 160 rpm for 18 h.

*Transformation of S. cerevisiae Cells*

For transformation of *S. cerevisiae* CEN.PK2 a lithium acetate/salmon testis carrier DNA/PEG mediated transformation protocol was used, as described by GIETZ AND WOODS in *Methods in Enzymology*.[266] In short, a 10 mL YPAD seed culture inoculated with a single colony was grown overnight for up to 18 h at 30 °C and 160 rpm. This seed culture was then topped with 40 mL YPAD medium and cultivated for another 4.5 h. After centrifugation (3000 x g, 5 min, 4 °C) the pellet was washed with 25 ml $ddH_2O$ and then again with 1 mL $ddH_2O$. After that water was added to the cells to a final volume of 1 mL. 100 µL aliquots of the resuspended cells were transferred to a clean microcentrifuge tube and centrifuged at max. speed for 1 min before the cell pellet was resuspended in 360 µL transformation mixture. The transformation mixture consisted of 240 µL PEG 3350 (50 % [w/v]), 36 µL LiOAc (1 M), 50 µL carrier DNA (2 $mg \cdot mL^{-1}$ salmon testis DNA) and 34 µL DNA plasmid mix. The DNA plasmid mix comprises *Asc*I digested *pTYGS*-arg/ade/met (600 ng) and PCR-derived insert genes (in equimolar concentrations) that contain a 30 bp 3´ and 5´ overlap with the cut sites of the plasmid. In case that recombination sites were not used for the introduction of gene cassettes small DNA-"patches" were amplified from the intact *pTYGS* plasmid that would result in the native plasmid sequence upon recombination. The resulting transformation mixture was incubated at 42 °C for 50 min, then centrifuged at 6000 x g for 1 min before the cell pellet was resuspended in 1 mL $ddH_2O$. 200 µL were streaked on SM-URA agar plates and incubated at 30 °C until colonies appeared after 2-3 d.

### 7.2.4. *Aspergillus oryzae* NSAR1

*Growth and Maintenance*

*Aspergillus oryzae* NSAR1 was cultured on DPY agar at 28 °C. Typically, after 5 - 7 d sporulation occurred upon which spores were taken up in a minimal amount of water, mixed with glycerol to a final concentration of 50 % (v/v) and stored indefinitely at - 80 °C. For the production of secondary metabolites mycelia/spores from a 7 d old plate was taken up in 4 ml of dd$H_2O$ and inoculated into 500 mL baffled flasks containing 100 ml DPY medium. The cultures were incubated at 28 °C and 110 rpm for 5 - 7 d.

*Transformation of A. oryzae NSAR1*

For transformation purposes mycelia/spores from a 7 d old plate were inoculated into a 250 mL flask containing 50 mL GN medium and incubated at 28 °C and 110 rpm for 20 h. The obtained conidia were collected *via* filtration through sterile miracloth, washed with approximately 50 mL 0.8 M NaCl and taken up in 10 mL protoplasting solution (100 mg Lysing Enzymes from *Trichoderma harzianum* [*Sigma Aldrich*] per 10 mL). Protoplast formation occurred during incubation of this mixture at 30 °C for 3 - 4 h, while being lightly rotated. Afterwards protoplasts were released from the hyphal strands *via* pipetting with a 1000 µL pipette equipped with cut-tips. After filtration through sterile miracloth the filtrate was centrifuged at 3000 x g for 5 min at 4 °C. The supernatant was carefully decanted and the (usually) hardly visible protoplasts were resuspended in 600 µL solution 1. 10 µL of plasmid DNA (>> 1000 ng) were then mixed with 100 µL of the protoplast solution and rested on ice for 2 min. After addition of 1 mL solution 2 and light rotation the mixture was rested at ambient temperature for 20 min. The protoplasts were then taken up in molten CZD/SX soft agar and overlaid on the corresponding CZD/SX agar to be incubated for 3 - 5 d until transformants appeared. The choice of CZD/SX agar depended on the nature of the transformed plasmid(s) and is described in Table 7.6.

**Table 7.6** Choice of CZD/SX agar for transformation of *Aspergillus oryzae* NSAR1.

| agar | # of transformed plasmids | plasmids |
|---|---|---|
| CZD/S | 1 | *pTYGS*-arg |
| CZD/S1 | 2 | *pTYGS*-are; *pTYGS*-ade |
| CZD/S1 - methionine | 3 | *pTYGS*-are; *pTYGS*-ade; *pTYGS*-met |

*Selection of Transformants*

The transformation process was typically repeated until ≥ 12 transformants were obtained. Each transformant was transferred to a fresh secondary selection plate and allowed to grow at 28 °C for 3 - 5 d. This selection was repeated on a tertiary plate before a small number of spores was transferred to DPY agar for regeneration. All transformants were analysed *via* extraction of gDNA

and subsequent PCR analysis for correct integration of all transformed gene cassettes. An overview of the conducted transformation experiments is displayed in Table 7.7.

**Table 7.7** Plasmids transformed in the individual experiments.

| exp. ID | labbook ID | transformed plasmids with *pTYGS*-arg/ade/met backbone |
|---|---|---|
| #Exp. 1 | CS I 112 | RS I 96 1-6, CS I 03 I, CS I 82 7-1 |
| #Exp. 2 | CS I 119 | RS I 96 1-6, CS I 03 I, CS I 106 1-2 |
| #Exp. 3 | CS I 134 | RS I 96 1-6, CS I 03 I, CS I 80 1-1 |
| #Exp. 4 | CS I 136 | RS I 96 1-6, CS I 128 2-1, *pTYGS*-met |
| #Exp. 5 | CS I 114 | RS I 63 2-5, CS I 03 I, CS I 110 1-5 |
| #Exp. 6 | CS I 130 | RS I 63 2-5, CS I 03 I, CS I 125 1-1 |
| #Exp. 7 | CS I 92 | RS I 63 2-5, CS I 75 3-2, *pTYGS*-met |
| #Exp. 8 | CS I 96 | RS I 63 2-5, CS I 93 2-4, *pTYGS*-met |
| #Exp. 9 | CS I 104 | RS I 63 2-5, CS I 93 2-4, CS I 101 1-1 |
| #Exp.10 | CS I 62 | RS I 63 2-5, CS I 58 1-2, *pTYGS*-met |

### 7.2.5. *Phaeosphaeriaceae* sp. CF-150626

*Growth and Maintenance*

*Phaeosphaeriaceae* sp. CF-150626 was cultured on PD agar and at ambient temperature for 14 d. For long term storage mycelia was resuspended in a minimal amount of water, mixed with glycerol to a final concentration of 25 % (v/v) and stored at - 80 °C indefinitely. For production of secondary metabolites a solid plate was cut into small agar disks and inoculated into a 250 mL flask containing 50 mL TOM medium. This seed culture was cultivated at ambient temperature for 3 - 5 d (upon which a visible change of colour occurs from red to brown). 4 mL of this seed culture were inoculated into 500 mL flasks containing 100 mL PM medium. The producing cultures were kept at 22 °C and 220 rpm for 13 d.

## 7.3. Molecular Biology Methods

### 7.3.1. gDNA Extraction

Fungal gDNA of *Aspergillus oryzae* NSAR1 and *Phaeosphaeriaceae* sp. CF-150626 was extracted using the GenElute™ Plant Genomic DNA Miniprep Kit (*Sigma Aldrich*) following the manufacturer´s instructions. Typically, a fungal culture was Büchner filtrated and approximately 100 mg of dried sample were used for the extraction.

For genome sequencing higher quality genomic DNA was needed. Prior to extraction of gDNA the fungal mycelia was freeze-dried overnight. The resulting mycelia was then grinded in liquid nitrogen to obtain a very fine powder that was used in multiple parallel extractions as described above. Combined gDNA samples were further purified using a sodium acetate-based purification protocol: To 100 µL total DNA volume 33 µL 3 M sodium acetate (pH 5.5) was added, followed

by addition of 2 volumes of 100% EtOH. The resulting mixture was incubated at - 20 °C for 30 min prior to centrifugation at 12000 x g for 10 min at 4 °C. The supernatant was carefully decanted, and the pellet was vortexed in 200 µL of 70% EtOH (- 20 °C). After centrifugation (12000 x g, > 3 min, 4 °C) the supernatant was pipetted off and the pellet was allowed to dry on air for at least 5 min. Finally, the DNA pellet was resuspended in nuclease free water and stored at - 20 °C.

### 7.3.2. RNA Extraction and cDNA Preparation

Total RNA was isolated from fungal mycelia (typically 100 mg) using the TRIzol® Reagent (*Thermo Fischer Scientific*) according to the manufacturer´s instructions. Total RNA was frozen in liquid nitrogen and stored at - 80 °C or directly converted to cDNA.

RNA was transcribed into cDNA using the High Capacity RNA-to-cDNA Kit (*Thermo Fischer Scientific*). The manufacturer´s instructions were followed.

### 7.3.3. Oligonucleotides

All oligonucleotides used in this study were obtained from *Sigma Aldrich* as concentrated stock solutions. Stocks were diluted in dd$H_2O$ [1:10] and stored at - 20 °C indefinitely.

**Table 7.8** Oligonucleotides used in this study.

| #labbook ID | 5´-3´sequence |
|---|---|
| 83 | ATGGCAGCTCATGGGCAAAC |
| 87 | CTTCTTAAATATCGTTGTAACTGTTCCTGA |
| 88 | CGAAGTATATTGGGAGACTATAGCTACTAG |
| 89 | ATTCACCACTATTATTCCCACCCTATAATA |
| 90 | GAGACGAAACAGACTTTTTCATCGCTAAAA |
| 91 | CTTTTCTTTTCTCTTTCTTTTCCCATCTTC |
| 92 | TGACCTCCTAAAACCCCAGTG |
| 322 | CTGCTGGCTTAACACGTGC |
| 421 | TTCTTTCAACACAAGATCCCAAAGTCAAAGATGGACAGCCCAGAAGTAT |
| 423 | TTTCATTCTATGCGTTATGAACATGTTCCCTTAAAGAGTATAGCCGCC |
| 424 | ACAGCTACCCCGCTTGAGCAGACATCACCGATGGGCAGCCTCACTGAT |
| 426 | TACGACAATGTCCATATCATCAATCATGACCTATGGTAGCACTACTGGC |
| 427 | CGACTGACCAATTCCGCAGCTCGTCAAAGGATGGCTCTCGCACAGCAA |
| 428 | CAGGTTGGCTGGTAGACGTCATATAATCATTCATTCTTTGTCCGAGCG |
| 645 | TTTCTTTCAACACAAGATCCCAAAGTCAAAATGCCGCAACTAAAGGTTCT |
| 650 | GGTTGGCTGGTAGACGTCATATAATCATACTCATGCCTCAAACTCCAGCT |
| 706 | TTTCATTCTATGCGTTATGAACATGTTCCCTCACTCCTTGAGAAGCTCTG |
| 709 | CGACTGACCAATTCCGCAGCTCGTCAAAGGATGACTGTGAAGATCCTTGT |
| 759 | GTCGACTGACCAATTCCGCAGCTCGTCAAAATGCCCGTTACTACCCCCAC |
| 760 | GGTTGGCTGGTAGACGTCATATAATCATACTTACCCAACAGCAGTTGTTA |
| 761 | GCCAACTTTGTACAAAAAAGCAGGCTCCGCATGCGTCGCAGTTTTCTTAT |
| 762 | TGCCAACTTTGTACAAGAAAGCTGGGTCGGCTAGAAGTGAAAGCCAGTCG |
| 892 | CGGAATTCATGCGTCGTAGCTTTCTGATT |
| 894 | ATAGTTTAGCGGCCGCTTAAAAGTGAAAACCCGTGG |

| #labbook ID | 5´-3´sequence |
|---|---|
| 1339 | TTCTTTCAACACAAGATCCCAAAGTCAAAATGAAGCTCCTCGCTACCGC |
| 1340 | TTCATTCTATGCGTTATGAACATGTTCCCTTCAGAAGTGAAAGCCAGTAG |
| 1341 | TAACAGCTACCCCGCTTGAGCAGACATCACATGACGGAAACCAGCCTGAA |
| 1342 | ACGACAATGTCCATATCATCAATCATGACCTCAAGAGACAACGGTAGCAT |
| 1440 | AACAGCTACCCCGCTTGAGCAGACATCACCATGCACAAGAACACTGCAAC |
| 1441 | ACGACAATGTCCATATCATCAATCATGACCCTACATGGTAAAGCCAGTAG |
| 1442 | GTCGACTGACCAATTCCGCAGCTCGTCAAAATGAAAGAAATGTGGGAAAC |
| 1443 | GGTTGGCTGGTAGACGTCATATAATCATACCTAAGGCGCTACAGTAATCG |
| 1450 | TTTCTTTCAACACAAGATCCCAAAGTCAAA ATGACGTCTCTCAGAGTTCT |
| 1451 | TTCATTCTATGCGTTATGAACATGTTCCCT TCACAACTCGTCATGAGTTG |
| 1455 | AGGCTTGCTATGGCGGGGTG |
| 1456 | CGGCATCGAACACCCCGCCATAGCAAGCCTCCACGACGCCATGTCAATCG |
| 1457R | CTCCCGGCACTTATTAACGT |
| 1457F | AGATGCACGAACGTTAATAAGTGCCGGGAGTGACACCGTCGCCATAACAT |
| 1458F | TAATGCCAACTTTGTACAAAAAAGCAGGCTATGTCATGGGGCTCGACAGC |
| 1459 | TATAATGCCAACTTTGTACAAGAAAGCTGGCTACTCCCCTCTCGGCCTAA |
| 1500 | TTTCTTTCAACACAAGATCCCAAAGTCAAAATGACGTCTTTTGGAACCAA |
| 1501 | TTCATTCTATGCGTTATGAACATGTTCCCTCTACCAGGGTTGCACCATAG |
| CSI001 | CTGAGAATCTTTATTTTCAGGGCGCCATGCCGGTTACCACAC |
| CSI002 | CAGTGGTGGTGGTGGTGGTGTTAACCAACTGCGGTGG |
| CSI011 $F_L$ | CAAAAGAAGCAATGGGCGTTC |
| CSI012 $F_S$ | CAAAAGAAGCAATGGGCGTTC |
| CSI013 $R_L$ | ACTGTGCAATGCTTGCC |
| CSI014 $R_S$ | ACTGTGCAATGCTTGCC |
| CSI017 $F_L$ | GTCTGCATTTTGTTCCGATGATGGATCGCTATCTGGAAC |
| CSI018 $F_S$ | TGATGGATCGCTATCTGGAAC |
| CSI019 $R_L$ | TCGGAACAAAATGCAGACCTGCGCTACCATAACGACG |
| CSI020 $R_S$ | CTGCGCTACCATAACGACG |
| CSI021 $F_L$ | ATATGGCAAGCGAAGCAATGGGTATTCTGCAGGGTGAAC |
| CSI022 $F_S$ | GGGTATTCTGCAGGGTGAAC |
| CSI023 $R_L$ | TGCTTCGCTTGCCATATCCAGGGTAACACACTGTGC |
| CSI024 $R_S$ | CCAGGGTAACACACTGTGC |

### 7.3.4. Polymerase Chain Reaction

Polymerase chain reaction (PCR) was performed to amplify DNA fragments. Two types of thermostable polymerase were used for this purpose. For analytical purposes (*e.g.,* colony PCR, assessment of integration of DNA cassettes into *A. oryzae* NSAR1 gDNA, *etc*.) One*Taq*® 2X Master Mix with Standard Buffer (*New England Biolabs*) was used. In case the amplified DNA fragment was used in downstream applications (*e.g.,* restriction-ligation based cloning, yeast homologous recombination, *etc*.) the Q5® High-Fidelity 2X Master Mix (*New England Biolabs*) was used. The manufacturer´s protocol was followed. Annealing temperatures were usually 58 °C for *One*Tag® and 60 °C for Q5® polymerase (alternatively primer specific annealing temperatures were calculated using the *New England Biolabs* TM calculator; https://tmcalculator.neb.com).

### 7.3.5. Obtained and Constructed Vectors

A series of fungal and bacterial expression vectors were constructed in the course of this thesis. Details on cloning procedures (*in vitro* cloning and *in vivo* homologous recombination) are specified below.

**Table 7.9** Overview of obtained and purchased plasmids.

| labbook ID | vector backbone | gene insert | origin |
|---|---|---|---|
| RS I 96 1-6 | *pTYGS*-arg | *asPKS1*, *asL1, asL3, asR2* | RAISSA SCHOR[123] |
| RS I 63 2-5 | *pTYGS*-arg | *asPKS1, asL1, asL3* | RAISSA SCHOR[123] |
| RS I 97 1-1 | *pTYGS*-met | *asL4, asL6* | RAISSA SCHOR[123] |
| RS I 112 2-3 | *pTYGS*-ade | *asR6* | RAISSA SCHOR[123] |
| RS I 100 3-1 | pE-YA | *asR5* | RAISSA SCHOR[123] |
| - | pETM-11 | - | DR. SIMONE HÖFLER |
| - | pET100/D-TOPO | *N-His$_6$-asR6* | *Thermo Fischer* |
| - | pET100/D-TOPO | *N-His$_6$-asR5* | *Thermo Fischer* |
| - | pET100/D-TOPO | *N-His$_6$-pycR6* | *Thermo Fischer* |
| - | pET100/D-TOPO | *N-His$_6$-eupR3* | *Thermo Fischer* |
| - | pET100/D-TOPO | *N-His$_6$-asR6$_{L285M}$* | *Thermo Fischer* |
| - | pET100/D-TOPO | *N-His$_6$-asR6$_{L285A}$* | *Thermo Fischer* |
| - | pET100/D-TOPO | *N-His$_6$-asR6$_{L285V}$* | *Thermo Fischer* |
| - | pET100/D-TOPO | *N-His$_6$-asR6$_{L285N_K289S}$* | *Thermo Fischer* |
| - | pET100/D-TOPO | *N-His$_6$-asR6$_{Y351A}$* | *Thermo Fischer* |

**Table 7.10** Oligonucleotides (ONC) used for construction of bacterial expression vectors.

| labbook ID | name | vector backbone | ONC |
|---|---|---|---|
| CS I 01 J | pET32_asR5 | pET32 | 892 + 894* |
| CS I 135 1-7 | pETM-11_asR6 | pETM-11 | CSI001 + CSI002 |
| CS I 158 1-2 | pET100/D-TOPO_eupR3$_{M261L}$ | pET100/D-TOPO | CSI011 $F_L$ + CSI012 $F_S$<br>CSI013 $R_L$ + CSI014 $R_S$ |
| CS I 161 1-2 | pET100/D-TOPO_asR6$_{Y344F}$ | pET100/D-TOPO | CSI017 $F_L$ + CSI018 $F_S$<br>CSI019 $R_L$ + CSI020 $R_S$ |
| CS I 162 1-3 | pET100/D-TOPO_asR6$_{K289S}$ | pET100/D-TOPO | CSI021 $F_L$ + CSI022 $F_S$<br>CSI023 $R_L$ + CSI024 $R_S$ |

*Cloning strategy devised by RAISSA SCHOR. Experiments partly conducted together.

