The Biosynthesis of Phyllostictine A and Sch-642305 from *Phyllosticta cirsii* and *Phomopsis CMU-LMA*

Von der Naturwissenschaftlichen Fakultät der

Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des Grades

Doktor der Naturwissenschaften (Dr. rer. nat.)

genehmigte Dissertation

von

**Francesco Trenti, Dottore magistrale (Italien)**

2018

The truth is,
most of us discover where we are heading when we arrive.

Abstract

The presented work focuses on the biosynthesis of fungal natural products.

Structural analysis of phyllostictine A **168** revealed the literature structure to be erroneous and led to the revised bicyclic 3-methylene tetramic acid **190**. In particular, feeding ^{13}C-labeled precursors showed a haphazard incorporation into **168** and set the basis for the correct structure elucidation. The Biosynthetic Gene Cluster (BGC) was identified by total genome sequencing of the producer (*P. cirsii*) and *in silico* analyis. Targeted knock out experiments confirmed the BGC to be involved in **190** biosynthesis and produced the intermediate **199**.

In the attempt to define each biosynthetic step that takes part in Sch-642305 **169** biosynthesis, compounds **294** and **292** were observed after targeted knock out of a cytochrome P450 cytochrome and a flavoprotein oxidase, respectively. Genome assembly and *in silico* analysis showed over 150 BGC for this strain. The **169**-related BGC was defined by rational survey and homology comparison with other producing organisms, such as *Penicillium verrucosum* and *Penicillium brefeldianum*, producers of **169** and Brefeldin A **275** respectively. Although, other genes of **169** cluster were successfully disrupted, no other intermediate was observed.
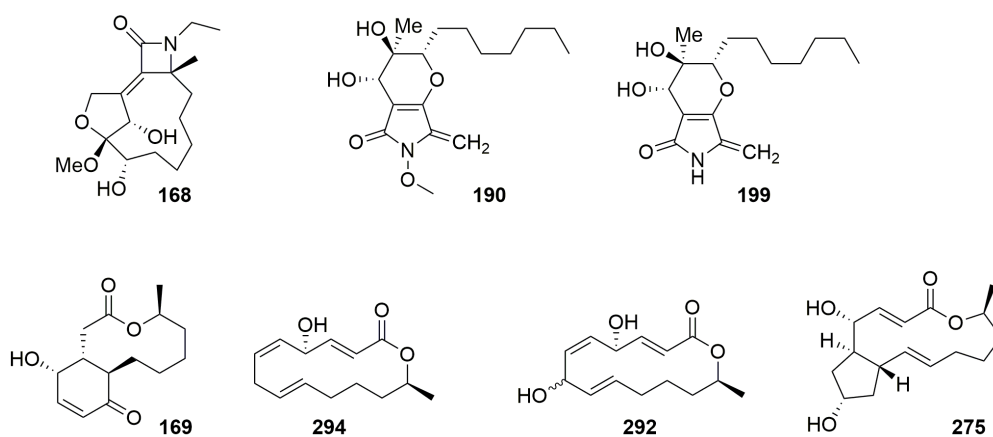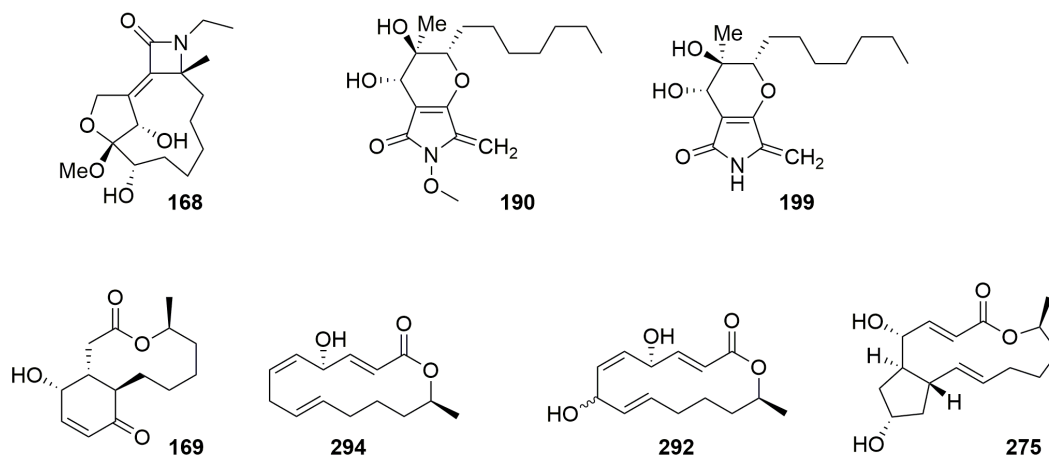
Keywords: Fungis; Biosynthesis; Natural products

Zusammenfassung

In der vorliegende Arbeit wurde die biosynthese von pilzlichen sekundärmetaboliten untersucht.

Die Strukturanalyse von phyllostictine A **168** zeigte, dass die in der ursprünglichen Veröffentlichung vorgeschlagene Struktur nicht korrekt ist und führte zur Revision der Struktur von **190** als bizyklisches 3-Methyltetramsäure Derivat. Besonders die Fütterungsexperimente mit ^{13}C-angereicherten Vorläufermolekülen zeigten eine willkürliche Inkorporation in **168** und etablierten damit die Basis zur Aufklärung der korrekten Struktur. Das entsprechende Biosynthese-Gen-Cluster (BGC) wurde durch die Genomsequenzierung des produzierenden Pilzes (*P. cirsii*) und durch eine *in silico* Analyse identifiziert. Gerichtete Gen-Knock-Out-Experimente bestätigten, dass das BGC und die Biosynthese von **190** verknüpft sind und führte zur Produktion des Intermediates **199**.