**Table 7.11** Oligonucleotides (ONC) used for construction of fungal expression vectors; gateway cloning was deployed as indicated in several examples.

| labbook ID | vector backbone | ONC *S. cerevisiae* | gateway® |
|---|---|---|---|
| CS I 03 I | *pTYGS*-ade | - | entry vector: RS I 100 3-1 (asR5)<br>destin. vector: RS 112 2-3 (asR6) |
| CS I 58 1-2 | *pTYGS*-ade | *pycR1*: 1339 + 1340<br>*pycR6*: 1341 + 1342<br>patch_eno: 87 + 88 | - |
| CS I 75 3-2 | *pTYGS*-ade | *eupR1*: 1440 + 1441<br>*eupR3*: 1442 + 1443<br>patch_padh: 89 + 90 | - |
| CS I 80 1-1 | *pTYGS*-met | *eupR5*: 1450 + 1451<br>patch_gdpd: 91 + 92<br>patch_eno: 87 + 88 | - |
| CS I 82 6-5 | pE-YA | *eupR6*: 1458F + 1455<br>+ 1456+ 1457R<br>+ 1457F + 1459 | - |
| CS I 82 7-1 | *pTYGS*-met | - | Entry vector: CS I 82 6-5<br>Destin. vector: pTYGS-met |
| CS I 93 2-4 | *pTYGS*-ade | *eupR1*: 1440 + 1441<br>*eupR3*: 1442 + 1443<br>*eupL4*: 1500 + 1501 | - |
| CS I 101 1-1 | *pTYGS*-met | - | Entry vector: CS I 82 6-5<br>Destin. vector: CS I 80 1-1 |

| | | | |
|---|---|---|---|
| CS I 106 1-2 | *pTYGS*-met | - | Entry vector: CS I 82 6-5<br>Destin. vector: RS I 97 1-1 |
| CS I 110 1-5 | *pTYGS*-met | *eupL4*: 1500 + 1501<br>patch_eno: 87 + 88<br>patch_gpdA: 91 + 92 | - |
| CS I 125 1-1 | *pTYGS*-met | - | Entry vector: CS I 82 6-5<br>Destin. vector: CS I 110 1-5 |
| CS I 128 1-10 | *pTYGS*-ade | *eupR3*: 1442 + 1443<br>patch_gdpd: 91 +92<br>patch_gadh: 89 + 90 | |
| CS I 128 2-1 | *pTYGS*-ade | - | Entry vector: RS I 100 3-1<br>Destin. vector: CS I 128 1-10 |

### 7.3.6. Agarose Gel Electrophoresis

Agarose gel electrophoresis was performed with a horizontal gel electrophoresis system from *Bio-Rad*. 1 % (w/v) agarose gels containing ROTI®-Safe GelStain (*Carl Roth*) were cast and overlayed with 0.5x TAE-buffer. Appropriate DNA samples (*e.g.,* DNA fragments, products of restriction digests or plasmid DNA) were mixed with 6x DNA loading dye (*New England Biolabs*) and loaded into the sample pockets of cast gels. Electrophoresis occurred at 120 V for 25 - 30 min. The 1 kb DNA ladder (*New England Biolabs*) was used as a reference to determine the size of DNA fragments. Visualization and documentation was achieved with the Molecular Imager® Gel Doc™ XR+ (*Bio-Rad*).

### 7.3.7. DNA Purification

PCR products were further purified using the NucleoSpin Gel and PCR Clean-up, Mini kit for gel extraction and PCR clean up (*Macherey-Nagel*). The manufacturer´s instructions were followed.

### 7.3.8. Determination of DNA Concentration

The concentration of DNA samples was determined with a DeNovix® DS-11+ Spectrophotometer.

### 7.3.9. Isolation of Plasmid DNA

Plasmid DNA from bacterial host cells was prepared using the NucleoSpin® Plasmid Mini Kit (*Macherey-Nagel*). The manufacturer´s instructions were followed. Plasmid DNA from *S. cerevisiae* cells was prepared using the Zymoprep™ Yeast Plasmid Miniprep II kit (*Zymo Research*), according to the manufacturer´s instructions.

### 7.3.10. Sequencing of DNA

Sequencing of DNA was performed by *Eurofins Genomics* (Mix2Seq OVERNIGHT).

### 7.3.11. Cloning Procedures

*Restriction Enzyme Digestion*

All restriction enzymes were used according to the manufacturer´s instructions. DNA sample, enzyme and buffers were thoroughly mixed and incubated at 37 °C for up to 12 h. After this heat inactivation was conducted at 60 °C or 80 °C for 20 min. The digested DNA fragments were purified as described in section 7.3.7.

*Dephosphorylation and Ligation*

For *in vitro* cloning digested plasmids were dephosphorylated with shrimp alkaline phosphatase (SAP). For this 1 µL of SAP-enzyme was added to the reaction mixture after the digestion was completed. After incubation at 37 °C for 1 h heat inactivation was achieved according to the manufacturer´s instructions. DNA ligation was performed with T4 DNA ligase (*New England Biolabs*) using a molar ration of 1:3 (plasmid:insert). All ligase reactions were performed overnight at 4 °C, followed by heat inactivation (65 °C) and transformation into *E. coli* TOP10.

*Gateway® Technology*

For recombinatorial cloning the Gateway LR Clonase enzyme mix II kit (*Invitrogen*) was used according to the manufacturer´s instructions.

*Site-Directed Mutagenesis*

Point mutants of selected terpene cyclases were either obtained as synthetic genes (*Thermo Fischer Scientific*) or the mutations were introduced manually. For manual introduction of point mutations the site-directed, ligase-independent mutagenesis (SLIM) protocol from TILLETT *et al*. was deployed.[224] In brief, a single PCR reaction with tailed primers containing the desired mutation in complementary overhangs at the respective terminus was used. SLIM PCR amplification was performed according to the reported protocol. After PCR 1 µL of *Dpn*I was added to the reaction mixture and incubated at 37 °C for 3 h. SLIM hybridization was then performed according to the reported protocol, but the PCR product was used without buffer exchange. 20 µL of the hybridization mixture were used for transformation of competent *E. coli* TOP10 cells.

## 7.4. Biochemistry Methods

### 7.4.1. Production of Recombinant Protein

For the heterologous expression of fungal genes in *E. coli* the *E. coli* BL21 (DE3) strain was used as an expression host. *E. coli* BL21 (DE3) is a common laboratory expression strain and deletion of the *lon* and *ompT* proteases makes is suitable for the production of high-level recombinant protein (recombinant protein can account for up to 50 % of total biomass).[267,268] Genes of interest were typically test-expressed on an analytical scale to find suitable expression conditions. For this 250 mL baffled flasks containing 50 mL medium (either 2TY or LB) were inoculated with 1 mL of an overnight bacterial seed culture. The culture was then cultivated at 37 °C and 225 rpm until the optical density $OD_{600}$ reached values between 0.4 and 0.6. Protein production was induced by addition of IPTG to a final concentration of 0.1 - 1.0 mM and the cultures were kept at 10 - 25 °C for 16 - 20 h. After cultivation the cells were collected by centrifugation (4000 x g, 20 min, 4 °C). The obtained pellet was resuspended in 15 mL $L^{-1}$ loading buffer, shock-frozen in liquid nitrogen and stored at - 20 °C until further usage.

For the production of recombinant protein on a preparative scale (1 - 10 L) 2 L flasks containing 1 L expression medium were inoculated with 5 mL of bacterial overnight seed culture and treated as described above.

Putative DIELS-ALDERase encoding gene *asR5* was expressed in *E. coli* arctic express (DE3) cells. The manufacturer´s (*Agilent Technologies*) instructions were followed.

### 7.4.2. Cell Lysis

Cell lysis was achieved by sonification (*Bandelin* SONOPULS) using the following standard settings: amplitude = 32 - 35 %, 2x 6 min, pulse rate 10 s on/10 s off. Cell debris was removed by centrifugation (15000 x g, 45 min, 4 °C) and the supernatant (total lysate) was carefully removed using a pipette.

### 7.4.3. Immobilized Metal Affinity Chromatography

All expression constructs used in this study were designed to encode for an *N*- or *C*-terminal polyhistidine tag. Isolation of $His_6$-tagged proteins was accomplished by $Ni^{2+}$ nitrilotriacetic (NTA)-affinity chromatography.

*Gravity-Flow Protocol for $Ni^{2+}$ NTA*

For small-scale purification of recombinant enzymes, the total lysate was mixed with 1.5 mL *PureCube* Ni-NTA agarose and incubated at 4 °C for 1 h under light rotation. The mixture was then applied to a CHROMABOND® empty column with PE filter element (15 mL; *Macherey-Nagel*). Manual pressure was used to increase the flow rate. After the lysate had passed through the filter element and the $Ni^{2+}$-bed had settled 15 mL lysis buffer were passed through to wash. After this the protein of interest was directly eluted with 5 mL elution buffer.

*FPLC Protocol for $Ni^{2+}$ NTA*

For large-scale purification of recombinant enzymes, a fast protein liquid chromatography (FPLC) protocol was used. An Äkta Pure FPLC system (*GE Healthcare*) connected to a HisTrap™ column (*GE Healthcare*; column volume [CV] = 5 mL) was used with the following settings: **A,** equilibration (lysis buffer; 5 CV; flow rate: 5 mL $min^{-1}$); **B,** sample application (flow rate: 1 mL $min^{-1}$); **C,** wash step A (lysis buffer; 5 CV; flow rate: 1.5 mL $min^{-1}$); **D,** wash step B (lysis buffer; 5 CV; flow rate: 5 mL $min^{-1}$); **E,** elution (lysis buffer and elution buffer; the concentration of elution buffer was raised to 100% across 8 CV; flow rate: 2.5 mL $min^{-1}$). A UV detector ($\lambda$ = 280 nm) was used to monitor the purification process.

Concentrated protein samples were obtained using an Amicon Ultra-15 centrifugal unit (*Millipore*; cut-off: 30 kDa). Repeated concentration/dilution was used for buffer exchange.

*Cleavage of Affinity Tags*

While fusion tags are essential tools for the purification of proteins of interest the tags themselves can influence or even impede activity or crystallisation.[269] For the crystallisation of AsR6 the $His_6$-Tag was cleaved off by the tobacco etch virus (TEV) nuclear inclusion proteinase (TEV-protease), that specifically recognizes the amino acid sequence ENLYFQ´(G/S) and cleaves between the Q and G/S residue.[270] TEV-protease was kindly provided by DR. SIMONE HÖFLER (CARLOMAGNO research group). Cleavage was performed overnight in 15 mL final volume (assay buffer; molar ratio of $His_6$-AsR6:TEV [3:1]) under slight rotation at 4 °C.

*Purification of Untagged Proteins*

Reverse $Ni^{2+}$ affinity chromatography was performed to separate mixtures of cleaved and uncleaved $His_6$-AsR6. For this, the same FPLC-program was used as for conventional $Ni^{2+}$ affinity chromatography but with assay buffer instead of lysis buffer. As the tag-less AsR6 is incapable of binding to the $Ni^{2+}$-matrix it elutes in the flow-through. The flow-through was collected and assessed by SDS-PAGE. Fractions containing the protein of interest were pooled, concentrated, shock-frozen in liquid nitrogen and stored at - 80 °C until crystallization.

### 7.4.4. Size-Exclusion Chromatography

Size-exclusion chromatography (SEC) is used to separate protein mixtures by their hydrodynamic volume. A HiLoad 26/600 Superdex 200 prep grade (pg) column (*GE Healthcare*) connected to an Äkta Pure FPLC system (*GE Healthcare*) was deployed for SEC. The deployed column uses a dextran matrix and highly crosslinked agarose to separate protein mixtures, as proteins with a small hydrodynamic volume prevail longer on the porous column matrix than those with a larger volume. For size-exclusion the manually packed column (storage in 20% EtOH) was washed with 1.2 CV $ddH_2O$ and then equilibrated with 1.2 CV of assay buffer. A typical protein sample was concentrated to a final volume of 1 mL and injected upon the column at 0.2 mL $min^{-1}$. After injection the flow-rate was adjusted to 1 mL $min^{-1}$ and elution occurred across 1.2 CV. Elution of proteins occurred at ambient temperatures and proteins of interest were monitored *via* the characteristic UV absorption at $\lambda$ = 280 nm.

### 7.4.5. SDS-Polyacrylamide Gel Electrophoresis

The purification of proteins was analyzed with 12% polyacrylamide gels (Table 7.12). A gel electrophoresis and gel casting system from *Bio-Rad* was used. Samples were prepared by mixing 10 µL 4x Lämmli buffer with 30 µL of protein solution before boiling the mixture at 95 °C for 5 - 10 minutes. Depending on the protein concentration of the sample 3 - 10 µL were loaded onto the gels. The Colour Prestained Protein Standard (Broad Range, 11 - 245 kDa; *New England Biolabs*) was used as a reference. Protein electrophoresis occurred at 80 mA for 45 - 50 min. For staining the gel was put into 20 mL of coomassie staining solution, briefly heated in the microwave and incubated for 10 min on a horizontal shaker. Destaining was achieved using coomassie bleach solution and fresh 20 mL aliquots were added every 20 - 30 min. Analysis and documentation was performed with the Molecular Imager Gel Doc XR+ system (*Bio-Rad*).

**Table 7.12** Composition of 12 % SDS-PAGE gels.

| **separating gel** Volume / mL | **stacking gel** Volume / mL | **composition** |
|---|---|---|
| 3.00 | 0.54 | 30 % acrylamide/bisacrylamide (Rotiphorese® Gel 30 [37,5:1]) |
| 2.45 | 1.70 | $ddH_2O$ |
| 1.90 | - | 1.5 M Tris-HCl, pH 8.8 |
| - | 0.25 | 0.5 M Tris-HCl, pH 6.8 |
| 0.075 | 0.002 | 10 % (w/v) SDS |
| 0.075 | 0.002 | 10 % (w/v) APS |
| 0.0003 | 0.0002 | TEMED |

### 7.4.6. Protein Identification (ESI Q-TOF)

Mass-spectrometry (ESI Q-TOF)-based identification of purified proteins was performed by DR. JENNIFER SENKLER (BRAUN research group, Institute of Plant Genetics, Leibniz University Hannover). Proteins of interest were analyzed by SDS-PAGE and the corresponding band was cut from the gel, dried and submitted for analysis.

### 7.4.7. Determination of Protein Concentration

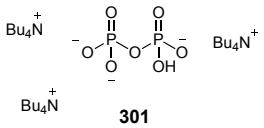
The concentration of protein samples was determined with a DeNovix® DS-11+ Spectrophotometer. The Expasy ProtParam tool (https://web.expasy.org/protparam/) was used to calculate protein molecular weight and extinction coefficients.

### 7.4.8. Enzyme Activity Assays with Terpene Cyclases

Enzymatic activity of terpene cyclases was determined on a 500 µL scale. In a 1.5 mL microcentrifuge tube the respective enzyme (0.1 - 1.0 mg $mL^{-1}$) was incubated with the respective diphosphate (150 µM) and magnesium chloride (5 mM) in assay buffer (pH 8.0). The assay was incubated at 28 °C for 30 - 45 min. The enzymatic assays were then extracted with 300 µL of GCMS-grade n-pentane/n-hexane and directly analysed by GCMS.

## 7.5. Chemical Synthesis

### 7.5.1. Synthesis of pyrophosphate 301




**301**

**Figure 7.1** Chemical structure of tris(tetrabutylammonium)hydrogenpyrophosphate **301**.

Tris(tetrabutylammonium)hydrogenpyrophosphate **301** was synthesized according to a protocol by KIRSCHNING *et al.*[153] Disodium-dihydrogenpyrophosphate (3.13 g, 14 mmol) was dissolved in 25 ml deionized water containing 1 mL concentrated ammonia (25 % in $H_2O$) and eluted with water through a cartridge containing ion exchange resin (Dowex AG 50WX8, $H^+$-form). The solution was then titrated with tetra-*n*-butylammoniumhydroxide to pH 7.3. Lyophilization afforded the title compound as a flocculant colourless solid material (yield not determined).

**$^1$H NMR** ($CDCl_3$, 400 MHz): $\delta$ = 3.15 - 3.26 (m, 24H), 1.60 - 1.72 (m, 24H), 1.37 (sex, $J$ = 7.4 Hz, 24H), 0.96 (t, $J$ = 7.4 Hz, 36H) ppm; **$^{31}$P NMR** ($CDCl_3$, 162 MHz): $\delta$ = -7.6 (s, 2P) ppm. See Appendix 8.3 for NMR spectra.

## 7.5.2. Synthesis of FPP 3

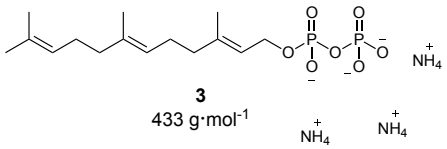



**Figure 7.2** Chemical structure of FPP **3**.

FPP **3** was synthesized according to a protocol by KIRSCHNING *et al*.[153] The reaction was carried out under $N_2$ atmosphere. To tris(tetrabutylammonium)hydrogen pyrophosphate **301** (1.0 g, 1.05 mmol, 2.5 eq.) in acetonitrile (5.0 mL) was added dropwise a solution of *trans, trans*-farnesylbromide (120 mg, 0.42 mmol, 1.0 eq.) in acetonitrile (3.5 mL). The reaction mixture was stirred for 19 h at ambient temperature. After this the solvent was removed under reduced pressure and the residual residue was taken up in 5 mL buffer (0.25 mM $NH_4HCO_2$ in 2 % (*v/v*) isopropylalcohol). This solution was eluted with buffer through a cartridge containing ion exchange resin (Dowex AG 50WX8, $NH_4^+$-form). After lyophilization the crude residue was dissolved in a minimal amount of buffer (50 mM $NH_4HCO_2$), mixed with 20 mL acetonitrile:isopropylalcohol (1:1) and vortexed. After centrifugation (2000 x g, 5 min) the supernatant was taken off and the extraction was repeated three times. Combined supernatants were concentrated to 5 mL. Lyophilization afforded FPP **3** as a flocculent white powder (150 mg, 34 mmol, 81%).

**$^1$H NMR** ($D_2O$, 400 MHz): $\delta$ = 5.48 (t, $J$ = 7.0 Hz, 1H), 5.16 - 5.28 (m, 2H), 4.50 (t, $J$ = 6.5 Hz, 2H), 1.93 - 2.24 (m, 8H), 1.74 (s, 3H), 1.71 (s, 3H), 1.64 (s, 6H) ppm; **$^{31}$P NMR** ($D_2O$, 162 MHz): $\delta$ = -8.30 (d, $J$ = 20.9 Hz, 1P), -10.19 (d, $J$ = 20.9 Hz, 1P) ppm. See Appendix 8.3 for NMR spectra.

## 7.5.3. Synthesis of (*S*)-NPP 4b

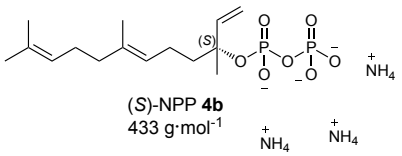



**Figure 7.3** Chemical structure of (*S*)-NPP **4b**.

(*S*)-NPP **4b** was derived from the direct phosphorylation of (*S*)-nerolidol according to a known protocol.[49,214] The reaction was carried out under $N_2$-atmosphere. To a stirring solution of (*S*)-nerolidol (15 mg, 0.07 mmol; $[\alpha]_D^{22} = +9.89$ [$c = 0.438$ g 100 mL$^{-1}$, $CHCl_3$]) in $CCl_3CN$ (0.6 mL) *bis*-triethylammonium phosphate (TEAP) solution (0.15 mL; prepared by the slow addition of 1.82 mL solution 1 [2.5 mL orthophosphoric acid, 9.4 mL acetonitrile] to 3 mL of solution 2 [11 mL triethylamine, 10 mL acetonitrile]) was added. After five and 10 minutes TEAP (0.15 mL) was added again before the reaction mixture was directly chromatographed on silica gel ($^i$PrOH:25 % $NH_3$:$H_2O$ [6:2.5:1]). Formation of the mono- and diphosphate was observed by thin-layer chromatography (TLC; stained with a solution of phosphomolybdic acid in EtOH (10 g·L$^{-1}$). Lyophilization afforded (*S*)-nerolidyl pyrophosphate **4b** (4.6 mg, 0.01 mmol, 15 %) as a light brown solid.

**$^1$H NMR** (DMSO-$d_6$, 400 MHz): $\delta$ = 6.00 (dd, $^3J_{H,H}$ = 17.4 Hz, $^3J_{H,H}$ = 10.8 Hz, 1H), 5.05 - 5.12 (m, 3H), 4.94 (d, $^3J_{H,H}$ = 12.4 Hz, 1H), 1.87 - 2.06 (m, 6H), 1.73 - 1.82 (m, 1H), 1.63 (s, 3H), 1.50 - 1.60 (m, 1H), 1.56 (s, 3H), 1.54 (s, 3H), 1.42 (s, 3H) ppm; **$^{31}$P NMR** (DMSO-$d_6$, 162 MHz): $\delta$ = -9.4 (d, $^2J_{P,P}$ = 12.2 Hz, 1P), -11.7 (d, $^2J_{P,P}$ = 12.2 Hz, 1P) ppm. See Appendix 8.3 for NMR spectra.