In dem Versuch alle Schritte der Biosynthese von Sch-642305 **169** zu analysieren, wurden die neuen Verbindungen **294** und **292** nach gerichtetem Gen-Knock-Out eines häm-abhängigen Cytrochrome P450, bzw. einer Flavoprotein-abhängigen Oxidase beobachtet. Die Genomassemblierung und *in silico* Analyse zeigten über 150 putative BGC für diesen Pilzstamm. Das *sch* BGC wurde durch eine rationale Vergleichsuntersuchung auf bekannte homologe BGC anderer Organsimen, wie *Penicillium verrucosum* und *Penicillium brefeldianum*, die ebenfalls bekannte Produzenten von **169** oder des verwandten Brefeldin A **275** sind, identifiziert. Obwohl weitere Gene des **169** BGC erfolgreich deaktiviert wurden, konnten keine weiteren Intermediate observiert werden.

Schlüsselwörter: Pilze; Biosynthese; Sekundärmetaboliten

Acknowledgement

List of Abbreviations

A: Adenylation domain
AAA: α-Aminoadipate Pathway
ACP: Acyl Carrier Protein domain
ACT: Artemis Comparison Tool
AntiSMASH: Antibiotics & Secondary Metabolites Analysis Shell
AT: Acyl Transferase domain
BGC: Biosynthetic Gene Cluster
BLAST: Basic Local Alignment Search Tool
C: Condensation domain
CDA: Czapek Dox Agar
CDD: Conserved Domains Database
C-Met: Carbon-MethylTransferase
CoA: Coenzyme A
COSY: Correlation Spectroscopy
DAL: Dihydroxyphenyl Acetic acid Lactone
DH: Dehydratase domain
DKC: Dieckmann release domain
DMAPP: Dimethylallyl-pyrophosphate
eGFP: Enhanced Green Fluorescent Protein
ELSD: Evaporative Light Scattering Detector
ER: Enoyl Reductase (domain)
ESI: Electron Spray Ionisation
extPCR: External PCR
FAD: Flavin Adenine Dinucleotide
FAS: Fatty Acid Synthase
FMN: Flavin Mononucleotide
Fw: Forward primer
gDNA: Genomic DNA
HIV: Human Immunodeficiency Virus
HMBC: Heteronuclear Multiple Bond Correlation
hp: hypothetical protein
HRMS: High Resolution Mass Sprectrometry
hrPKS: Highly Reducing Polyketide Synthase
HSQC: Heteronuclear Single Quantum Coherence spectroscopy
HygB: Hygromycin B
$HygR$: Hygromycin B Resistance construct $(P_{gdpA} + hph)$
IPP: Isopentenyl-PyroPhosphate
ITS: Internal Transcribed Spacer
KO: Knock Out
KR: Keto Reductase domain
KS: Keto Synthase domain

LCMS: Liquid Chromatography Mass Spectrometry
MCPA: 2-methyl-4-chlorophenoxyacetic acid
MeT: Methyltransferase
MS: Mass Spectrometry
m/z: Mass over Charge ratio
NAD(P): Nicotinamide Adenine Dinucleotide (Phosphate)
NGS: Next Generation Sequencing
NMR: Nuclear Magnetic Resonance
NOESY: Nuclear Overhauser Effect Spectroscopy
NRPS: Non-Ribosomal Peptide Synthetase
nt: nucleotide(s)
PCR: Polymerase Chain Reaction
PAL: Phenylalanine Ammonia Lyase
PCP: Peptide Carrier Protein domain
PDB: Potato Dextrose Broth
PDB: Protein Data Bank
PE: Pair End
PE: Peak Enhancement
PEP: Phospho Enoyl Pyruvate
PKS: Polyketide Synthase
PKS-NRPS: Polyketide Synthase-Non Ribosomal Peptide Synthase
PLP: Pyridoxal Phosphate
PP: PyroPhosphate
PPant: PhosphoPantheinyl
PR: Partially Reduced
prep-LCMS: Preparative LCMS
RAL: Resorcylic Acid Lactone
RCM: Ring Closing Metathesis
Re: Reverse primer
SAM: S-Adenosyl Methionine
SAT: Starting Acyl Transferase domain
spp: species
SQD: Single Quadrupole Detector
T: Thiolation domain
TE: Thiolesterase (domain)
TH: Thiohydrolase
UV: Ultra Violet
wPCR: Whole PCR
WT: Wild Type

Compound numbering

1 salicin
2 salicylic acid
3 acetylsalicylic acid
4 cystobactamid A
5 artemisinin
6 azadirachtin
7 paclitaxel
8 10-deacetylbaccatin III
9 muscarin
10 acetylcholine
11 alternariol
12 curvularin
13 mycophenolic acid
14 geodin
15 emodin
16 stipitatic acid
17 [*methyl*, ^{13}C]-methionine
18 actinorhodin
19 norsolorinic acid
20 cladosporin
21 erythromicyn A
22 lovastatin
23 nystatin
24 brevetoxin B2
25 acetyl-CoA
26 malonyl-CoA
27 coenzyme A
30 acetate
31 proprionate
55 monacolin J
56 naringenin
57 *p*-hydroxybenzalacetone
62 penicillin G
63 cytochalasin E
64 fusarin C
65 cyclosporin A
78 prefusarin
79 psilocybin
80 nicotine
81 quinine
82 ephedrine
83 strychnine
84 morphine
85 purine
86 pyrimidine
87 histidine
88 lysine

89 phenylalanine
90 tyrosine
91 tryptophan
92 anthranilic acid
93 ornithine
94 cocaine
109 isopentenyl-pyrophosphate (IPP)
110 dimethylallyl-pyrophosphate (DMAPP)
111 geranyl-PP
112 farnesyl-PP
113 geranylgeranyl-PP
114 farnesylfarnesyl-PP
115 menthane
116 pinane
117 bornane
118 fenchane
119 isocamphane
120 carane
121 thujane
122 menthol
123 camphor
124 xenovulene A
125 tetrahydrocannabinol (THC)
126 lysergic acid
140/**141** NAD$^+$/NADH
142/**143** NADP$^+$/NADPH
144/**145** FAD/FADH$_2$
146 FADH (semiquinone)
150 *S*-adenosyl methionine
163 phalloidin
164 α-amanitin
165 muscimol
166 cephalosporin
167 strobilurin A
168 phyllostictine A (erroneus)
169 Sch-642305
170 MCPA
171 clopyralyd
172 fusaric acid
173-175 phyllostictine B-D (erroneus)
178 clavam
179 carbapenem
180 monocyclic β-lactam
181 phaeosphaeride A = pyhllostictine B (re-assigned)
182-184 paraphaeosphaerides A-C
186 scytolide

Contents

1 Introduction

Natural products are small molecules (<1500 Da) produced by bacteria, fungi, algae, plants and even by some animals, that give an evolutionary advantage to competition in the surrounding environment.[1,2] Since they do not take part in maintaining homeostasis and are not necessary for life, they are also known as secondary, or specialised, metabolites.[3] Given the vast array of biological activities displayed by natural products, humans have learned to deploy them since the beginning of history for treatment of disease and injuries, hunting, transcendental magic rituals, perfumes, food additives and coloring agents.[4–6] Today natural compounds constitute valuable starting materials for drug development because the preexisting scaffolds of bioactive molecules can be chemically modified to enhance their biological activity, stability and intake.[7–9]

A classic example is salicin **1** and salicylic acid **2** (Figure 1.1), found in the bark of willow trees and a variety of plants such as meadowsweet (*Spiraea ulmaria*). The pain relieving and anti-inflammatory effects of salicylate-rich plants were already known by the Assyrians (4000 b.C.) and Sumerians (3500 b.C.).[10] The active compound salicin was purified in 1828 by the German chemist Johann Buchner[11] and the derivative acetylsalicylic acid (**3**, Figure 1.1) was chemically synthetised in 1853 by the french chemist Charles Gerhardt,[12] opening the way to the first synthetic pharmaceutical drug, aspirin (acetylsalicylic acid **3**).
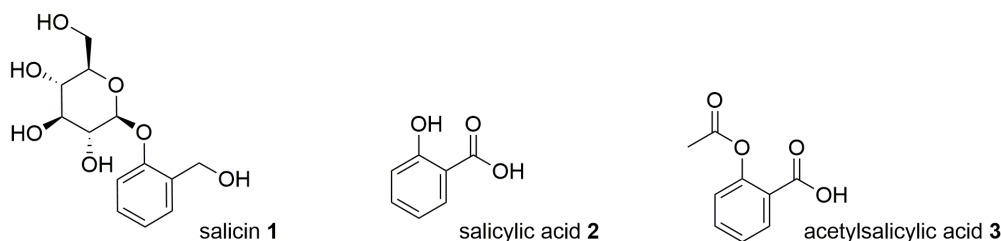
Figure 1.1: Salicylates in ancient and modern medicine

Similarly, the antibiotic penicillins, isolated from the fungus *Penicillium chrysogenum* in 1928, played a fundamental role in suppressing Gram-negative infections such as gonorrhea, meningitis and pharyngitis during the Second World War, but their true potential was reached through modification of their scaffold. Many classes of natural products are now used as medicines, either as the isolated compounds themselves, or as chemically modified analogs (Table 1.1).[13]

Molecule	Producer organism	Biological activity
Amphotericin B	*Streptomyces nodosus*	Antifungal[14]
Bialaphos	*Streptomyces hygroscopicus*	Herbicide[15]
Cephalosporins	*Cephalosporium chrysogenum*	Antibiotic[16]
Cyclosporin	*Trichoderma polysporum*	Immunosuppressant[17]
Daunorubicin	*Streptomyces spp*	Antitumoral[18]
Doxorubicin	*Streptomyces peucetius*	Antitumoral[19]
Erythromycin	*Streptomyces erythreus*	Antibiotic[20]
Gentamicin	*Micromonospora purpurea*	Antibiotic[21]
Kanamycin	*Streptomyces canus*	Antibiotic[22]
Lovastatin	*Aspergillus terreus*	Cholesterol lowering[23]
Mitosane	*Streptomyces caespitosus*	Antitumoral[24]
Monensin	*Streptomyces cinnamonensis*	Antibacterial[25]
Rifamycins	*Amycolatopsis rifamycinica*	Antibiotic[26]
Salinomycin	*Streptomyces albus*	Antibacterial[27]
Tacrolimus	*Streptomyces spp*	Immunosuppressant[28]
Tetracyclines	*Streptomyces aureofaciens*	Antibacterial[29]
Vancomycin	*Amycolatopsis orientalis*	Antibiotic[30]

Table 1.1: Molecules isolated from natural sources with biological activities.