## 7.5.4. Synthesis of (*rac*)-NPP 4

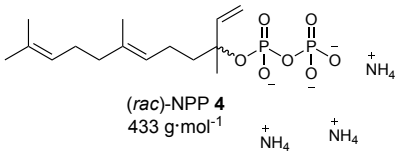



**Figure 7.4** Chemical structure of (*rac*)-NPP **4**.

Under the same conditions as for **4b** (*rac*)-nerolidol (15 mg, 0.07 mmol) was converted to (*rac*)-nerolidyl pyrophosphate **4** (5.0 mg, 0.01 mmol, 16 %).

**$^1$H NMR** (DMSO-$d_6$, 400 MHz): $\delta$ = 6.00 (m, 1H), 5.01 - 5.18 (m, 3H), 4.95 (d, $^3J_{H,H}$ = 10.8 Hz, 1H), 1.84 - 2.07 (m, 6H), 1.73 - 1.83 (m, 1H), 1.63 (s, 3H), 1.50 - 1.60 (m, 1H), 1.56 (s, 3H), 1.54 (s, 3H), 1.42 (s, 3H) ppm; **$^{31}$P NMR** (DMSO-$d_6$, 162 MHz): $\delta$ = -8.8 (m, 1P), -11.6 (m, 1P) ppm. See Appendix 8.3 for NMR spectra.

## 7.6. Chemical Analysis

### 7.6.1. Liquid Chromatography Mass Spectrometry

*Analytical LCMS*

Analytical LCMS data was obtained on either of the following machines:

A Waters 2795HT HPLC (equipped with **i,** a Phenomenex Kinetex column [2.6 µ; $C_{18}$; 100 Å; 4.6 x 100 mm] and **ii,** a Phenomenex Security Guard precolumn [Luna $C_5$; 300 Å] using a flowrate of 1 mL $min^{-1}$). A Waters 996 Diode Array Detector (spanning 210 – 600 nm) plus Waters ZQ mass detector (spanning *m/z* 100 – 1000; operating in $ES^+$ and $ES^-$ mode) was used for analysis.

A Waters 2767 sample manager connected to Waters 2545 pumps and SFO (equipped with **i,** a Phenomenex Kinetex column [2.6 µ; $C_{18}$; 100 Å; 4.6 x 100 mm] and **ii,** a Phenomenex Security Guard precolumn [Luna $C_5$; 300 Å] using a flowrate of 1 mL $min^{-1}$). A Waters 2998 Diode Array Detector (spanning 210 – 600 nm) plus Waters 2424 ELSD and SQD-2 mass detector (spanning *m/z* 100 – 1000; operating in $ES^+$ and $ES^-$ mode) was used for analysis.

Both machines were run with HPLC-grade solvents; **A,** $H_2O$ (with 0.05% formic acid) and **B,** acetonitrile (with 0.045% formic acid). Typical settings: 15 min gradient starting at 10% **B**; ramping to 90% **B** until 10 min; 90% **B** for 2 min; three minutes at 10% **B**.

*Preparative LCMS and Compound Isolation*

Preparative isolation of tropolone sesquiterpenoids and related compounds was achieved using a Waters mass-directed autopurification system (equipped with **i)** a Phenomenex Kinetex Axia column [5µ; $C_{18}$; 100 Å; 21.2 x 250 mm] and **ii)** a Phenomenex Security Guard column [Luna $C_5$; 300 Å]), a Waters 2767 autosampler, a Waters 2454 pump system using a flowrate of 1 mL $min^{-1}$. The flow was split after the column (100 : 1); the minority flow was compensated with [acetonitrile:water (1:1)] at 0.8 mL $min^{-1}$. Analysis of the minority flow was achieved using a Waters 2998 Diode Array Detector, a Waters 2424 ELSD and SQD-2 mass detector (spanning *m/z* 100 – 1000; operating in $ES^+$ and $ES^-$ mode).

Preparative LCMS was run with HPLC-grade solvents; **A,** $H_2O$ (with 0.05% formic acid) and **B,** acetonitrile (with 0.045% formic acid). Typical settings: 20 min gradient starting at 50% **B**; 50% **B** for 1 min; ramping to 90% **B** in 14 min; 90% **B** for 2 min; ramping to 50% **B** in 2 min; two minutes at 50% **B**.

Detected peaks were collected according to mass and collected in glass test tubes. Identical fraction-containing tubes were pooled. Residual acetonitrile was removed *in vacuo* prior to freeze-drying. All compounds were stored at - 20 °C until HRMS and/or NMR analysis. Mass data was analysed with MassLynx$^{TM}$ (*Waters*).

### 7.6.2. Nuclear Magnetic Resonance

NMR data was acquired by technical staff at the NMR core facility at the institute for organic chemistry (Hannover) on either of the following spectrometers: **i)** a Bruker Ascend 600 MHz with Avance NEO console, sample case and cryo-cooled probe DUL; **ii)** a Bruker Ultrashield 500 MHz with Avance IIIHD console, Sample Xpress and cryo-cooled probe TCI; **iii)** a Bruker Ascend 400 MHz with Avance III console, Sample Xpress and Prodigy BBFO probe; **iv)** a Bruker Ascend 400 MHz with Avance IIIHD console, BACS and BB; **v)** a Bruker Ultrashield 400 MHz with Avance I console, BACS and DUL. Data was acquired at 400/500/600 MHz ($^{1}$H) and 125 MHz ($^{13}$C) respectively. Standard parameters were deployed for 2D NMR spectra.

### 7.6.3. High Resolution Mass Spectrometry

HRMS data was acquired by technical staff at the MS core facility at the institute for organic chemistry (Hannover) using a HR UPLC MS/MS; Water QTof premier (ESI- and APCI-MS/MS) with UPLC (Waters Acquity incl. TUV detector).

### 7.6.4. Optical Rotation

Optical rotations $[\alpha]_D$ were acquired on a P3000 polarimeter (*A. Krüss* Optotronic, $\lambda$ = 589 nm).

### 7.6.5. Gas Chromatography Mass Spectrometry

GCMS data was obtained on a HP 6890 gas chromatograph that was connected to a 5973 mass detector (*Agilent*). A AG19091J-433 capillary column (*Agilent*; 30 m, 0.32 mm i.d., 0.25-µm film) was used. The deployed instrumental parameters were as follows: **1**, He at 2.0 mL min$^{-1}$; **2,** injection volume 1 µL; **3,** transfer line 250 °C; **4,** electron energy 69.9 eV. The GC was set to **1**, initial temperature: 100 °C; initial time: 2 min; **2,** temperature gradient A: 100 → 180 °C at 2 °C min$^{-1}$; **3**, temperature gradient B: 180 → 300 °C at 40 °C min$^{-1}$; **3,** final temperature: 300 °C; final hold time: 7 min; total run time: 52 min. GCMS was operated in splitless-mode.

### 7.6.6. Extraction of *Phaeosphaeriaceae* sp. CF-150626 Cultures

Submerged liquid cultures of *Phaeosphaeriaceae* sp. CF-150626 were mixed with equal volumes of acetone and stirred at ambient temperature for 1 h. After filtration, the acetone was removed from the filtrate under reduced pressure. The remaining aqueous phase was extracted two times with equal volumes EtOAc. EtOAc layers were combined, dried over anhydrous magnesium sulfate and organic solvent was removed under reduced pressure. The crude extract was either stored at - 20 °C or taken up in [dichloromethane:methanol (1:1) or acetonitrile:water (9:1)] to give a final concentration of 10 mg $mL^{-1}$ for analytical LCMS and 50 mg $mL^{-1}$ for preparative LCMS. Notably, after stirring the mycelia in acetone and subsequent filtration the residual mycelia still contained substantial amounts of tropolone sesquiterpenoids. The residual mycelia was stirred in EtOAc at ambient temperature for at least 20 min and filtrated again. The organic solvent was dried over anhydrous magnesium sulfate, removed under reduced pressure and the crude extract was combined with the extract obtained from the previous step.

### 7.6.7. Extraction of *Aspergillus oryzae* NSAR1 Cultures

Submerged liquid cultures of *A. oryzae* NSAR1 were acidified to pH = ~2 (with 2 M HCl), mixed with equal volumes of EtOAc, blended to homogeneity and stirred at ambient temperature for at least 20 min. After Büchner filtration organic and aqueous layers were separated. The aqueous layer was extracted two additional times with EtOAc. All organic layers were combined, dried over anhydrous magnesium sulfate and the solvent was removed under reduced pressure. The crude extract was either stored at - 20 °C or prepared for LCMS. In the latter case the extract was dissolved in [dichloromethane:methanol (1:1) or acetonitrile:water (9:1)] to give a final concentration of 10 mg $mL^{-1}$ for analytical LCMS and 50 mg $mL^{-1}$ for preparative LCMS.

# References


[1] K. C. L. Mulder, F. Mulinari, O. L. Franco, M. S. F. Soares, B. S. Magalhães, N. S. Parachin, *Biotechnol. Adv.* **2015**, *33*, 648–665.

[2] N. Srivastava, A. Akhila, *J. Plant Interact.* **2011**, *6*, 265–273.

[3] J. Davies, K. S. Ryan, *ACS Chem. Biol.* **2012**, *7*, 252–259.

[4] C. Greco, N. P. Keller, A. Rokas, *Curr. Opin. Microbiol.* **2019**, *51*, 22–29.

[5] G. Yim, H. H. Wang, J. Davies, *Philos. Trans. R. Soc. B* **2007**, *362*, 1195–1200.

[6] C. C. Loron, C. François, R. H. Rainbird, E. C. Turner, S. Borensztajn, E. J. Javaux, *Nature* **2019**, *570*, 232–235.

[7] F. Vinale, K. Sivasithamparam, E. L. Ghisalberti, M. Ruocco, S. Woo, M. Lorito, *Nat. Prod. Commun.* **2012**, *7*, 1545–1550.

[8] A. Rokas, M. E. Mead, J. L. Steenwyk, H. A. Raja, N. H. Oberlies, *Nat. Prod. Rep.* **2020**, *37*, 868–878.

[9] J. Kominek, D. T. Doering, D. A. Opulente, X. X. Shen, X. Zhou, J. DeVirgilio, A. B. Hulfachor, M. Groenewald, M. A. Mcgee, S. D. Karlen, et al., *Cell* **2019**, *176*, 1356–1366.

[10] S. Rudolph, J. G. Maciá-Vicente, H. Lotz-Winter, M. Schleuning, M. Piepenbring, *Stud. Mycol.* **2018**, *89*, 95–104.

[11] P. Kumar, D. K. Mahato, M. Kamle, T. K. Mohanta, S. G. Kang, *Front. Microbiol.* **2017**, *7*, 1–10.

[12] A. Casadevall, *Pathog. Immun.* **2018**, *3*, 183–196.

[13] M. T. Robey, L. K. Caesar, M. T. Drott, N. P. Keller, N. L. Kelleher, *bioRxiv* **2020**, DOI https://doi.org/10.1101/2020.09.21.307157.

[14] M. H. Medema, R. Kottmann, P. Yilmaz, M. Cummings, J. B. Biggins, K. Blin, I. De Bruijn, Y. H. Chooi, J. Claesen, R. C. Coates, et al., *Nat. Chem. Biol.* **2015**, *11*, 625–631.

[15] A. G. McInnes, D. G. Smith, J. A. Walter, L. C. Vining, J. L. C. Wright, *J. Chem. Soc. Chem. Commun.* **1974**, 282–284.

[16] K. M. Fisch, W. Bakeer, A. A. Yakasai, Z. Song, J. Pedrick, Z. Wasil, A. M. Bailey, C. M. Lazarus, T. J. Simpson, R. J. Cox, *J. Am. Chem. Soc.* **2011**, *133*, 16635–16641.

[17] H. C. Lin, Y. Tsunematsu, S. Dhingra, W. Xu, M. Fukutomi, Y.-H. Chooi, D. E. Cane, A. M. Calvo, K. Watanabe, Y. Tang, *J. Am. Chem. Soc.* **2014**, *136*, 4426–4436.

[18] N. Sin, L. Meng, M. Q. W. Wang, J. J. Wen, W. G. Bornmann, C. M. Crews, *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 6099–6103.

[19] C. Wallwey, S. M. Li, *Nat. Prod. Rep.* **2011**, *28*, 496–510.

[20] R. Geris, T. J. Simpson, *Nat. Prod. Rep.* **2009**, *26*, 1063–1094.

[21] K. J. Weissman, P. F. Leadlay, *Nat. Rev. Microbiol.* **2005**, *3*, 925–936.

[22] R. Nofiani, K. de Mattos-Shipley, K. E. Lebe, L.-C. Han, Z. Iqbal, A. M. Bailey, C. L. Willis, T. J. Simpson, R. J. Cox, *Nat. Commun.* **2018**, *9*, 1–11.

[23] F. Trenti, K. E. Lebe, E. Adelin, J. Ouazzani, C. Schotte, R. J. Cox, *RSC Adv.*

**2020**, *10*, 27369–27376.

[24] K. E. Lebe, R. J. Cox, *Chem. Sci.* **2019**, *10*, 1227–1231.

[25] R. J. Cox, *Org. Biomol. Chem.* **2007**, *5*, 2010–2026.

[26] A. J. Birch, F. W. Donovan, *Aust. J. Chem.* **1953**, *6*, 360–368.

[27] A. J. Birch, R. A. Massy-Westropp, C. J. Moye, *Aust. J. Chem.* **1955**, *8*, 539–544.

[28] B. Shen, *Curr. Opin. Chem. Biol.* **2003**, *7*, 285–295.

[29] J. Staunton, K. J. Weissman, *Nat. Prod. Rep.* **2001**, *18*, 380–416.

[30] Y. H. Chooi, Y. Tang, *J. Org. Chem.* **2012**, *77*, 9933–9953.

[31] V. Y. Alekseyev, C. W. Liu, D. E. Cane, J. D. Puglisi, C. Khosla, *Protein Sci.* **2007**, *16*, 2093–2107.

[32] J. Beld, E. C. Sonnenschein, C. R. Vickery, J. P. Noel, M. D. Burkart, *Nat. Prod. Rep.* **2014**, *31*, 61–108.

[33] J. M. Crawford, A. L. Vagstad, K. P. Whitworth, K. C. Ehrlich, C. A. Townsend, *ChemBioChem* **2008**, *9*, 1019–1023.

[34] J. F. Barajas, G. Shakya, G. Moreno, H. Rivera Jr., D. R. Jackson, C. L. Topper, A. L. Vagstad, J. J. La Clair, C. A. Townsend, M. D. Burkart, et al., *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, E4142–E4148.

[35] L. Du, L. Lou, *Nat. Prod. Rep.* **2010**, *27*, 255–278.

[36] J. M. Crawford, T. P. Korman, J. W. Labonte, A. L. Vagstad, E. A. Hill, O. Kamari-Bidkorpeh, S. C. Tsai, C. A. Townsend, *Nature* **2009**, *461*, 1139–1143.

[37] J. Kennedy, K. Auclair, S. G. Kendrew, C. Park, J. C. Vederas, C. Richard Hutchinson, *Science (80-. ).* **1999**, *284*, 1368 LP – 1372.

[38] I. Soehano, L. Yang, F. Ding, H. Sun, Z. J. Low, X. Liu, Z. X. Liang, *Org. Biomol. Chem.* **2014**, *12*, 8542–8549.

[39] D. W. Christianson, *Chem. Rev.* **2017**, *117*, 11570–11648.

[40] S. Moser, H. Pichler, *Appl. Microbiol. Biotechnol.* **2019**, *103*, 5501–5516.

[41] J. Jiang, D. E. Cane, *J. Am. Chem. Soc.* **2008**, *130*, 428–429.

[42] X. Wang, J. Wu, J. Chen, L. Xiao, Y. Zhang, F. Wang, X. Li, *J. Agric. Food Chem.* **2020**, *68*, 8381–8390.

[43] C. Schmidt-Dannert, in *Biotechnol. Isoprenoids* (Eds.: J. Schrader, J. Bohlmann), Springer International Publishing, Cham, **2015**, pp. 19–61.

[44] M. C. Lemfack, W. Brandt, K. Krüger, A. Gurowietz, J. Djifack, J.-P. Jung, M. Hopf, H. Noack, B. Junker, S. von Reuß, et al., *Sci. Rep.* **2021**, *11*, 3182.

[45] S. Von Reuss, D. Domik, M. C. Lemfack, N. Magnus, M. Kai, T. Weise, B. Piechulla, *J. Am. Chem. Soc.* **2018**, *140*, 11855–11862.

[46] B. Engels, U. Heinig, T. Grothe, M. Stadler, S. Jennewein, *J. Biol. Chem.* **2011**, *286*, 6871–6878.

[47] J. Rinkel, J. S. Dickschat, *Org. Lett.* **2019**, *21*, 2426–2429.

[48] M. B. Quin, C. M. Flynn, C. Schmidt-Dannert, *Nat. Prod. Rep.* **2014**, *31*, 1449–1473.

[49] J. Rinkel, J. S. Dickschat, *Beilstein J. Org. Chem.* **2019**, *15*, 789–794.

[50] G. T. Wawrzyn, M. B. Quin, S. Choudhary, F. López-Gallego, C. Schmidt-

Dannert, *Chem. Biol.* **2012**, *19*, 772–783.

[51] H. A. Gennadios, V. Gonzalez, L. Di Costanzo, A. Li, F. Yu, D. J. Miller, R. K. Allemann, D. W. Christianson, *Biochemistry* **2009**, *48*, 6175–6183.

[52] D. B. Little, R. B. Croteau, *Arch. Biochem. Biophys.* **2002**, *402*, 120–135.

[53] E. Y. Shishova, L. Di Costanzo, D. E. Cane, D. W. Christianson, *Biochemistry* **2007**, *46*, 1941–1951.

[54] J. M. Caruthers, I. Kang, M. J. Rynkiewicz, D. E. Cane, D. W. Christianson, *J. Biol. Chem.* **2000**, *275*, 25533–25539.

[55] P. Moosmann, F. Ecker, S. Leopold-Messer, J. K. B. Cahn, C. L. Dieterich, M. Groll, J. Piel, *Nat. Chem.* **2020**, *12*, 968–972.

[56] M. Jiang, Z. Wu, L. Liu, S. Chen, *Org. Biomol. Chem.* **2020**, *19*, 1644–1704.

[57] M. D. Sintchak, M. A. Fleming, O. Futer, S. A. Raybuck, S. P. Chambers, P. R. Caron, M. A. Murcko, K. P. Wilson, *Cell* **1996**, *85*, 921–930.

[58] L. M. Shaw, H. W. Sollinger, P. Halloran, R. E. Morris, R. W. Yatscoff, J. Ransom, I. Tsina, P. Keown, D. W. Holt, R. Lieberman, et al., *Ther. Drug Monit.* **1995**, *17*, 690–699.

[59] Y. Matsuda, I. Abe, *Nat. Prod. Rep.* **2016**, *33*, 26–53.

[60] M. C. McCowen, M. E. Callender, J. F. Lawlis, *Science (80-. ).* **1951**, *113*, 202–203.

[61] J.-M. Molina, M. Tourneur, C. Sarfati, S. Chevret, A. de Gouvello, J.-G. Gobert, S. Balkan, F. Derouin, *N. Engl. J. Med.* **2002**, *346*, 1963–1969.

[62] J. E. Sears, D. L. Boger, *Acc. Chem. Res.* **2015**, *48*, 653–662.

[63] L. Barra, I. Abe, *Nat. Prod. Rep.* **2021**, *38*, 566–585.

[64] Y. Matsuda, T. Wakimoto, T. Mori, T. Awakawa, I. Abe, *J. Am. Chem. Soc.* **2014**, *136*, 15326–15336.

[65] H. C. Lo, R. Entwistle, C. J. Guo, M. Ahuja, E. Szewczyk, J. H. Hung, Y. M. Chiang, B. R. Oakley, C. C. C. Wang, *J. Am. Chem. Soc.* **2012**, *134*, 4709–4720.

[66] T. Itoh, K. Tokunaga, E. K. Radhakrishnan, I. Fujii, I. Abe, Y. Ebizuka, T. Kushiro, *ChemBioChem* **2012**, *13*, 1132–1135.

[67] Y. Matsuda, T. Awakawa, I. Abe, *Tetrahedron* **2013**, *69*, 8199–8204.

[68] E. I. Solomon, S. Goudarzi, K. D. Sutherlin, *Biochemistry* **2016**, *55*, 6363–6374.

[69] M. Toplak, A. Matthews, R. Teufel, *Arch. Biochem. Biophys.* **2021**, *698*, 108732.