Many active compounds were isolated and identified by virtue of their biological activities, such as the antibacterial cystobactamid A **4**,[31,32] the antimalarial agent artemisin **5** (Figure 1.2)[33,34] and the insecticide azadirachtin **6**.[35,36]
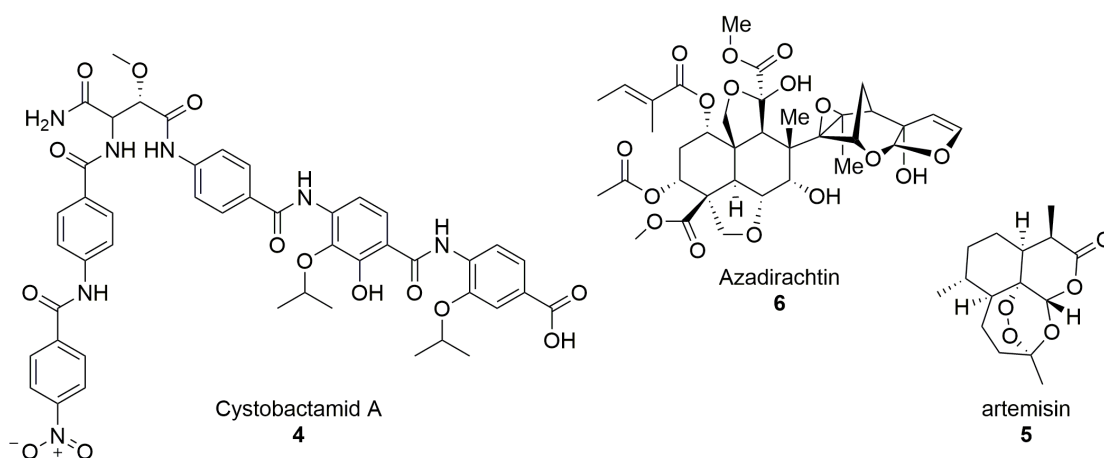
Figure 1.2: Bioactive natural products.

The impact of natural active compounds in modern medicine is noteworthy as they have evolved to interact with specific membrane proteins and enzymes within the cell: a review of small-molecule drugs revealed 63% to be derived from or based on compounds found in nature, with a relatively small number of totally synthetic origin.[8] Being able to access

2

a given compound without depending on the source is a challenge for chemists, who establish total or partial synthesis starting from a more available precursor. Paclitaxel **7**, for example, is the active compound in taxol drug, used in cancer therapy as it inhibits the depolymerization of the tubulin cytoskeleton. Total synthesis of **7** has been achieved, but with yields so low that it is not feasible in an industrial context.[37,38] Nowadays paclitaxel is obtained by semisynthesis starting from the precursor 10-deacetylbaccatin III **8** isolated from the European yew (Scheme 1.1).[39]

10-deacetylbaccatin III
8
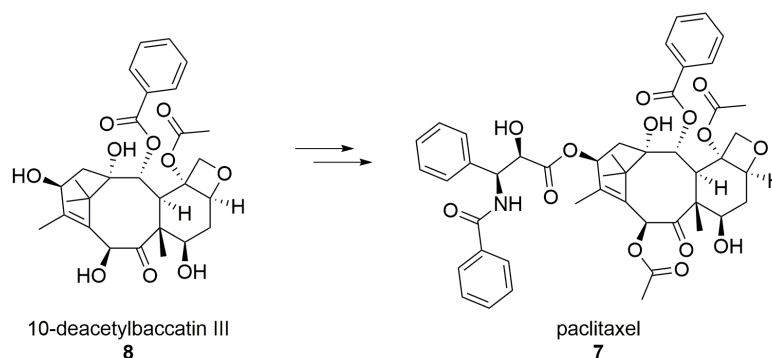
paclitaxel
7

Scheme 1.1: Semisynthesis of paclitaxel from 10-deacetylbaccatin III.

The ability of certain compounds to activate or suppress the activity of specific proteins make them powerful tools in understanding the molecular biology of life. For example, muscarin **9**, found in the mushroom *Amanita muscaria*, has high affinity for cholinergic receptors in the central and peripheral parasympathetic nervous system as it mimics the neurotransmitter acetylcholine **10** by virtue of its quaternary amine (Figure 1.3). Muscarin played a fundamental role in discerning the different classes of receptor in the nervous system.[40]

muscarine **9**
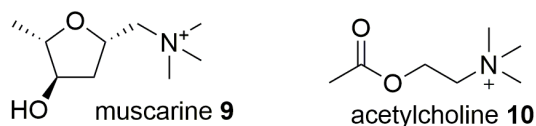
acetylcholine **10**

Figure 1.3: Muscarine and acetylcholine bind the same receptor in the nervous system.

Harold Raistrick (1890-1971) was a pioneer in natural product isolation and characterisation. Raistick and collaborators could define the structure of a variety of compounds isolated from plants, fungi and bacteria with exquisite precision considering that mass spectrometry techniques were still developing and NMR could yet reach frequencies high enough for a reliable structure assignment. At that time, structure characterisation was established by the reactivity of the purified compound: reaction with O_3 discerned olefins, bromination defined the aromaticity of the compound, methylating agents would spot free

hydroxy groups. Assembling all the pieces of information to deduce atom-connectivity was a puzzle so complex that many compounds would wait eventually decades to have a full structure characterisation. Among the molecules investigated by Raistrick we can find alternariol **11**, curvularin **12**, mycophenolic acid **13**, geodin **14**, emodin **15** and the tropolone stipitatic acid **16**.[41]
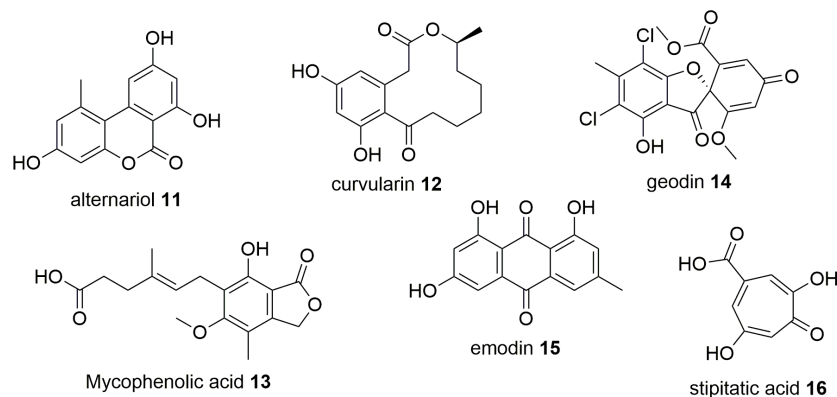
Figure 1.4: Compounds investigated by Raistrick.

Modern approaches for the isolation and characterisation include high/low resolution mass spectrometry-coupled liquid chromatography (LCMS), 1D/2D nuclear magnetic resonance (NMR) and crystallography to unravel the absolute structure of the natural compounds. Classic ^{13}C and ^{18}O-labelling experiments, where heavy precursors are fed to the producing organism gives insights on atoms connectivity and rearrangement during the biosynthesis. For example, the mechanism behind stipitatic acid **16** biosynthesis *via* oxidative ring expansion was proposed thanks to labeling experiments: feeding [*methyl*, ^{13}C]-methionine **17** resulted in the labeling of one carbon in the tropolone ring (Scheme 1.2).[42]
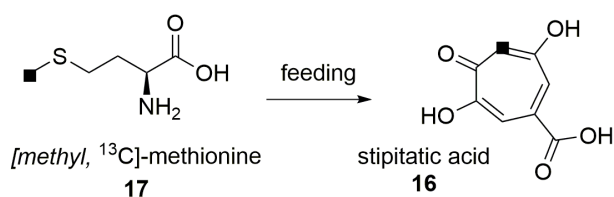
Scheme 1.2: Incorporation of labeled methyl into the tropolone ring.

In an era of exponentially progressing genome sequencing, biologists and chemists have started digging deeper into the molecular mechanisms of natural product biosynthesis: thanks to genome accessibility and manipulation, the study of biosynthetic pathways investigates the production of secondary metabolites *in vivo* through targeted genome modifications and heterologous expression. The ability to interfere with secondary metabolite

pathways, along with expressing entire gene clusters in host organisms, gives insight to dissect the chemistry behind each biosynthetic step. *In vitro* studies shed light on the enzymatic dynamics and mechanisms that lead to enantioselective, complex structures, including the programming of *cis*-acting catalytic domains of iterative PKS and post elongation tailoring enzymes.