[70] L. Kahlert, E. F. Bassiony, R. J. Cox, E. J. Skellam, *Angew. Chemie - Int. Ed.* **2020**, *59*, 5816–5822.

[71] P. R. Ortiz De Montellano, *Chem. Rev.* **2010**, *110*, 932–948.

[72] B. Meunier, S. P. de Visser, S. Shaik, *Chem. Rev.* **2004**, *104*, 3947–3980.

[73] C. Wang, K. Becker, S. Pfütze, E. Kuhnert, M. Stadler, R. J. Cox, E. Skellam, *Org. Lett.* **2019**, *21*, 8756–8760.

[74] Y. Nakashima, T. Mitsuhashi, Y. Matsuda, M. Senda, H. Sato, M. Yamazaki, M. Uchiyama, T. Senda, I. Abe, *J. Am. Chem. Soc.* **2018**, *140*, 9743–9750.

[75] H. Nakamura, Y. Matsuda, I. Abe, *Nat. Prod. Rep.* **2018**, *35*, 633–645.

[76] H. J. Liao, J. Li, J. L. Huang, M. Davidson, I. Kurnikov, T. S. Lin, J. L. Lee, M. Kurnikova, Y. Guo, N. L. Chan, et al., *Angew. Chemie - Int. Ed.* **2018**, *57*, 1831–

1835.

[77] N. Ishikawa, H. Tanaka, F. Koyama, H. Noguchi, C. C. C. Wang, K. Hotta, K. Watanabe, *Angew. Chemie - Int. Ed.* **2014**, *53*, 12880–12884.

[78] F. Mayerl, Q. Gao, S. Huang, S. E. Klohr, J. A. Matson, D. R. Gustavson, D. M. Pirnik, R. L. Berry, C. Fairchild, W. C. Rose, *J. Antibiot. (Tokyo).* **1993**, *46*, 1082–1088.

[79] T. Bunyapaiboonsri, S. Veeranondha, T. Boonruangprapa, S. Somrithipol, *Phytochem. Lett.* **2008**, *1*, 204–206.

[80] Y. Zhai, Y. Li, J. Zhang, Y. Zhang, F. Ren, X. Zhang, G. Liu, X. Liu, Y. Che, *Fungal Genet. Biol.* **2019**, *129*, 7–15.

[81] T. El-Elimat, H. A. Raja, S. Ayers, S. J. Kurina, J. E. Burdette, Z. Mattes, R. Sabatelle, J. W. Bacon, A. H. Colby, M. W. Grinstaff, et al., *Org. Lett.* **2019**, *21*, 529–534.

[82] Z. Y. Al Subeh, N. Q. Chu, J. T. Korunes-Miller, L. L. Tsai, T. N. Graf, Y. P. Hung, C. J. Pearce, M. W. Grinstaff, A. H. Colby, Y. L. Colson, et al., *J. Control. Release* **2021**, *331*, 260–269.

[83] G. H. Harris, K. Hoogsteen, K. C. Silverman, S. L. Raghoobar, G. F. Bills, R. B. Lingham, J. L. Smith, H. W. Dougherty, C. Cascales, F. Peláez, *Tetrahedron* **1993**, *49*, 2139–2144.

[84] C.-J. Hsiao, S.-H. Hsiao, W.-L. Chen, J.-H. Guh, G. Hsiao, Y.-J. Chan, T.-H. Lee, C.-L. Chung, *Chem. Biol. Interact.* **2012**, *197*, 23–30.

[85] P. Cai, D. Smith, B. Cunningham, S. Brown-Shimer, B. Katz, C. Pearce, D. Venables, D. Houck, *J. Nat. Prod.* **1998**, *61*, 791–795.

[86] P. Pittayakhajonwut, M. Theerasilp, P. Kongsaeree, A. Rungrod, M. Tanticharoen, Y. Thebtaranonth, *Planta Med.* **2002**, *68*, 1017–1019.

[87] M. E. Raggatt, T. J. Simpson, M. Inês Chicarelli-Robinson, *Chem. Commun.* **1997**, 2245–2246.

[88] F. Yu, S. Okamto, K. Nakasone, K. Adachi, S. Matsuda, H. Harada, N. Misawa, R. Utsumi, *Planta* **2008**, *227*, 1291–1299.

[89] U. Neuenschwander, B. Czarniecki, I. Hermans, *J. Org. Chem.* **2012**, *77*, 2865–2869.

[90] N. L. Brock, K. Huss, B. Tudzynski, J. S. Dickschat, *ChemBioChem* **2013**, *14*, 311–315.

[91] D. B. Amby, T. Manczak, M. A. Petersen, T. Sundelin, C. Weitzel, M. Grajewski, H. T. Simonsen, B. Jensen, *Microbiology* **2016**, *162*, 1773–1783.

[92] M. B. Quin, C. M. Flynn, G. T. Wawrzyn, S. Choudhary, C. Schmidt-Dannert, *ChemBioChem* **2013**, *14*, 2480–2491.

[93] J. Rinkel, J. S. Dickschat, *Beilstein J. Org. Chem.* **2015**, *11*, 2493–2508.

[94] R. Schoenheimer, D. Rittenberg, in *Science (80-. ).*, **1935**, pp. 156–158.

[95] D. J. Wilkinson, *Mass Spectrom. Rev.* **2018**, *37*, 57–80.

[96] A. Bacher, C. Rieder, D. Eichinger, D. Arigoni, G. Fuchs, W. Eisenreich, *FEMS Microbiol. Rev.* **1999**, *22*, 567–598.

[97] T. J. Simpson, in *Biosynth. Polyketides Vitam.* (Eds.: F.J. Leeper, J.C. Vederas), Springer Berlin Heidelberg, Berlin, Heidelberg, **1998**, pp. 1–48.

[98] T. J. Simpson, *Tetrahedron Lett.* **1981**, *22*, 3785–3788.

[99] C. R. McIntyre, T. J. Simpson, D. J. Stenzel, A. J. Bartlett, E. O'Brien, J. S. E. Holker, *J. Chem. Soc. Chem. Commun.* **1982**, 781–782.

[100] C. A. Jones, P. J. Sidebottom, R. J. P. Cannell, D. Noble, B. A. M. Rudd, *J. Antibiot. (Tokyo).* **1992**, *45*, 1492–1498.

[101] F. Trenti, R. J. Cox, *J. Nat. Prod.* **2017**, *80*, 1235–1240.

[102] G. Li, Y. W. Guo, J. S. Dickschat, *Angew. Chemie - Int. Ed.* **2021**, *60*, 1488–1492.

[103] M. N. Heneghan, A. A. Yakasai, L. M. Halo, Z. Song, A. M. Bailey, T. J. Simpson, R. J. Cox, C. M. Lazarus, *ChemBioChem* **2010**, *11*, 1508–1512.

[104] E. Skellam, *Trends Biotechnol.* **2019**, *37*, 416–427.

[105] K. A. K. Pahirulzaman, K. Williams, C. M. Lazarus, *Methods Enzymol.* **2012**, *517*, 241–260.

[106] Z. Song, W. Bakeer, J. W. Marshall, A. A. Yakasai, R. M. Khalid, J. Collemare, E. Skellam, D. Tharreau, M. H. Lebrun, C. M. Lazarus, et al., *Chem. Sci.* **2015**, *6*, 4837–4845.

[107] V. Hantke, C. Wang, E. J. Skellam, R. J. Cox, *RSC Adv.* **2019**, *9*, 35797–35802.

[108] A. M. Bailey, F. Alberti, S. Kilaru, C. M. Collins, K. De Mattos-Shipley, A. J. Hartley, P. Hayes, A. Griffin, C. M. Lazarus, R. J. Cox, et al., *Sci. Rep.* **2016**, *6*, 25202.

[109] F. Alberti, K. Khairudin, E. R. Venegas, J. A. Davies, P. M. Hayes, C. L. Willis, A. M. Bailey, G. D. Foster, *Nat. Commun.* **2017**, *8*, 1831.

[110] M. Stadler, D. Hoffmeister, *Front. Microbiol.* **2015**, *6*, 1–4.

[111] J. Fricke, F. Blei, D. Hoffmeister, *Angew. Chemie - Int. Ed.* **2017**, *56*, 12352–12355.

[112] N. Milne, P. Thomsen, N. Mølgaard Knudsen, P. Rubaszka, M. Kristensen, I. Borodina, *Metab. Eng.* **2020**, *60*, 25–36.

[113] L. Kahlert, C. Schotte, R. J. Cox, *Synthesis (Stuttg).* **2021**, DOI 10.1055/a-1401-2716.

[114] B. Wang, F. Guo, S. H. Dong, H. Zhao, *Nat. Chem. Biol.* **2019**, *15*, 111–114.

[115] C. J. B. Harvey, M. Tang, U. Schlecht, J. Horecka, C. R. Fischer, H.-C. Lin, J. Li, B. Naughton, J. Cherry, M. Miranda, et al., *Sci. Adv.* **2018**, *4*, eaar5459.

[116] D. Jarczynska, J. K. H. Rendsvig, N. Pagels, V. R. Viana, C. S. Nødvig, F. H. Kirchner, T. Strucko, M. L. Nielsen, U. H. Mortensen, **2021**, *10*, 579–588.

[117] F. J. Jin, J. I. Maruyama, P. R. Juvvadi, M. Arioka, K. Kitamoto, *FEMS Microbiol. Lett.* **2004**, *239*, 79–85.

[118] O. Yamada, B. Rho Lee, K. Gomi, *Biosci. Biotechnol. Biochem.* **1997**, *61*, 1367–1369.

[119] O. Hara, H. Anzai, S. Imai, Y. Kumada, T. Murakami, R. Itoh, E. Takano, A. Satoh, K. Nagaoka, *J. Antibiot. (Tokyo).* **1988**, *41*, 538–547.

[120] I. E. Mattern, P. J. Punt, C. A. M. J. J. Van Den Hondel, *Fungal Genet. Rep.* **1988**, *35*.

[121] R. Schor, C. Schotte, D. Wibberg, J. Kalinowski, R. J. Cox, *Nat. Commun.* **2018**, *9*, 1963.

[122] C. Schotte, L. Li, D. Wibberg, J. Kalinowski, R. J. Cox, *Angew. Chemie - Int. Ed.* **2020**, *59*, 23870–23878.

[123] R. Schor, Biosynthesis of Xenovulenes in Acremonium Strictum, Gottfried Wilhelm Leibniz-Universität Hannover, Diss., **2018**.

[124] A. M. Ainsworth, M. I. Chicarelli-Robinson, B. R. Copp, U. Fauth, P. J. Hylands, J. A. Holloway, M. Latif, G. B. O´Beirne, N. Porter, D. V Renno, et al., *J. Antibiot. (Tokyo).* **1995**, *48*, 568–573.

[125] P. Thomas, H. Sundaram, B. J. Krishek, P. Chazot, X. Xie, P. Bevan, S. J. Brocchini, C. J. Latham, P. Charlton, M. Moore, et al., *J. Pharmacol. Exp. Ther.* **1997**, *282*, 513–20.

[126] W. E. Haefely, *Int. Anesthesiol. Clin.* **1988**, *26*, 262–272.

[127] R. Bentley, *J. Biol. Chem.* **1963**, *238*, 1895–1902.

[128] Z. Zhang, C. S. Jamieson, Y. L. Zhao, D. Li, M. Ohashi, K. N. Houk, Y. Tang, *J. Am. Chem. Soc.* **2019**, *141*, 5659–5663.

[129] V. Hantke, E. J. Skellam, R. J. Cox, *Chem. Commun.* **2020**, *56*, 2925–2928.

[130] T. Ose, K. Watanabe, T. Mie, M. Honma, H. Watanabe, M. Yao, H. Oikawa, I. Tanaka, *Nature* **2003**, *422*, 185–189.

[131] L. Kahlert, R. Cox, E. Skellam, *Chem. Commun.* **2020**, *56*, 10934–10937.

[132] L. Gao, C. Su, X. Du, R. Wang, S. Chen, Y. Zhou, C. Liu, X. Liu, R. Tian, L. Zhang, et al., *Nat. Chem.* **2020**, *12*, 620–628.

[133] H. Weenen, M. H. H. Nkunya, A. A. El-Fadl, S. Harkema, B. Zwanenburg, *J. Org. Chem.* **1990**, *55*, 5107–5109.

[134] R. M. Adlington, J. E. Baldwin, G. J. Pritchard, A. J. Williams, D. J. Watkin, *Org. Lett.* **1999**, *1*, 1937–1939.

[135] R. M. Adlington, J. E. Baldwin, A. V. W. Mayweg, G. J. Pritchard, *Org. Lett.* **2002**, *4*, 3009–3011.

[136] P.-J. Li, G. Dräger, A. Kirschning, *Org. Lett.* **2019**, *21*, 998–1001.

[137] A. M. Bailey, R. J. Cox, K. Harley, C. M. Lazarus, T. J. Simpson, E. Skellam, *Chem. Commun.* **2007**, 4053–4055.

[138] R. W. Bryant, R. J. Light, *Biochemistry* **1974**, *13*, 1516–1522.

[139] H. Guo, D. Roman, C. Beemelmanns, *Nat. Prod. Rep.* **2019**, *36*, 1137–1155.

[140] J. Davison, A. al Fahad, M. Cai, Z. Song, S. Y. Yehia, C. M. Lazarus, A. M. Bailey, T. J. Simpson, R. J. Cox, *Proc. Natl. Acad. Sci.* **2012**, *109*, 7642–7647.

[141] T. Doyon, J., K. Skinner, D. Yang, L. Mallik, T. Wymore, M. Koutmos, P. M. Zimmerman, A. R. H. Narayan, *ChemRxiv* **2020**, 11–13.

[142] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, *J. Mol. Biol.* **1990**, *215*, 403–10.

[143] C. L. M. Gilchrist, Y.-H. Chooi, *Bioinformatics* **2021**, 1–3.

[144] J. L. Hartley, G. F. Temple, M. A. Brasch, *Genome Res.* **2000**, *10*, 1788–1795.

[145] K. H. M. Nazmul Hussain Nazir, H. Ichinose, H. Wariishi, *Arch. Microbiol.* **2010**, *192*, 395–408.

[146] A. L. Mitchell, T. K. Attwood, P. C. Babbitt, M. Blum, P. Bork, A. Bridge, S. D. Brown, H. Y. Chang, S. El-Gebali, M. I. Fraser, et al., *Nucleic Acids Res.* **2019**, *47*, D351–D360.

[147] A. Marchler-Bauer, M. K. Derbyshire, N. R. Gonzales, S. Lu, F. Chitsaz, L. Y. Geer, R. C. Geer, J. He, M. Gwadz, D. I. Hurwitz, et al., *Nucleic Acids Res.* **2014**, *43*, D222–D226.

[148] L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, M. J. E. Sternberg, *Nat. Protoc.* **2015**, *10*, 845–858.

[149] A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F. T. Heer, T. A. P. De Beer, C. Rempfer, L. Bordoli, et al., *Nucleic Acids Res.* **2018**, *46*, W296–W303.

[150] K. Terpe, *Appl. Microbiol. Biotechnol.* **2003**, *60*, 523–533.

[151] D. Hartinger, S. Heinl, H. E. Schwartz, R. Grabherr, G. Schatzmayr, D. Haltrich, W. D. Moll, *Microb. Cell Fact.* **2010**, *9*, 1–14.

[152] M. Ferrer, T. N. Chernikova, M. M. Yakimov, P. N. Golyshin, K. N. Timmis, *Nat. Biotechnol.* **2003**, *21*, 1266–1267.

[153] C. Oberhauser, V. Harms, K. Seidel, B. Schröder, K. Ekramzadeh, S. Beutel, S. Winkler, L. Lauterbach, J. S. Dickschat, A. Kirschning, *Angew. Chemie - Int. Ed.* **2018**, *57*, 11802–11806.

[154] K. Nishihara, M. Kanemori, M. Kitagawa, H. Yanagi, T. Yura, *Appl. Environ. Microbiol.* **1998**, *64*, 1694–1699.

[155] L. Käll, A. Krogh, E. L. L. Sonnhammer, *Nucleic Acids Res.* **2007**, *35*, W429–W432.

[156] T. Doyon, Development and Characterization of Non-Heme Iron Biocatalysts for Complex Molecule Synthesis, University of Michigan, Diss., **2020**.

[157] C. S. Jamieson, M. Ohashi, F. Liu, Y. Tang, K. N. Houk, *Nat. Prod. Rep.* **2019**, *36*, 698–713.

[158] Y. Cai, Y. Hai, M. Ohashi, C. S. Jamieson, M. Garcia-Borras, K. N. Houk, J. Zhou, Y. Tang, *Nat. Chem.* **2019**, *11*, 812–820.

[159] K. Auclair, A. Sutherland, J. Kennedy, D. J. Witter, J. P. Van den Heever, C. R. Hutchinson, J. C. Vederas, *J. Am. Chem. Soc.* **2000**, *122*, 11519–11520.

[160] H. Oikawa, K. Katayama, Y. Suzuki, A. Ichihara, *J. Chem. Soc. Chem. Commun.* **1995**, 1321–1322.

[161] K. Kasahara, T. Miyamoto, T. Fujimoto, H. Oguri, T. Tokiwano, H. Oikawa, Y. Ebizuka, I. Fujii, *ChemBioChem* **2010**, *11*, 1245–1252.

[162] H. J. Kim, M. W. Ruszczycky, S. H. Choi, Y. N. Liu, H. W. Liu, *Nature* **2011**, *473*, 109–112.

[163] M. Ohashi, F. Liu, Y. Hai, M. Chen, M. Tang, Z. Yang, M. Sato, K. Watanabe, K. N. Houk, Y. Tang, *Nature* **2017**, *549*, 502–506.

[164] F. Madeira, Y. M. Park, J. Lee, N. Buso, T. Gur, N. Madhusoodanan, P. Basutkar, A. R. N. Tivey, S. C. Potter, R. D. Finn, et al., *Nucleic Acids Res.* **2019**, *47*, W636–W641.

[165] F. Yu, S. Okamoto, H. Harada, K. Yamasaki, N. Misawa, R. Utsumi, *Cell. Mol. Life Sci.* **2011**, *68*, 1033–1040.

[166] M. C. Tang, Y. Zou, K. Watanabe, C. T. Walsh, Y. Tang, *Chem. Rev.* **2017**, *117*, 5226–5333.

[167] C. Zaehle, M. Gressler, E. Shelest, E. Geib, C. Hertweck, M. Brock, *Chem. Biol.* **2014**, *21*, 719–731.

[168] E. L. F. Ferreira, D. E. Williams, L. P. Ióca, R. P. Morais-Urano, M. F. C. Santos, B. O. Patrick, L. M. Elias, S. P. Lira, A. G. Ferreira, M. R. Z. Passarini, et al., *Org. Lett.* **2015**, *17*, 5152–5155.

[169] C. A. Helliwell, P. M. Chandler, A. Poole, E. S. Dennis, W. J. Peacock, *Proc.*

*Natl. Acad. Sci. U. S. A.* **2001**, *98*, 2065–2070.

[170] Z. Yunt, K. Reinhardt, A. Li, M. Engeser, H. M. Dahse, M. Gütschow, T. Bruhn, G. Bringmann, J. Piel, *J. Am. Chem. Soc.* **2009**, *131*, 2297–2305.

[171] Y. Tsunematsu, N. Ishikawa, D. Wakana, Y. Goda, H. Noguchi, H. Moriya, K. Hotta, K. Watanabe, *Nat. Chem. Biol.* **2013**, *9*, 818–825.

[172] Q. Y. Qi, L. Bao, J. W. Ren, J. J. Han, Z. Y. Zhang, Y. Li, Y. J. Yao, R. Cao, H. W. Liu, *Org. Lett.* **2014**, *16*, 5092–5095.

[173] X. W. Yang, Y. P. Li, J. Su, W. G. Ma, G. Xu, *Org. Lett.* **2016**, *18*, 1876–1879.

[174] S. Ayers, D. L. Zink, J. S. Powell, C. M. Brown, A. Grund, G. F. Bills, G. Platas, D. Thompson, S. B. Singh, *J. Nat. Prod.* **2008**, *71*, 457–459.

[175] Q. Chen, J. Gao, C. Jamieson, J. Liu, M. Ohashi, J. Bai, D. Yan, B. Liu, Y. Che, Y. Wang, et al., *J. Am. Chem. Soc.* **2019**, *141*, 14052–14056.

[176] S. Chen, Z. Liu, H. Tan, Y. Chen, S. Li, H. Li, S. Zhu, H. Liu, W. Zhang, *Org. Chem. Front.* **2020**, *7*, 557–562.

[177] J. Zhang, Y. Li, F. Ren, Y. Zhang, X. Liu, L. Liu, Y. Che, *J. Nat. Prod.* **2019**, *82*, 1678–1685.

[178] J. T. Lin, W. H. Liu, *J. Agric. Food Chem.* **2006**, *54*, 7564–7569.

[179] J. A. Ballantine, C. H. Hassall, B. D. Jones, *Phytochemistry* **1968**, *7*, 1529–1534.

[180] K. Fushimi, K. Anzai, S. Tokuyama, Y. Kiriiwa, N. Matsumoto, A. Sekiya, D. Hashizume, K. Nagasawa, H. Hirai, H. Kawagishi, *Tetrahedron* **2012**, *68*, 1262–1265.

[181] I. V. Grigoriev, R. Nikitin, S. Haridas, A. Kuo, R. Ohm, R. Otillar, R. Riley, A. Salamov, X. Zhao, F. Korzeniewski, et al., *Nucleic Acids Res.* **2014**, *42*, D699–D704.

[182] C. L. Schoch, K. A. Seifert, S. Huhndorf, V. Robert, J. L. Spouge, C. A. Levesque, W. Chen, E. Bolchacova, K. Voigt, P. W. Crous, et al., *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 6241–6246.

[183] R. Phookamsak, D. N. Wanasinghe, S. Hongsanan, C. Phukhamsakda, S. K. Huang, D. S. Tennakoon, C. Norphanphoun, E. Camporesi, T. S. Bulgakov, I. Promputtha, et al., *Fungal Divers.* **2017**, *87*, 299–339.

[184] M. Hernández-Restrepo, R. K. Schumacher, M. J. Wingfield, I. Ahmad, L. Cai, T. A. Duong, J. Edwards, J. Gené, J. Z. Groenewald, S. Jabeen, et al., *Sydowia* **2016**, *68*, 193–230.

[185] J. Zhang, L. Liu, B. Wang, Y. Zhang, L. Wang, X. Liu, Y. Che, *J. Nat. Prod.* **2015**, *78*, 3058–3066.

[186] M. Ahuja, Y. M. Chiang, S. L. Chang, M. B. Praseuth, R. Entwistle, J. F. Sanchez, H. C. Lo, H. H. Yeh, B. R. Oakley, C. C. C. Wang, *J. Am. Chem. Soc.* **2012**, *134*, 8212–8221.

[187] A. Gurevich, V. Saveliev, N. Vyahhi, G. Tesler, *Bioinformatics* **2013**, *29*, 1072–1075.

[188] D. Wibberg, M. Stadler, C. Lambert, B. Bunk, C. Spröer, C. Rückert, J. Kalinowski, R. J. Cox, E. Kuhnert, *Fungal Divers.* **2021**, *106*, 7–28.

[189] K. Blin, T. Wolf, M. G. Chevrette, X. Lu, C. J. Schwalen, S. A. Kautsar, H. G. Suarez Duran, E. L. C. De Los Santos, H. U. Kim, M. Nave, et al., *Nucleic Acids Res.* **2017**, *45*, W36–W41.

[190] K. Blin, S. Shaw, K. Steinke, R. Villebro, N. Ziemert, S. Y. Lee, M. H. Medema,

T. Weber, *Nucleic Acids Res.* **2019**, *47*, W81–W87.

[191] H. Shirahama, G. S. Arora, T. Matsumoto, *Chem. Lett.* **1983**, *12*, 281–282.

[192] T. Shiina, T. Ozaki, Y. Matsu, S. Nagamine, C. Liu, M. Hashimoto, A. Minami, H. Oikawa, *Org. Lett.* **2020**, *22*, 1997–2001.

[193] Z. Qin, R. Devine, T. J. Booth, E. H. E. Farrar, M. N. Grayson, M. I. Hutchings, B. Wilkinson, *Chem. Sci.* **2020**, *11*, 8125–8131.

[194] K. Breese, G. Fuchs, *Eur. J. Biochem.* **1998**, *251*, 916–923.

[195] M. Boll, *Biochim. Biophys. Acta - Bioenerg.* **2005**, *1707*, 34–50.

[196] M. Hirama, S. Itô, *Tetrahedron Lett.* **1976**, *17*, 2339–2342.

[197] C.-M. Wang, R. Hopson, X. Lin, D. E. Cane, *J. Am. Chem. Soc.* **2009**, *131*, 8360–8361.

[198] D. Arigoni, *Pure Appl. Chem.* **1975**, *41*, 219–245.

[199] C. L. Steele, J. Crock, J. Bohlmann, R. Croteau, *J. Biol. Chem.* **1998**, *273*, 2078–2089.

[200] D. E. Cane, R. Iyengar, M. S. Shiao, *J. Am. Chem. Soc.* **1981**, *103*, 914–931.

[201] S. C. Wang, D. J. Tantillo, *Org. Lett.* **2008**, *10*, 4827–4830.

[202] Z. Li, R. Gao, Q. Hao, H. Zhao, L. Cheng, F. He, L. Liu, X. Liu, W. K. W. Chou, H. Zhu, et al., *Biochemistry* **2016**, *55*, 6599–6604.

[203] D. E. Cane, Q. Xue, *J. Am. Chem. Soc.* **1996**, *118*, 1563–1564.

[204] F. Lopez-Gallego, S. A. Agger, D. A. Pella, M. D. Distefano, C. Schmidt-Dannert, *ChemBioChem* **2010**, *11*, 1093–1106.

[205] M. J. Rynkiewicz, D. E. Cane, D. W. Christianson, *Proc. Natl. Acad. Sci.* **2001**, *98*, 13543–13548.

[206] E. Hochuli, W. Bannwarth, H. Dobeli, R. Gentz, D. Stüber, *Nat. Biotechnol.* **1988**, *6*, 1321–1325.

[207] D. R. Smyth, M. K. Mrozkiewicz, W. J. McGrath, P. Listwan, B. Kobe, *Protein Sci.* **2003**, *12*, 1313–1322.

[208] M. P. Strub, F. Hoh, J. F. Sanchez, J. M. Strub, A. Böck, A. Aumelas, C. Dumas, *Structure* **2003**, *11*, 1359–1367.

[209] F. W. Studier, *Protein Expr. Purif.* **2005**, *41*, 207–34.

[210] C. I. Liu, G. Y. Liu, Y. Song, F. Yin, M. E. Hensler, W. Y. Jeng, V. Nizet, A. H. J. Wang, E. Oldfield, *Science (80-. ).* **2008**, *319*, 1391–1394.

[211] M. Chen, N. Al-Lami, M. Janvier, E. L. D'Antonio, J. A. Faraldos, D. E. Cane, R. K. Allemann, D. W. Christianson, *Biochemistry* **2013**, *52*, 5441–5453.

[212] R. M. Phan, C. D. Poulter, *J. Org. Chem.* **2001**, *66*, 6705–6710.

[213] G. Ramamoorthy, R. M. Phan, C. D. Poulter, *J. Org. Chem.* **2016**, *81*, 5093–5100.

[214] R. K. Keller, R. Thompson, *J. Chromatogr. A* **1993**, *645*, 161–167.

[215] E. R. Soares, F. M. A. Da Silva, R. A. De Almeida, B. R. De Lima, H. H. F. Koolen, C. C. Lourenço, M. J. Salvador, A. Flach, L. A. M. A. Da Costa, A. Q. L. De Souza, et al., *Nat. Prod. Res.* **2015**, *29*, 1285–1288.

[216] L. Holm, *Protein Sci.* **2020**, *29*, 128–140.

[217] P. Baer, P. Rabe, K. Fischer, C. A. Citron, T. A. Klapschinski, M. Groll, J. S.

Dickschat, *Angew. Chemie - Int. Ed.* **2014**, *53*, 7652–7656.

[218] J. A. Aaron, X. Lin, D. E. Cane, D. W. Christianson, *Biochemistry* **2010**, *49*, 1787–1797.

[219] R. Driller, S. Janke, M. Fuchs, E. Warner, A. R. Mhashal, D. T. Major, M. Christmann, T. Brück, B. Loll, *Nat. Commun.* **2018**, *9*, 3971.

[220] W. Liu, X. Feng, Y. Zheng, C. H. Huang, C. Nakano, T. Hoshino, S. Bogue, T. P. Ko, C. C. Chen, Y. Cui, et al., *Sci. Rep.* **2014**, *4*, 6125.

[221] G. E. Crooks, G. Hon, J. M. Chandonia, S. E. Brenner, *Genome Res.* **2004**, *14*, 1188–1190.

[222] P. Rabe, T. Schmitz, J. S. Dickschat, *Beilstein J. Org. Chem.* **2016**, *12*, 1839–1850.

[223] C. Jenson, W. L. Jorgensen, *J. Am. Chem. Soc.* **1997**, *119*, 10846–10854.

[224] J. Chiu, P. E. March, R. Lee, D. Tillett, *Nucleic Acids Res.* **2004**, *32*, e174.

[225] J. Chappell, R. M. Coates, *Compr. Nat. Prod. II Chem. Biol.* **2010**, *1*, 609–641.

[226] T. E. O'Brien, S. J. Bertolani, D. J. Tantillo, J. B. Siegel, *Chem. Sci.* **2016**, *7*, 4009–4015.

[227] R. Li, W. K. W. Chou, J. A. Himmelberger, K. M. Litwin, G. G. Harris, D. E. Cane, D. W. Christianson, *Biochemistry* **2014**, *53*, 1155–1168.

[228] L. S. Vedula, Y. Zhao, R. M. Coates, T. Koyama, D. E. Cane, D. W. Christianson, *Arch. Biochem. Biophys.* **2007**, *466*, 260–266.

[229] K. Zhou, R. J. Peters, *Phytochemistry* **2009**, *70*, 366–369.

[230] C. Iannuzzi, M. Adrover, R. Puglisi, R. Yan, P. A. Temussi, A. Pastore, *Protein Sci.* **2014**, *23*, 1208–1219.

[231] T. G. Köllner, J. Gershenzon, J. Degenhardt, *Phytochemistry* **2009**, *70*, 1139–1145.

[232] J. Rinkel, J. S. Dickschat, *Beilstein J. Org. Chem.* **2019**, *15*, 1008–1019.

[233] D. E. Cane, M. Tandon, *J. Am. Chem. Soc.* **1995**, *117*, 5602–5603.

[234] J. Rinkel, P. Rabe, X. Chen, T. G. Köllner, F. Chen, J. S. Dickschat, *Chem. - A Eur. J.* **2017**, *23*, 10501–10505.

[235] P. Baer, P. Rabe, C. A. Citron, C. C. De Oliveira Mann, N. Kaufmann, M. Groll, J. S. Dickschat, *ChemBioChem* **2014**, *15*, 213–216.

[236] V. Gonzalez, S. Touchet, D. J. Grundy, J. A. Faraldos, R. K. Allemann, *J. Am. Chem. Soc.* **2014**, *136*, 14505–14512.

[237] D. J. Newman, G. M. Cragg, *J. Nat. Prod.* **2020**, *83*, 770–803.

[238] H. B. Bode, R. Müller, *Angew. Chemie - Int. Ed.* **2005**, *44*, 6828–6846.

[239] H. G. Floss, *J. Biotechnol.* **2006**, *124*, 242–257.

[240] D. A. Hopwood, F. Malpartida, H. M. Kieser, H. Ikeda, J. Duncant, I. Fujiii, B. A. M. Rudd, H. G. Floss, S. Omura, *Nature* **1985**, *314*, 642–644.

[241] H. Sun, Z. Liu, H. Zhao, E. L. Ang, *Drug Des. Devel. Ther.* **2015**, *9*, 823–833.

[242] K. Bhattacharjee, N. R. Palepu, K. M. Rao, S. R. Joshi, *3 Biotech* **2018**, *8*, 31.

[243] K. Scherlach, D. Boettger, N. Remme, C. Hertweck, *Nat. Prod. Rep.* **2010**, *27*, 869–886.

[244] C. Wang, C. Lambert, M. Hauser, A. Deuschmann, C. Zeilinger, K. Rottner, T. E.

B. Stradal, M. Stadler, E. J. Skellam, R. J. Cox, *Chem. – A Eur. J.* **2020**, *26*, 13578–13583.

[245] R. Huisgen, *Proc. Chem. Soc.* **1961**, 357–396.

[246] K. A. J. Bozhüyük, F. Fleischhacker, A. Linck, F. Wesche, A. Tietze, C. P. Niesert, H. B. Bode, *Nat. Chem.* **2018**, *10*, 275–281.

[247] A. S. Eustáquio, D. O´Hagan, B. S. Moore, *J. Nat. Prod.* **2010**, *73*, 378–382.

[248] K. Tsukada, S. Shinki, A. Kaneko, K. Murakami, K. Irie, M. Murai, H. Miyoshi, S. Dan, K. Kawaji, H. Hayashi, et al., *Nat. Commun.* **2020**, *11*, 1830.

[249] H. Kikuchi, T. Hoshi, M. Kitayama, M. Sekiya, Y. Katou, K. Ueda, Y. Kubohara, H. Sato, M. Shimazu, S. Kurata, et al., *Tetrahedron* **2009**, *65*, 469–477.

[250] J. C. Lee, E. Lobkovsky, N. B. Pliam, G. Strobel, J. Clardy, *J. Org. Chem.* **1995**, *60*, 7076–7077.

[251] Y. He, B. Wang, W. Chen, R. J. Cox, J. He, F. Chen, *Biotechnol. Adv.* **2018**, *36*, 739–783.

[252] S. Hoefgen, J. Lin, J. Fricke, M. C. Stroe, D. J. Mattern, J. E. Kufs, P. Hortschansky, A. A. Brakhage, D. Hoffmeister, V. Valiante, *Metab. Eng.* **2018**, *48*, 44–51.

[253] C. J. Paddon, P. J. Westfall, D. J. Pitera, K. Benjamin, K. Fisher, D. McPhee, M. D. Leavell, A. Tai, A. Main, D. Eng, et al., *Nature* **2013**, *496*, 528–532.

[254] K. Zhou, K. Qiao, S. Edgar, G. Stephanopoulos, *Nat. Biotechnol.* **2015**, *33*, 377–383.

[255] C. M. Lazarus, K. Williams, A. M. Bailey, *Nat. Prod. Rep.* **2014**, *31*, 1339–1347.

[256] V. Ter-Hovhannisyan, A. Lomsadze, Y. O. Chernoff, M. Borodovsky, *Genome Res.* **2008**, *18*, 1979–1990.

[257] C. Y. Bemis, C. N. Ungarean, A. S. Shved, C. S. Jamieson, T. Hwang, K. S. Lee, K. N. Houk, D. Sarlah, *J. Am. Chem. Soc* **2021**, *143*, 6006–6017.

[258] D. P. Overy, F. Berrue, H. Correa, N. Hanif, K. Hay, M. Lanteigne, K. Mquilian, S. Duffy, P. Boland, R. Jagannathan, et al., *Mycology* **2014**, *5*, 130–144.

[259] D. H. Scharf, P. Chankhamjon, K. Scherlach, J. Dworschak, T. Heinekamp, M. Roth, A. A. Brakhage, C. Hertweck, *ChemBioChem* **2021**, *22*, 336–339.

[260] T. Saruwatari, F. Yagishita, T. Mino, H. Noguchi, K. Hotta, K. Watanabe, *ChemBioChem* **2014**, *15*, 656–659.

[261] K. I. Kusumoto, M. Matsushita-Morita, I. Furukawa, S. Suzuki, Y. Yamagata, Y. Koide, H. Ishida, M. Takeuchi, Y. Kashiwagi, *J. Appl. Microbiol.* **2008**, *105*, 1711–1719.

[262] R. Fujii, A. Minami, K. Gomi, H. Oikawa, *Tetrahedron Lett.* **2013**, *54*, 2999–3002.

[263] R. Fujii, T. Ugai, H. Ichinose, M. Hatakeyama, T. Kosaki, K. Gomi, I. Fujii, A. Minami, H. Oikawa, *Biosci. Biotechnol. Biochem.* **2016**, *80*, 426–431.

[264] H. Zhang, V. Hantke, P. Bruhnke, E. Skellam, R. J. Cox, *Chem. - A Eur. J.* **2021**, *27*, 3106–3113.

[265] F. Imeri, K. A. Nolan, A. M. Bapst, S. Santambrogio, I. Abreu-Rodríguez, P. Spielmann, S. Pfundstein, S. Libertini, L. Crowther, I. M. C. Orlando, et al., *Kidney Int.* **2019**, *95*, 375–387.

[266] R. D. Gietz, R. A. Woods, *Methods Enzymol.* **2002**, *350*, 87–96.

[267] H. Jeong, H. J. Kim, S. J. Lee, *Genome Announc.* **2015**, *3*, e00134-15.

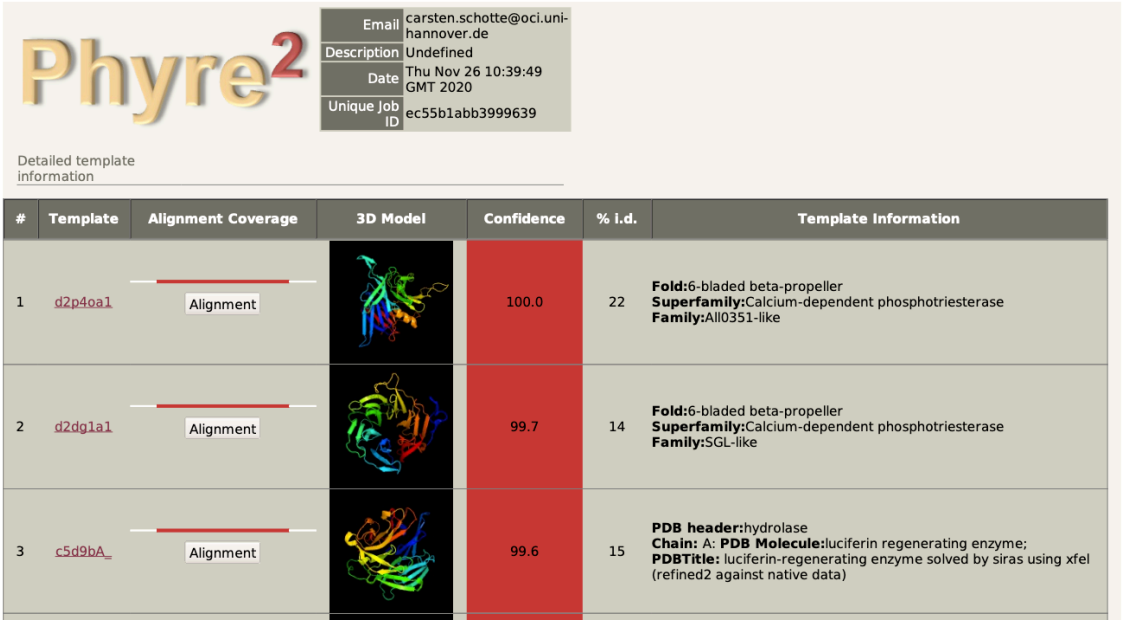
[268] K. Graumann, A. Premstaller, *Biotechnol. J.* **2006**, *1*, 164–186.

[269] D. S. Waugh, *Trends Biotechnol.* **2005**, *23*, 316–320.

[270] T. D. Parks, K. K. Leuther, E. D. Howard, S. A. Johnston, W. G. Dougherty, *Anal. Biochem.* **1994**, *216*, 413–417.