1.1 Polyketides and Polyketide Synthases

Polyketide natural products are a remarkable class of molecules. They exhibit a wide range of functionality and structural diversity, and are fundamental active substances in modern medicine, including antibiotic, anticancer, antifungal, antiparasitic and immuno-suppressive properties.[43] Some examples can be found in polycyclic aromatics, such as actinorhodin **18**, norsolorinic acid **19** and cladosporin **20**; macrolactone polyketides such as erythromicyn A **21**; the decalin lovastatin **22**; polyene polyketides, such as nystatin **23**; and polyether systems such as the brevetoxin class **24** (Figure 1.5). The simple chemical assembly of scaffolds of this entire family stems from fatty acid biosynthesis in primary metabolism: the core enzyme Polyketide Synthase (PKS) is chemically and structurally related to the Fatty Acid Synthases (FAS) involved in primary metabolism.[43–48]
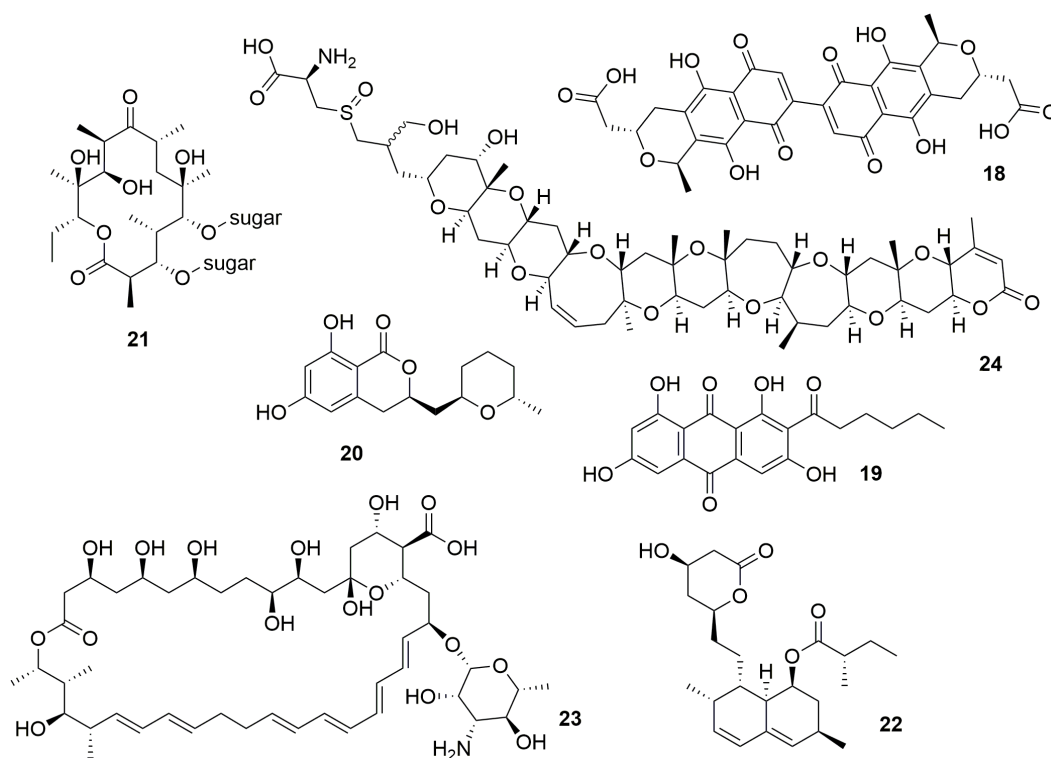
Figure 1.5: Examples of polyketide natural products.

Polyketides are assembled by decarboxylative thioclaisen condensation by Polyketide Synthases (PKS).[49] The biosynthetic process has two distinct phases: the chain extension and the β-processing cycle. The first step includes loading of the building blocks (usually acetyl-CoA **25** and malonyl-CoA **26**) and formation of C-C bonds, while the later steps process the β-ketone. PKS are composed by domains: Acyl Carrier Protein (ACP) binds the substrate; ketosynthase (KS) and acyltransferase (AT) mediates the loading and elongation; ketoreductase (KR), C-methyltransferase (C-MeT), dehydratase (DH), enoylreductase (ER) mediates the α/β-processing. An important feature of PKS is the post-translation addition of the prosthetic group phosphopantetine (PPant) donated by coenzyme A (CoA) **27** on a serine residue of the ACP domain, converting the *apo*-enzyme to the active *holo*-enzyme (**28** to **29**, Scheme 1.3). The PPant group mimics the CoA in the formation of the thiolester bond with the starting unit and acts as swinging arm that feeds the catalytic domains through every *cis*-tailoring step. The thiolester linkage stabilize both the α-carbanion and the elecrophile thiolester of the elongating chain. This is possible by the nature of the sulfur atom in the thiolester linkage, which is a weak electron donor and acts as a optimal leaving group during each cycle of thioclaisen condensation.
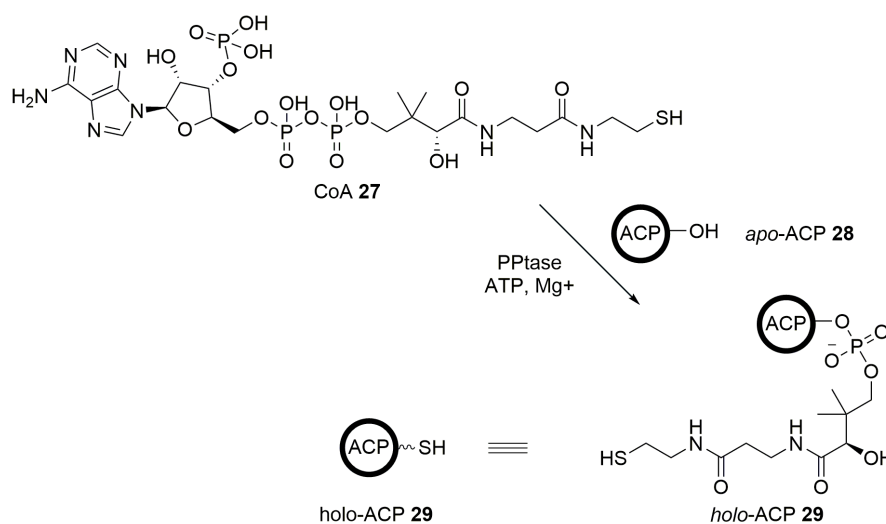
Scheme 1.3: Conversion of *apo*-ACP to *holo*-form with consumption of CoA and ATP by the phosphopantetinase enzyme PPtase.

The chain elongation requires a starter unit, normally acetate **30** or propionate **31**, that is activated by covalent bonding to CoA mediated by the acetyl-CoA synthetase enzyme with consumption of ATP (**30** to **25**, Scheme 1.4). The activated starting unit is then covalently tethered to the starting Acyl Transferase (AT or SAT) domain of the PKS (**29** to **32**, Scheme 1.4). The starter unit is then loaded onto the KS domain **33**. The KS mediates the C-C bond formation with a malonyl-CoA extender unit **26**, which is transferred by the AT domain onto the ACP (**34** to **35**, Scheme 1.4). The KS

domain extends the chain by thioclaisen condensation: malonyl-ACP can easily undergo decarboxylation and act as nucleophile in the formation of the β-keto thiolester bond (**35** to **36**, Scheme 1.4).