# 8. Appendix

## 8.1. Xenovulene-Type Tropolone Sesquiterpenoids



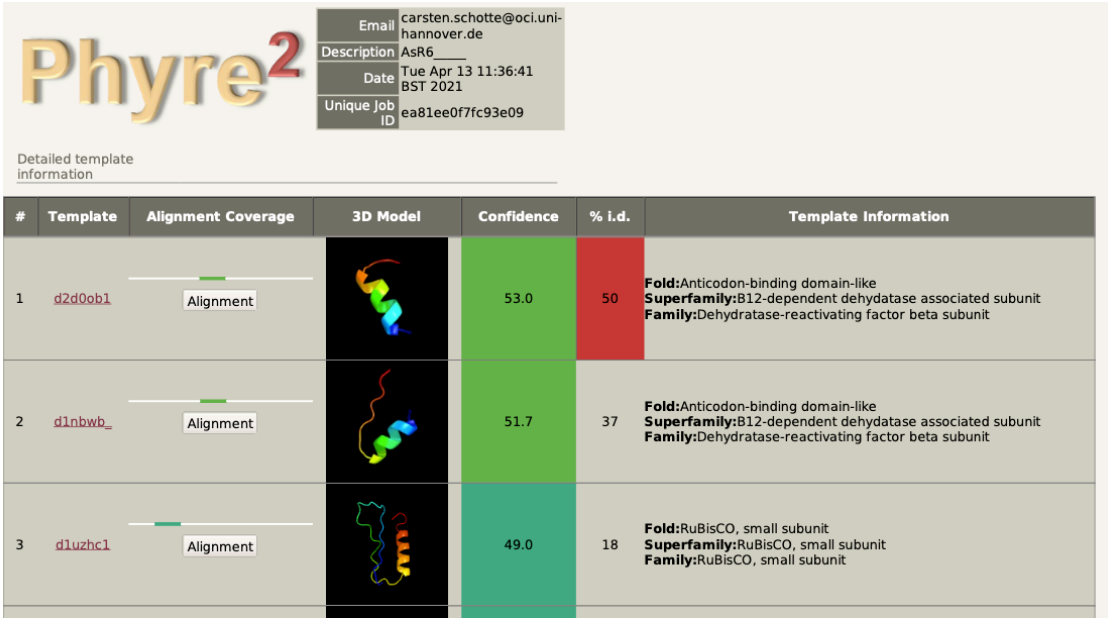
Phyre2

| | |
|---|---|
| Email | carsten.schotte@oci.uni-hannover.de |
| Description | Undefined |
| Date | Thu Nov 26 10:39:49 GMT 2020 |
| Unique Job ID | ec55b1abb3999639 |

Detailed template information

| # | Template | Alignment Coverage | 3D Model | Confidence | % i.d. | Template Information |
|---|---|---|---|---|---|---|
| 1 | d2p4oa1 | Alignment | | 100.0 | 22 | **Fold:**6-bladed beta-propeller<br>**Superfamily:**Calcium-dependent phosphotriesterase<br>**Family:**All0351-like |
| 2 | d2dg1a1 | Alignment | | 99.7 | 14 | **Fold:**6-bladed beta-propeller<br>**Superfamily:**Calcium-dependent phosphotriesterase<br>**Family:**SGL-like |
| 3 | c5d9bA_ | Alignment | | 99.6 | 15 | **PDB header:**hydrolase<br>**Chain:** A: **PDB Molecule:**luciferin regenerating enzyme;<br>**PDBTitle:** luciferin-regenerating enzyme solved by siras using xfel (refined2 against native data) |

**Figure 8.1** Phyre2 analysis of AsR5.



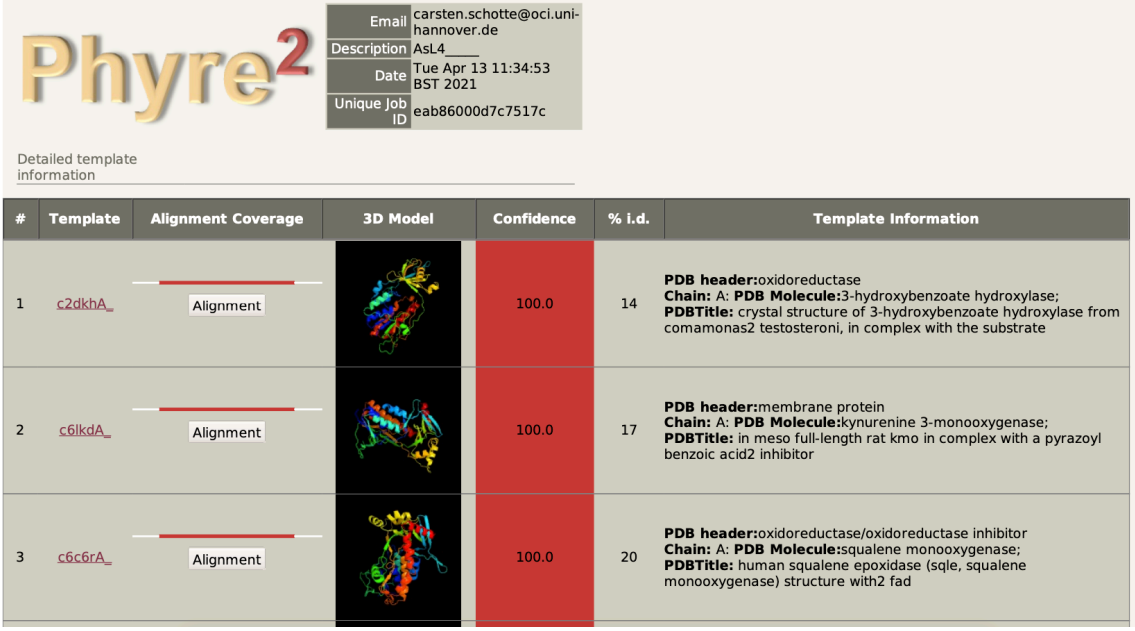
Phyre2

| | |
|---|---|
| Email | carsten.schotte@oci.uni-hannover.de |
| Description | AsR6____ |
| Date | Tue Apr 13 11:36:41 BST 2021 |
| Unique Job ID | ea81ee0f7fc93e09 |

Detailed template information

| # | Template | Alignment Coverage | 3D Model | Confidence | % i.d. | Template Information |
|---|---|---|---|---|---|---|
| 1 | d2d0ob1 | Alignment | | 53.0 | 50 | **Fold:**Anticodon-binding domain-like<br>**Superfamily:**B12-dependent dehydatase associated subunit<br>**Family:**Dehydratase-reactivating factor beta subunit |
| 2 | d1nbwb_ | Alignment | | 51.7 | 37 | **Fold:**Anticodon-binding domain-like<br>**Superfamily:**B12-dependent dehydatase associated subunit<br>**Family:**Dehydratase-reactivating factor beta subunit |
| 3 | d1uzhc1 | Alignment | | 49.0 | 18 | **Fold:**RuBisCO, small subunit<br>**Superfamily:**RuBisCO, small subunit<br>**Family:**RuBisCO, small subunit |

**Figure 8.2** Phyre2 analysis of AsR6.



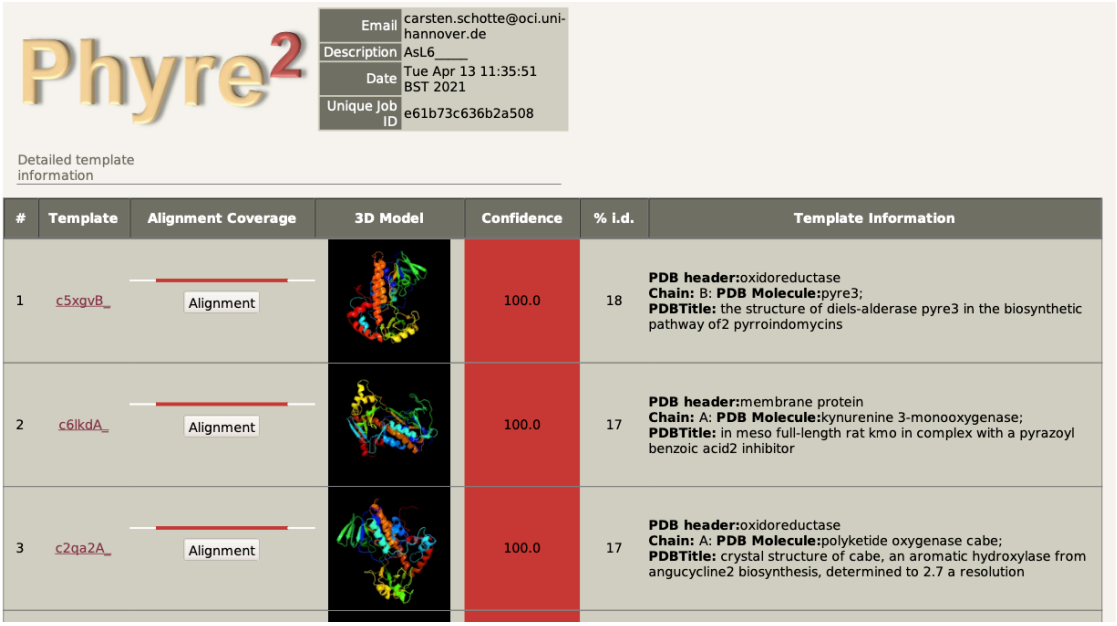
Phyre2

| | |
|---|---|
| Email | carsten.schotte@oci.uni-hannover.de |
| Description | AsL4____ |
| Date | Tue Apr 13 11:34:53 BST 2021 |
| Unique Job ID | eab86000d7c7517c |

Detailed template information

| # | Template | Alignment Coverage | 3D Model | Confidence | % i.d. | Template Information |
|---|---|---|---|---|---|---|
| 1 | c2dkhA_ | Alignment | | 100.0 | 14 | **PDB header:**oxidoreductase **Chain:** A: **PDB Molecule:**3-hydroxybenzoate hydroxylase; **PDBTitle:** crystal structure of 3-hydroxybenzoate hydroxylase from comamonas2 testosteroni, in complex with the substrate |
| 2 | c6lkdA_ | Alignment | | 100.0 | 17 | **PDB header:**membrane protein **Chain:** A: **PDB Molecule:**kynurenine 3-monooxygenase; **PDBTitle:** in meso full-length rat kmo in complex with a pyrazoyl benzoic acid2 inhibitor |
| 3 | c6c6rA_ | Alignment | | 100.0 | 20 | **PDB header:**oxidoreductase/oxidoreductase inhibitor **Chain:** A: **PDB Molecule:**squalene monooxygenase; **PDBTitle:** human squalene epoxidase (sqle, squalene monooxygenase) structure with2 fad |

**Figure 8.3** Phyre2 analysis of AsL4.



Phyre2

| | |
|---|---|
| Email | carsten.schotte@oci.uni-hannover.de |
| Description | AsL6____ |
| Date | Tue Apr 13 11:35:51 BST 2021 |
| Unique Job ID | e61b73c636b2a508 |

Detailed template information

| # | Template | Alignment Coverage | 3D Model | Confidence | % i.d. | Template Information |
|---|---|---|---|---|---|---|
| 1 | c5xgvB_ | Alignment | | 100.0 | 18 | **PDB header:**oxidoreductase **Chain:** B: **PDB Molecule:**pyre3; **PDBTitle:** the structure of diels-alderase pyre3 in the biosynthetic pathway of2 pyrroindomycins |
| 2 | c6lkdA_ | Alignment | | 100.0 | 17 | **PDB header:**membrane protein **Chain:** A: **PDB Molecule:**kynurenine 3-monooxygenase; **PDBTitle:** in meso full-length rat kmo in complex with a pyrazoyl benzoic acid2 inhibitor |
| 3 | c2qa2A_ | Alignment | | 100.0 | 17 | **PDB header:**oxidoreductase **Chain:** A: **PDB Molecule:**polyketide oxygenase cabe; **PDBTitle:** crystal structure of cabe, an aromatic hydroxylase from angucycline2 biosynthesis, determined to 2.7 a resolution |

**Figure 8.4** Phyre2 analysis of AsL6.

**80.93% coverage, 28 unique peptides (highlighted in red):**

MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDHPFTMTETSLKMMWEGLKEYFGDGFAPGSAPF
RNNVHYCDLHPAPKRDANFRYGVEVEEEAMPLFSVLGDTCAPPCTCFEVTKLVEHIDSFIATAS
DYHADDYAITSGKNDATIDQVCFYAMRDTLSWWVHWGGGLRPSADYWKHVYVAFAAVPDDFQLP
PRDFLSGQYTLLGHTWDDCREGLLREGVPPETVKMAEMALWRQMLTQYLEKVDPTLRSLLVART
SMMTLYRVQTANTLGCAALLLSSKGVNVNGLEDGALEVASIAQCVFLDISKEAMGILKGEKTET
ITGNRAQMKREMLWVYTRCMDYLGSQPNADFLRRYASAGMHYVPMMDRYRERMNGHSRFAIPPA
MARILLPYVKNPSALSGVLPIEDTTTGVSIHEAVNFLPSANATVVS

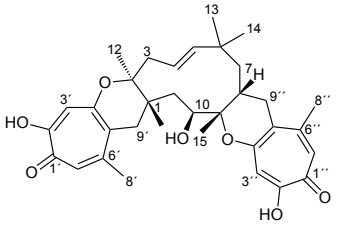
**Figure 8.5** Results protein mass spectrometry $His_6$-PycR6.

## 8.2. Eupenifeldin-Type Tropolone Sesquiterpenoids

AGTGTTATGTTGCTTAGTGGACGCTGATATAGCTGGCGAGAACCGCAAATTGTGCTGCGCTTCAAACCAATACACCGG
CTGCCAATGAATTTAAGGCGAGTCCAGACACCAGGGAAAGGACAAACACCCAACACCAAGCAAAGCTTGAGGGTACAA
ATGACGCTCGAACAAGCATGCCCCATGGAATACCAAGGGGCGCAATGTGCGTTCAAAGATTCGATGATTCACTGAATT
CTGCAATTCACACTACTTATCGCATTTCGCTGCGTTCTTCATCGATGCCAGAACCAAGAGATCCGTTGTTGAAAGTTG
TAATTATTATAGTTATTCAGACGCTGATTGAAAATTAAAAGGTTGTTGAGTTGTCCGACCGGCAGGCAAGCCCACCGA
GGAAACAAGAGTGCTCAAAAAGACAAGGGTTCAGACAGAGCGCGTACTCGCAAACCAGCTGCCGAAGCACGCTGGAAA
C

**Figure 8.6** ITS sequence of *Phaeosphaeriaceae* sp. CF-150626.

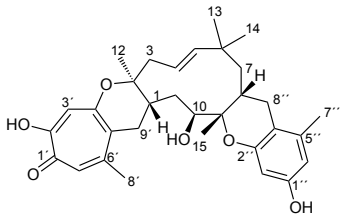
**Table 8.1** Summarized NMR data for eupenifeldin **5** labelled with [2-$^{13}$C]-acetate and non-labelled (natural abundance). norm = normalized; PE = peak enhancement; incorporation events highlighted in green.




eupenifeldin **5**
(548 g·mol$^{-1}$)

| Eupenifeldin (natural abundance) **5** | | | | Eupenifeldin (labeled) **5** | | | |
|---|---|---|---|---|---|---|---|
| **Pos.** | **$\delta_C$** | **Intensity** | **norm. to C-3´´** | **$\delta_C$ type** | **Intensity** | **norm. to C-3´´** | **PE** |
| **1** | 41.6 | 177.9 | 0.07 | 41.5 | 160.35 | 0.53 | 7.35 |
| **2** | 80.5 | 3009.01 | 1.22 | 80.5 | 255.69 | 0.85 | 0.69 |
| **3** | 46.3 | 1290.71 | 0.52 | 46.3 | 1183.15 | 3.92 | 7.48 |
| **4** | 125.9 | 429.39 | 0.17 | 125.9 | 69.49 | 0.23 | 1.32 |
| **5** | 144.1 | 2023.44 | 0.82 | 144.1 | 1808.45 | 5.99 | 7.29 |
| **6** | 35.0 | 1760.37 | 0.72 | 35.0 | 175.04 | 0.58 | 0.81 |
| **7** | 46.6 | 60625 | 0.25 | 46.6 | 84.86 | 0.28 | 1.14 |
| **8** | 32.0 | 1775.99 | 0.72 | 32.0 | 1285.28 | 4.26 | 5.91 |
| **9** | 82.0 | 4204.58 | 1.71 | 82.0 | 384.33 | 1.27 | 0.75 |
| **10** | 70.8 | 1839.77 | 0.75 | 70.8 | 1650.43 | 5.47 | 7.32 |
| **11** | 30.2 | 388.89 | 0.16 | 30.3 | 59.39 | 0.20 | 1.25 |
| **12** | 19.4 | 3294.09 | 1.34 | 19.4 | 2666.92 | 8.84 | 6.61 |
| **13** | 29.7 | 1840.71 | 0.75 | 29.7 | 1553.96 | 5.15 | 6.89 |
| **14** | 27.3 | 415.84 | 0.17 | 27.2 | 423.77 | 1.40 | 8.32 |
| **15** | 16.1 | 1620.74 | 0.66 | 16.1 | 1400.98 | 4.64 | 7.05 |
| **1´** | 172.6 | 2288.98 | 0.93 | 172.7 | 232.18 | 0.77 | 0.83 |
| **2´** | 163.2 | 2470.32 | 1.00 | 163.1 | 2178.66 | 7.22 | 7.20 |
| **3´** | 113.5 | 2709.01 | 1.10 | 113.4 | 315.99 | 1.05 | 0.95 |
| **4´** | 160.3 | 2369.71 | 0.96 | 160.2 | 199.32 | 0.66 | 0.69 |
| **5´** | 122.5 | 824.15 | 0.33 | 122.3 | 647.60 | 2.15 | 6.41 |
| **6´** | 150.5 | 2521.18 | 1.02 | 150.5 | 228.95 | 0.76 | 0.74 |
| **7´** | 124.7 | 2767.78 | 1.12 | 124.8 | 2116.90 | 7.02 | 6.24 |
| **8´** | 27.4 | 4551.74 | 1.85 | 27.5 | 2978.82 | 9.87 | 5.34 |
| **9´** | 33.0 | 1714.44 | 0.70 | 33.1 | 180.96 | 0.60 | 0.86 |
| **1´´** | 173.3 | 2199.75 | 0.89 | 173.4 | 230.52 | 0.76 | 0.86 |
| **2´´** | 162.9 | 2045.41 | 0.83 | 162.9 | 2153.22 | 7.14 | 8.59 |
| **3´´** | 112.8 | 2461.56 | 1 | 112.7 | 301.66 | 1 | 1 |
| **4´´** | 159.5 | 2288.51 | 0.93 | 159.4 | 180.55 | 0.60 | 0.64 |
| **5´´** | 118.7 | 2301.47 | 0.93 | 118.6 | 1633.59 | 5.42 | 5.79 |
| **6´´** | 151.5 | 2774.16 | 1.13 | 151.4 | 256.43 | 0.85 | 0.75 |
| **7´´** | 125.6 | 2102.25 | 0.85 | 125.7 | 1678.03 | 5.56 | 6.51 |
| **8´´** | 27.5 | 4567.07 | 1.86 | 27.5 | 3177.05 | 10.53 | 5.68 |
| **9´´** | 34.4 | 2074.85 | 0.84 | 34.4 | 218.88 | 0.73 | 0.86 |

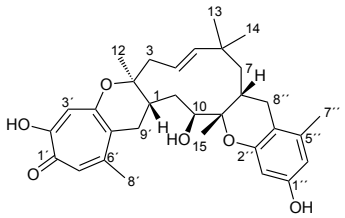
**Table 8.2** Summarized NMR data for noreupenifeldin B **6** labelled with [1-$^{13}$C]-acetate and natural abundance. norm = normalized; PE = peak enhancement; incorporation events highlighted in green; carbon atoms in red could not be distinguished.