The catalytic domains of PKS are often covalently bound *via* peptidic bond and perform, beside the chain extension, a series of α/β-processing events: S-adenosylmethionine (SAM)-dependent α-methyltransfer mediated by the methyl transferase domain (C-MeT); β-keto reduction by the NADPH-dependent ketoreductase domain (KR, **36** to **37**, Scheme 1.4); dehydratation by the DH domain to give β-olefin (**37** to **38**, Scheme 1.4); and further reduction of the β-olefin by the ER domain to yield the saturated alkyl chain (**38** to **39**, Scheme 1.4). A fully processed polyketide chain is released hydrolytically, with the formation of carboxylic acid **40** or reductively, generating an aldehyde **41** or a primary alcohol **42**. A *cis*-thiolesterase domain (TE) can mediate the release of the polyketide chain, but it is also a common feature for PKS to rely on *trans*-acting thiolesterases. The multitude of possible post- and in-assembly line decoration of oxidative and reducing events give rise the enormous varieties of these compounds in nature.
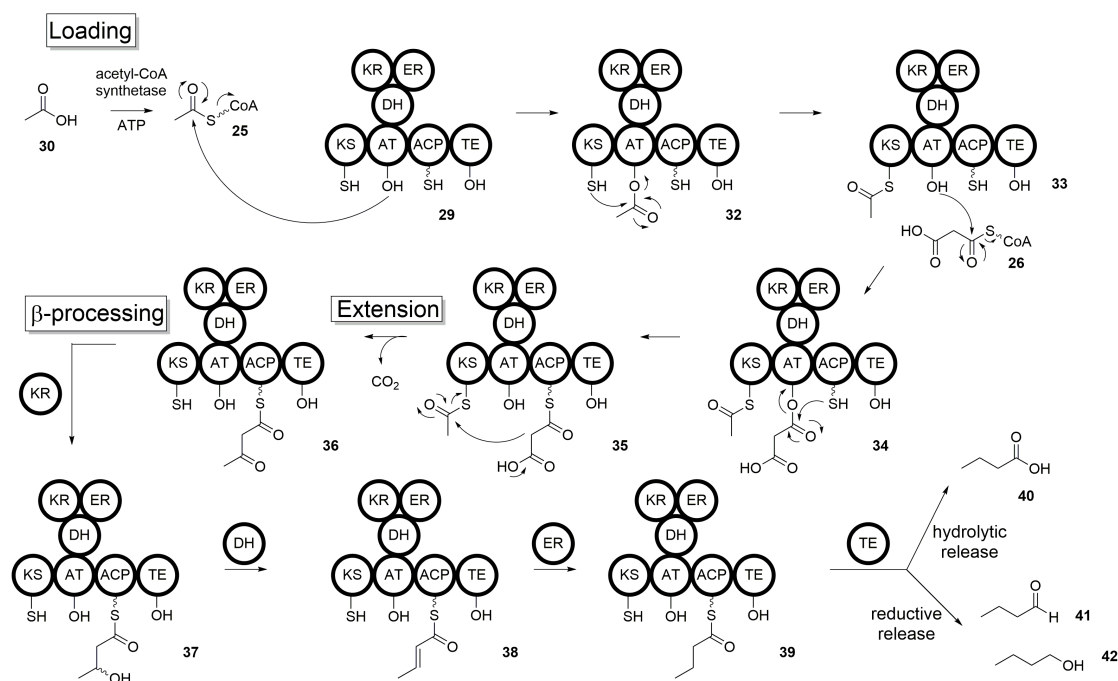
Scheme 1.4: Chain elongation and β-processing of polyketide chains.

Type I PKS are subdivided into two main categories: modular and iterative.
Modular type I PKS are gargantuan enzymes, which form by an assembly line of *cis*-working modules. They are supramolecular assemblies constituted by discrete proteins with modular domains that work in an assembly line fashion during the biosynthesis of

7

the growing polyketide. Every module contains a set of catalytic domains to introduce variability in the formation of functional groups.

A well characterised example is erythromycin A **21**, a macrolactone polyketide synthethised by DEBS modular PKS in *S. erythreus* (Scheme 1.5).[50,51] DEBS is constituted by three proteins (DEBS1-3) non-covalently bound, which contain seven modules (Loading + modules 1-6) for a total of 28 domains working in line. Each module has the ACP and KS domains for loading and chain extension, plus a combination of β-processing domains (KR, DH and ER). Propionyl-ACP **43** is the starting unit, which is elongated with S-methylmalonyl-CoA **44**.[52] Six events of chain elongation yield a C_{15} chain decorated by methyl groups derived from S-methylmalonyl-CoA, and reductions that depend on the β-processing domain arsenal of each module (**45** to **50**, Scheme 1.5). The chain is released by macrolactonisation mediated by the *cis*-acting TE domain to yield **51** and post-elongation tailoring events mediated by P450 cytochromes and glycosyltransferases finally give **21**.[52]

Scheme 1.5: Modular organization of DEBS megasynthase.