noreupenifeldin B **6**
(520 g·mol$^{-1}$)

| noreupenifeldin **6** (natural abundance) | | | | noreupenifeldin **6** (labelled) | | | |
|---|---|---|---|---|---|---|---|
| **Pos.** | **$\delta_C$** | **Intensity** | **norm. to C-3´** | **$\delta_C$ type** | **Intensity** | **norm. to C-3´** | **PE** |
| **1** | 41.6 | 0.06 | 0.05 | 41.5 | 0.07 | 0.04 | 0.77 |
| **2** | 80.9 | 1.42 | 1.17 | 80.9 | 9.12 | 5.16 | 4.42 |
| **3** | 46.3 | 0.76 | 0.63 | 46.3 | 2.30 | 1.30 | 2.08 |
| **4** | 125.3 | 0.18 | 0.15 | 125.2 | 1.58 | 0.89 | 6.12 |
| **5** | 144.6 | 0.99 | 0.81 | 144.7 | 1.30 | 0.74 | 0.90 |
| **6** | 35.1 | 0.75 | 0.61 | 35.0 | 5.77 | 3.26 | 5.33 |
| **7** | 46.3 | 0.76 | 0.63 | 46.3 | 2.30 | 1.30 | 2.08 |
| **8** | 32.4 | 0.63 | 0.51 | 32.4 | 0.80 | 0.45 | 0.88 |
| **9** | 80.6 | 1.71 | 1.41 | 80.5 | 10.46 | 5.92 | 4.21 |
| **10** | 71.0 | 0.69 | 0.56 | 71.0 | 0.76 | 0.43 | 0.76 |
| **11** | 30.4 | 0.11 | 0.091 | 30.4 | 1.08 | 0.61 | 6.86 |
| **12** | 19.3 | 1.61 | 1.32 | 19.3 | 2.25 | 1.27 | 0.96 |
| **13** | 29.8 | 0.86 | 0.70 | 29.8 | 1.48 | 0.84 | 1.19 |
| **14** | 27.2 | 0.16 | 0.13 | 27.2 | 0.18 | 0.10 | 0.78 |
| **15** | 15.8 | 0.73 | 0.60 | 15.8 | 1.09 | 0.62 | 1.03 |
| **1´** | 172.4 | 0.98 | 0.81 | 172.2 | 3.30 | 1.86 | 2.31 |
| **2´** | 163.3 | 1.24 | 1.02 | 163.3 | 0.67 | 0.38 | 0.37 |
| **3´** | 113.7 | 1.22 | 1 | 113.9 | 1.77 | 1 | 1 |
| **4´** | 160.6 | 1.19 | 0.98 | 160.8 | 6.62 | 3.74 | 3.82 |
| **5´** | 122.9 | 0.32 | 0.27 | 123.2 | 0.35 | 0.20 | 0.75 |
| **6´** | 150.7 | 1.18 | 0.97 | 150.9 | 6.42 | 3.63 | 3.75 |
| **7´** | 124.7 | 1.27 | 1.04 | 124.6 | 1.23 | 0.70 | 0.67 |
| **8´** | 27.5 | 2.19 | 1.80 | 27.5 | 2.51 | 1.42 | 0.79 |
| **9´** | 33.2 | 0.71 | 0.59 | 33.2 | 5.78 | 3.27 | 5.58 |
| **1´´** | 154.9 | 1.45 | 1.19 | 155.2 | 1.29 | 0.73 | 0.62 |
| **2´´** | 101.6 | 1.31 | 1.08 | 101.6 | 2.01 | 1.14 | 1.05 |
| **3´´** | 153.7 | 1.33 | 1.09 | 153.7 | 9.36 | 5.30 | 4.86 |
| **4´´** | 111.1 | 1.04 | 0.85 | 110.9 | 1.24 | 0.70 | 0.83 |
| **5´´** | 139.6 | 1.34 | 1.10 | 139.6 | 9.07 | 5.13 | 4.65 |
| **6´´** | 110.1 | 0.97 | 0.80 | 110.2 | 1.07 | 0.61 | 0.76 |
| **7´´** | 19.6 | 2.18 | 1.79 | 19.5 | 2.77 | 1.56 | 0.87 |
| **8´´** | 29.7 | 1.03 | 0.85 | 29.7 | 7.55 | 4.27 | 5.03 |

**Table 8.3** Summarized NMR data for noreupenifeldin B **6** labelled with [2-$^{13}$C]-acetate and non-labelled (natural abundance). norm = normalized; PE = peak enhancement; incorporation events highlighted in green; carbon atoms in red could not be distinguished.



noreupenifeldin B **6**
(520 g·mol$^{-1}$)

| noreupenifeldin **6** (natural abundance) | | | | noreupenifeldin B **6** (labelled) | | | |
|---|---|---|---|---|---|---|---|
| **Pos.** | **$\delta_C$** | **Intensity** | **norm. to C-3´** | **$\delta_C$ type** | **Intensity** | **norm. to C-3´** | **PE** |
| 1 | 41.6 | 0.06 | 0.05 | 41.5 | 0.56 | 0.42 | 8.09 |
| 2 | 80.9 | 1.42 | 1.17 | 80.9 | 1.49 | 1.11 | 0.95 |
| 3 | 46.3 | 0.76 | 0.63 | 46.3 | 4.79 | 3.59 | 5.73 |
| 4 | 125.3 | 0.18 | 0.15 | 125.3 | 0.29 | 0.22 | 1.48 |
| 5 | 144.6 | 0.99 | 0.81 | 144.7 | 7.51 | 5.62 | 6.91 |
| 6 | 35.1 | 0.75 | 0.61 | 35.1 | 0.98 | 0.73 | 1.20 |
| 7 | 46.3 | 0.76 | 0.63 | 46.3 | 4.79 | 3.59 | 5.73 |
| 8 | 32.4 | 0.63 | 0.51 | 32.4 | 4.46 | 3.34 | 6.50 |
| 9 | 80.6 | 1.71 | 1.41 | 80.6 | 1.89 | 1.42 | 1.01 |
| 10 | 71.0 | 0.69 | 0.56 | 71.0 | 4.10 | 3.07 | 5.44 |
| 11 | 30.4 | 0.11 | 0.091 | 30.4 | 0.15 | 0.11 | 1.25 |
| 12 | 19.3 | 1.61 | 1.32 | 19.3 | 12.01 | 9.00 | 6.80 |
| 13 | 29.8 | 0.86 | 0.70 | 29.7 | 6.95 | 5.21 | 7.41 |
| 14 | 27.2 | 0.16 | 0.13 | 27.2 | 1.41 | 1.05 | 8.14 |
| 15 | 15.8 | 0.73 | 0.60 | 15.7 | 6.02 | 4.51 | 7.51 |
| 1´ | 172.4 | 0.98 | 0.81 | 172.4 | 1.05 | 0.79 | 0.98 |
| 2´ | 163.3 | 1.24 | 1.02 | 163.3 | 5.43 | 4.07 | 3.98 |
| 3´ | 113.7 | 1.22 | 1 | 113.8 | 1.34 | 1 | 1 |
| 4´ | 160.6 | 1.19 | 0.98 | 160.6 | 1.23 | 0.92 | 0.94 |
| 5´ | 122.9 | 0.32 | 0.27 | 123.0 | 2.53 | 1.90 | 7.16 |
| 6´ | 150.7 | 1.18 | 0.97 | 150.8 | 0.40 | 0.30 | 0.31 |
| 7´ | 124.7 | 1.27 | 1.04 | 124.7 | 7.93 | 5.94 | 5.71 |
| 8´ | 27.5 | 2.19 | 1.80 | 27.5 | 13.48 | 10.10 | 5.62 |
| 9´ | 33.2 | 0.71 | 0.59 | 33.2 | 0.88 | 0.66 | 1.12 |
| 1´´ | 154.9 | 1.45 | 1.19 | 155.1 | 7.04 | 5.27 | 4.43 |
| 2´´ | 101.6 | 1.31 | 1.08 | 101.6 | 1.48 | 1.11 | 1.02 |
| 3´´ | 153.7 | 1.33 | 1.09 | 153.7 | 1.41 | 1.06 | 0.97 |
| 4´´ | 111.1 | 1.04 | 0.85 | 111.0 | 6.07 | 4.55 | 5.34 |
| 5´´ | 139.6 | 1.34 | 1.10 | 139.6 | 1.57 | 1.17 | 1.06 |
| 6´´ | 110.1 | 0.97 | 0.80 | 110.1 | 6.02 | 4.51 | 5.67 |
| 7´´ | 19.6 | 2.18 | 1.79 | 19.5 | 12.18 | 9.12 | 5.09 |
| 8´´ | 29.7 | 1.03 | 0.85 | 29.7 | 1.33 | 0.99 | 1.17 |

## 8.3. NMR Spectra

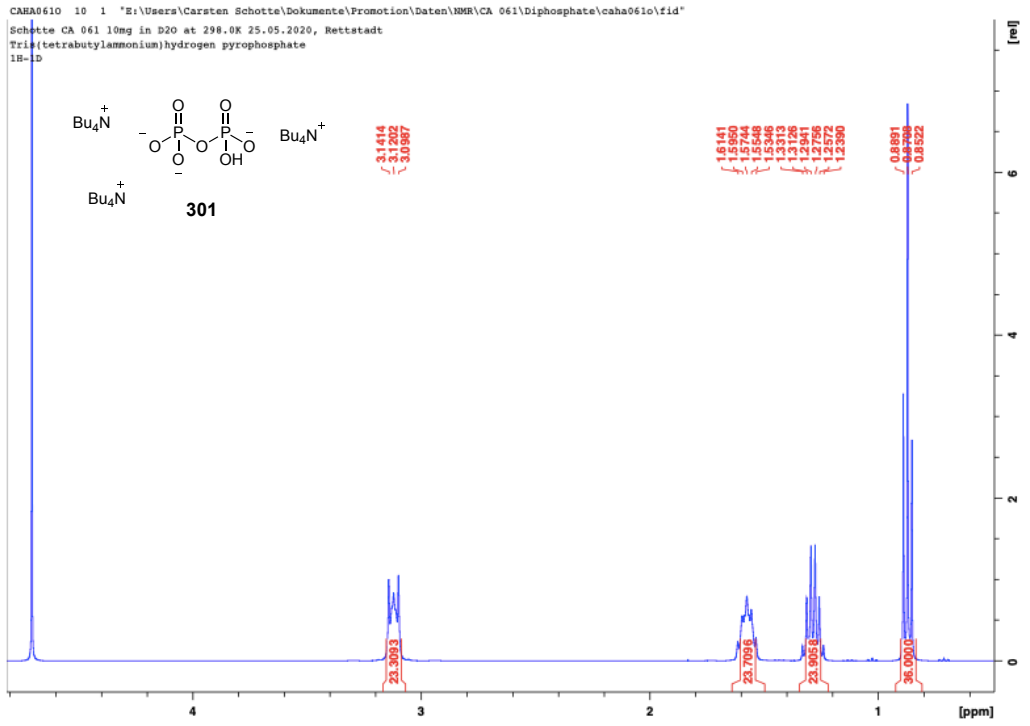



**Figure 8.7** $^{1}$H NMR spectrum of tris(tetrabutylammonium)hydrogen phosphate **301** in $D_2O$ (400 MHz).

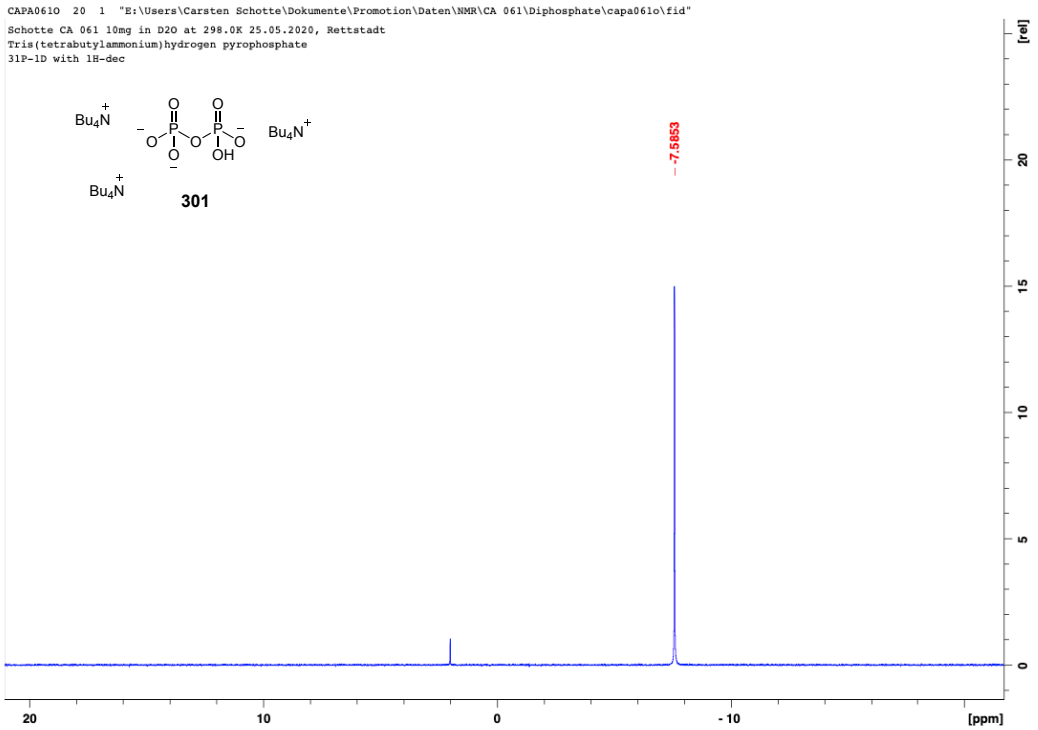



**Figure 8.8** $^{31}$P NMR spectrum of tris(tetrabutylammonium)hydrogen phosphate **301** in $D_2O$ (162 MHz).

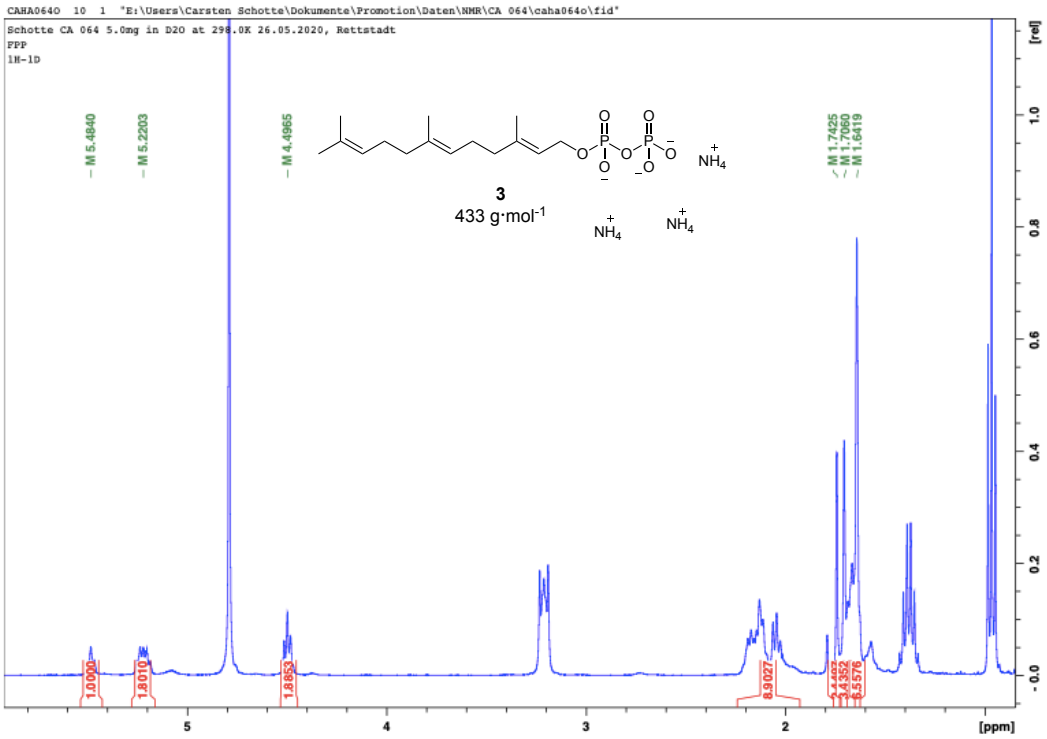



**Figure 8.9** $^{1}$H NMR spectrum of FPP **3** in $D_2O$ (400 MHz).

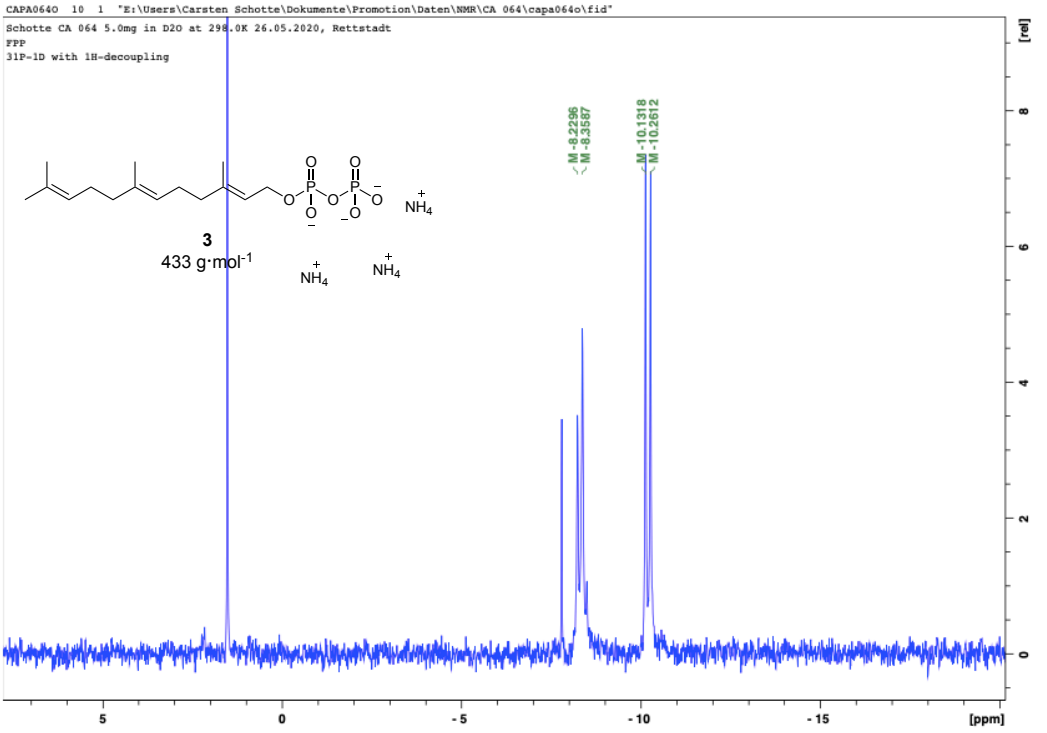



**Figure 8.10** $^{31}$P NMR spectrum of FPP **3** in $D_2O$ (162 MHz).

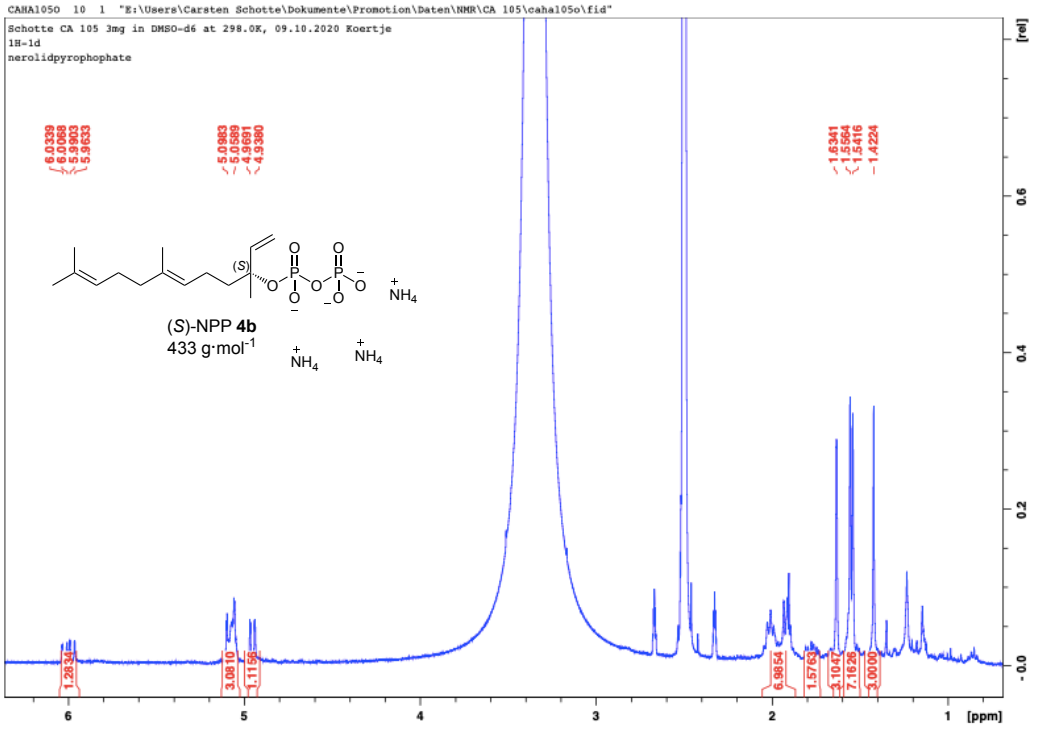



**Figure 8.11** $^{1}$H NMR spectrum of (*S*)-nerolidyl pyrophosphate **4b** in DMSO-$d_6$ (400 MHz).

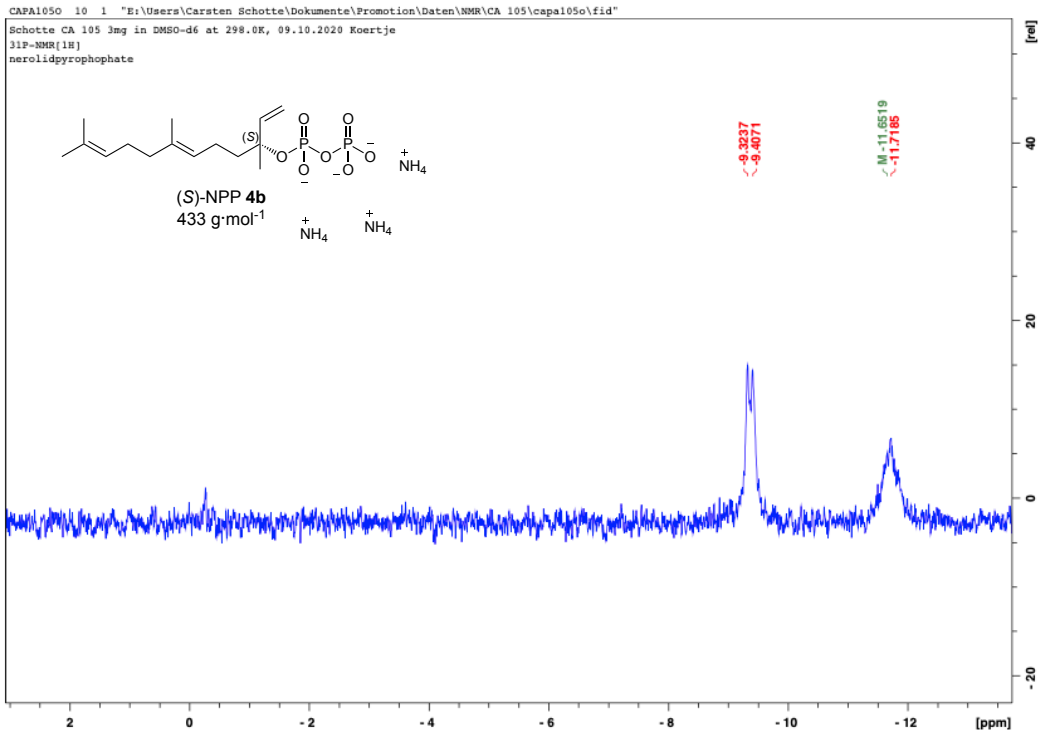



**Figure 8.12** $^{31}$P NMR spectrum of (*S*)-nerolidyl pyrophosphate **4b** in DMSO-$d_6$ (162 MHz).

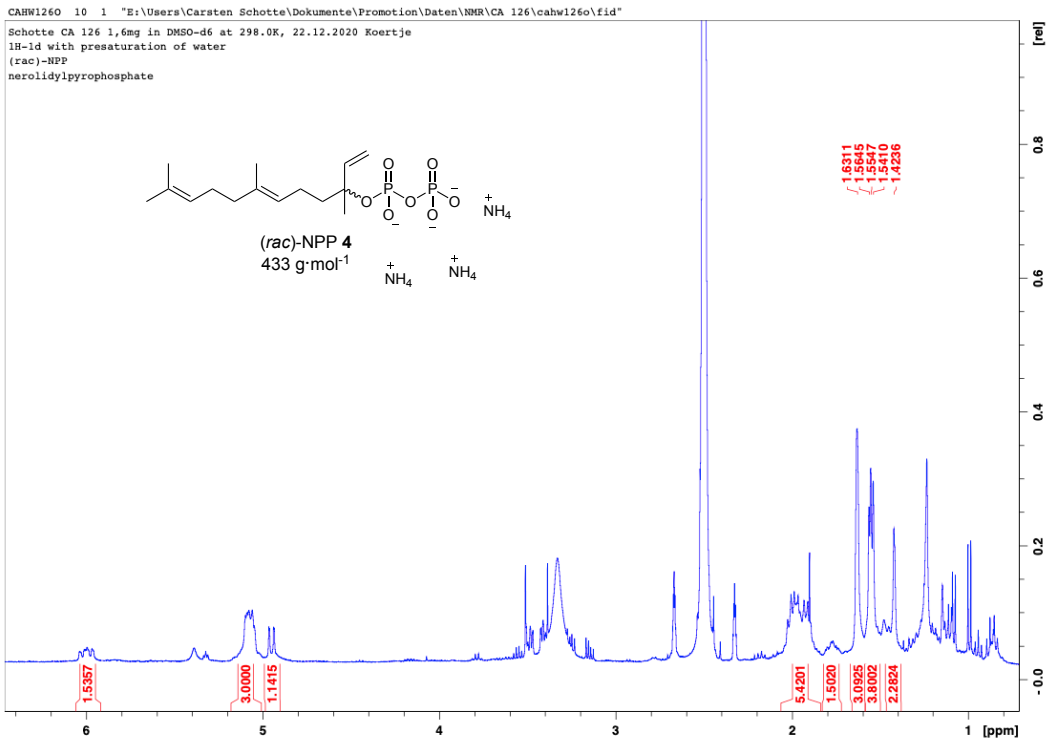



**Figure 8.13** $^{1}$H NMR spectrum of (*rac*)-nerolidyl pyrophosphate **4** in DMSO-$d_6$ (400 MHz).

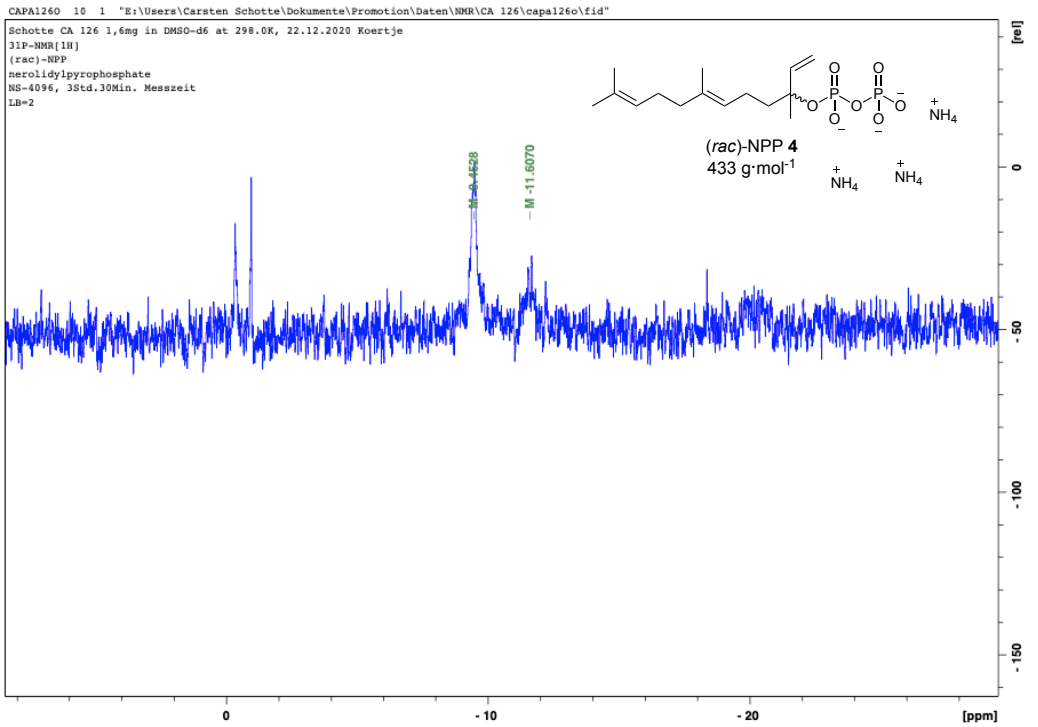



**Figure 8.14** $^{31}$P NMR spectrum of (*rac*)-nerolidyl pyrophosphate **4** in DMSO-$d_6$ (162 MHz).

## 8.4. Overview of deployed Labbook ID´s

To link key experiments to the corresponding labbook entries and to the corresponding raw data labbook identifiers (Labbook-IDs) were used throughout the thesis. Compare table 8.4 for detailed information regarding each labbook-ID. Detailed information on the construction of indicated plasmids and indicated transformation experiments are also given in the experimental section of this thesis.

**Table 8.4** Description of labbook-ID´s used in this thesis. Labbook-ID´s are ordered according to first appearance.

| Labbook-ID | Description of conducted experiments | Description of created plasmids/obtained transformants |
|---|---|---|
| CS I 01 J | Labbook ID "CS I 01 J" refers to the heterologous expression of hetero-Diels-Alderase encoding gene *asR5.* | Plasmid CS I 01 J was created in the scope of this thesis. It harbors the *asR5* gene sequence inserted into a pET32 vector backbone. Correct cloning was confirmed by DNA sequencing. |
| CS I 78 1-13 | Labbok ID "CS I 78 1-13" refers to the creation of plasmid CS I 78 1-13. | Plasmid CS I 78 1-13 was created in the scope of this thesis. It harbors the *eupR5* gene sequence inserted into a pET28a vector backbone. Correct cloning was confirmed by DNA sequencing. |
| CS I 135 1-7 | Labbok ID "CS I 135 1-7" refers to the creation of plasmid CS I 135 1-7. | Plasmid CS I 135 1-7 was created in the scope of this thesis. It harbors the codon-optimized gene sequence of *asR6* inserted into a pETM-11 vector backbone. Correct cloning was confirmed by DNA sequencing. |
| CS I 112 | Labbook ID "CS I 112" refers to the heterologous expression of *asPKS1, asL1, asL3, asR2, asR5, asR6* and *eupR6* in *A. oryzae* NSAR1. | Transformants carrying the indicated set of genes were confirmed by PCR and analyzed for metabolite production by LCMS. |
| CS I 119 | Labbook ID "CS I 119" refers to the heterologous expression of *asPKS1, asL1, asL3, asL4, asL6, asR2, asR5, asR6* and *eupR6* in *A. oryzae* NSAR1. | Transformants carrying the indicated set of genes were confirmed by PCR and analyzed for metabolite production by LCMS. |
| CS I 134 | Labbook ID "CS I 134" refers to the heterologous expression of *asPKS1, asL1, asL3, asR2, asR5, asR6* and *eupR5* in *A. oryzae* NSAR1. | Transformants carrying the indicated set of genes were confirmed by PCR and analyzed for metabolite production by LCMS. |
| CS I 136 | Labbook ID "CS I 136" refers to the heterologous expression of *asPKS1, asL1, asL3, asR2, asR5* and *eupR3* in *A. oryzae* NSAR1. | Transformants carrying the indicated set of genes were confirmed by PCR and analyzed for metabolite production by LCMS. |
| CS I 114 | Labbook ID "CS I 114" refers to the heterologous expression of *asPKS1, asL1, asL3, asR5, asR6* and *eupL4* in *A. oryzae* NSAR1. | Transformants carrying the indicated set of genes were confirmed by PCR and analyzed for metabolite production by LCMS. |
| CS I 130 | Labbook ID "CS I 130" refers to the heterologous expression of *asPKS1, asL1, asL3, asR5, asR6, eupL4* and *eupR6* in *A. oryzae* NSAR1. | Transformants carrying the indicated set of genes were confirmed by PCR and analyzed for metabolite production by LCMS. |
| CS I 92 | Labbook ID "CS I 92" refers to the heterologous expression of *asPKS1, asL1, asL3, eupR1* and *eupR3* in *A. oryzae* NSAR1. | Transformants carrying the indicated set of genes were confirmed by PCR and analyzed for metabolite production by LCMS. |

| | | |
|---|---|---|
| CS I 96 | Labbook ID "CS I 96" refers to the heterologous expression of *asPKS1, asL1, asL3, eupL4, eupR1* and *eupR3* in *A. oryzae* NSAR1. | Transformants carrying the indicated set of genes were confirmed by PCR and analyzed for metabolite production by LCMS. |
| CS I 104 | Labbook ID "CS I 104" refers to the heterologous expression of *asPKS1, asL1, asL3, eupL4, eupR1, eupR3, eupR5* and *eupR6* in *A. oryzae* NSAR1. | Transformants carrying the indicated set of genes were confirmed by PCR and analyzed for metabolite production by LCMS. |
| CS I 62 | Labbook ID "CS I 62" refers to the heterologous expression of *asPKS1, asL1, asL3, pycR1* and *pycR6* in *A. oryzae* NSAR1. | Transformants carrying the indicated set of genes were confirmed by PCR and analyzed for metabolite production by LCMS. |
| CS I 03 I | Labbok ID "CS I 03 I" refers to the creation of plasmid CS I 03 I. | Plasmid CS I 03 I was created in the scope of this thesis. It harbors the *asR5* and *asR6* gene sequences inserted into a pTYGS-*ade* vector backbone. Correct cloning was confirmed by DNA sequencing. |
| CS I 82 7-1 | Labbok ID "CS I 82 7-1" refers to the creation of plasmid CS I 82 7-1. | Plasmid CS I 82 7-1 was created in the scope of this thesis. It harbors the *eupR6* gene sequence inserted into a pTYGS-*met* vector backbone. Correct cloning was confirmed by DNA sequencing. |
| CS I 106 1-2 | Labbok ID "CS I 106 1-2" refers to the creation of plasmid CS I 106 1-2. | Plasmid CS I 106 1-2 was created in the scope of this thesis. It harbors the *eupR6, asL4* and *asL6* gene sequences inserted into a pTYGS-*met* vector backbone. Correct cloning was confirmed by DNA sequencing. |
| CS I 80 1-1 | Labbok ID "CS I 80 1-1" refers to the creation of plasmid CS I 80 1-1. | Plasmid CS I 80 1-1 was created in the scope of this thesis. It harbors the *eupR5* gene sequence inserted into a pTYGS-*met* vector backbone. Correct cloning was confirmed by DNA sequencing. |
| CS I 128 2-1 | Labbok ID "CS I 128 2-1" refers to the creation of plasmid CS I 128 2-1. | Plasmid CS I 128 2-1 was created in the scope of this thesis. It harbors the *eupR5* and *asR5* gene sequences inserted into a pTYGS-*ade* vector backbone. Correct cloning was confirmed by DNA sequencing. |
| CS I 110 1-5 | Labbok ID "CS I 110 1-5" refers to the creation of plasmid CS I 110 1-5. | Plasmid CS I 110 1-5 was created in the scope of this thesis. It harbors the *eupL4* gene sequence inserted into a pTYGS-*met* vector backbone. Correct cloning was confirmed by DNA sequencing. |
| CS I 125 1-1 | Labbok ID "CS I 125 1-1" refers to the creation of plasmid CS I 125 1-1. | Plasmid CS I 125 1-1 was created in the scope of this thesis. It harbors the *eupL4* and *eupR6* gene sequences inserted into a pTYGS-*met* vector backbone. Correct cloning was confirmed by DNA sequencing. |
| CS I 75 3-2 | Labbok ID "CS I 75 3-2" refers to the creation of plasmid CS I 75 3-2. | Plasmid CS I 75 3-2 was created in the scope of this thesis. It harbors the *eupR1* and *eupR3* gene sequences inserted into a pTYGS-*ade* vector backbone. Correct cloning was confirmed by DNA sequencing. |
| CS I 93 2-4 | Labbok ID "CS I 93 2-4" refers to the creation of plasmid CS I 93 2-4. | Plasmid CS I 93 2-4 was created in the scope of this thesis. It harbors the *eupL4, eupR1* and *eupR3* gene sequences inserted into a pTYGS-*ade* vector backbone. Correct cloning was confirmed by DNA sequencing. |

| | | |
|---|---|---|
| CS I 101 1-1 | Labbok ID "CS I 101 1-1" refers to the creation of plasmid CS I 101 1-1. | Plasmid CS I 101 1-1 was created in the scope of this thesis. It harbors the *eupR5* and *eupR6* gene sequences inserted into a pTYGS-*met* vector backbone. Correct cloning was confirmed by DNA sequencing. |
| CS I 58 1-2 | Labbok ID "CS I 58 1-2" refers to the creation of plasmid CS I 58 1-2. | Plasmid CS I 58 1-2 was created in the scope of this thesis. It harbors the *pycR1* and *pycR6* gene sequences inserted into a pTYGS-*ade* vector backbone. Correct cloning was confirmed by DNA sequencing. |
| CS I 158 1-2 | Labbok ID "CS I 158 1-2" refers to the creation of plasmid CS I 158 1-2. | Plasmid CS I 158 1-2 was created in the scope of this thesis. It harbors the $eupR3_{M261L}$ gene sequence inserted into a pET100/D-TOPO vector backbone. Correct cloning was confirmed by DNA sequencing. |
| CS I 161 1-2 | Labbok ID "CS I 161 1-2" refers to the creation of plasmid CS I 161 1-2. | Plasmid CS I 161 1-2 was created in the scope of this thesis. It harbors the $asR6_{Y344F}$ gene sequence inserted into a pET100/D-TOPO vector backbone. Correct cloning was confirmed by DNA sequencing. |
| CS I 82 6-5 | Labbok ID "CS I 82 6-5" refers to the creation of plasmid CS I 82 6-5. | Plasmid CS I 82 6-5 was created in the scope of this thesis. It harbors the *eupR6* gene sequence inserted into a pTYGS-*met* vector backbone. Correct cloning was confirmed by DNA sequencing. |
| CS I 128 1-10 | Labbok ID "CS I 128 1-10" refers to the creation of plasmid CS I 128 1-10. | Plasmid CS I 128 1-10 was created in the scope of this thesis. It harbors the *eupR3* gene sequence inserted into a pTYGS-*ade* vector backbone. Correct cloning was confirmed by DNA sequencing. |
| RS I 96 1-6 | Labbok ID "RS I 96 1-6" refers to the plasmid RS I 96 1-6 that was previously created by Raissa Schor and was obtained from the Cox Group strain collection.[123] | Plasmid RS I 96 1-6 harbors the *asPKS1, asL1, asL3* and *asR2* gene sequences inserted into a pTYGS-*arg* vector backbone. |
| RS I 63 2-5 | Labbok ID "RS I 63 2-5" refers to the plasmid RS I 63 2-5 that was previously created by Raissa Schor and was obtained from the Cox Group strain collection.[123] | Plasmid RS I 63 2-5 harbours the *asPKS1, asL1,* and *asL3* gene sequences inserted into a pTYGS-*arg* vector backbone. |

## *Curriculum Vitae*

| | |
|---|---|
| Jun 2017 – today | **PhD**, Microbial Chemistry, Gottfried Wilhelm Leibniz Universität Hannover, Germany |
| Oct 2014 – Mar 2017 | **Master of Science**, Chemistry, Philipps-University Marburg, Germany |
| Oct 2011 – Sep 2014 | **Bachelor of Science**, Chemistry, Philipps-University Marburg, Germany |
| Jun 2011 | **Abitur**, Albert-Einstein-Gymnasium Buchholz, Germany |

## List of Publications

[1] Kahlert, L; Schotte, C.; Cox, R.J. *Synthesis* **2021**, *53*, A-N.

[2a] Schotte, C.; Lei, L.; Wibberg, D.; Kalinowski, J.; Cox, R.J. *Angew. Chem. Int. Ed.* **2020**, *59*, 2-11.

[2b] Schotte, C.; Lei, L.; Wibberg, D.; Kalinowski, J.; Cox. R.J. *Angew. Chem.* **2020**, *132*, 2-11.

[3] Trenti, F.; Lebe, K.E.; Adelin, E.; Ouazzani, J.; Schotte, C.; Cox, R.J. *RSC Adv.* **2020**, *10*, 27369.

[4] Schor, R.; Schotte, C.; Wibberg, D.; Kalinowski, J.; Cox, R.J. *Nat. Commun.* **2018**, *9*.