8

Iterative type I PKS are enzymes composed of a single multidomain protein. In contrast to modular PKS, each domain acts iteratively on the substrate in a non-predictable fashion. Methylation or reductive events might happen or not at any given elongation event, leaving in place the ketone or reducing it to alcohol, olefin or methylene (Scheme 1.6).

Iterative PKS can be divided into three subclasses according to the reducing potential of their domain arsenal. Highly Reducing PKS (hr-PKS) posses all KR, DH and ER β-processing domains and can potentially fully reduce a ketone to methylene. Partially Reducing PKS (pr-PKS) lacks the ER domain and can at maximum reduce the carbonyl to an olefin. Non-Reducing PKS (nr-PKS) lacks all reducing domains and their products are true, fully oxidised, polyketides.
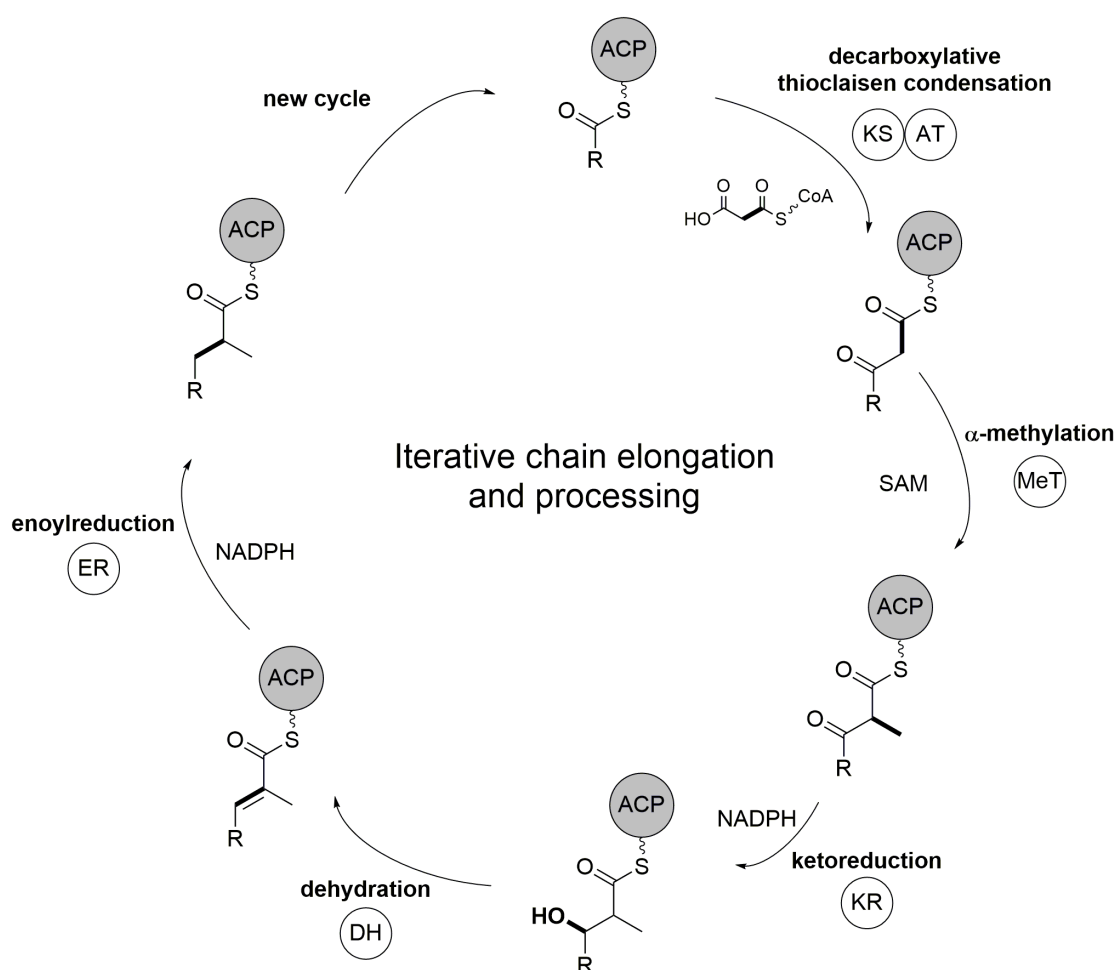
Scheme 1.6: Catalytic cycle of a highly reducing fungal PKS. A single set of domains is used iteratively, and each domain can introduce structural variety.

Iterative PKS are typical in fungal secondary metabolism. For example, the metabolite lovasatin **22**, isolated from the filamentous fungus *Aspergillus terreus*, is derived from two iterative PKS (LovB and LovF) products (Scheme 1.7).[53] In particular, LovB is the nonaketide synthase that is proposed to catalyse the decalin ring formation by Diels-Alder reaction at the hexaketide stage (**52** to **53**, Scheme 1.7). LovB works in conjunction with the *trans*-acting enoyl reductase LovC to fully reduce part of the elongating chain (**52** to **54**, Scheme 1.7). The polyketide chain is then offloaded from LovB by the *trans*-acting thiolesterase LovG and oxidised by the cytochrome P450 LovA to yield monacolin J **55**. A second highly reducing iterative PKS (LovF) produces the methylated diketide, which is transferred onto the LovA-derived hydroxy group by the acyltransferase LovD. Intramolecular condensation, finally yields **22** (Scheme 1.7).[54]
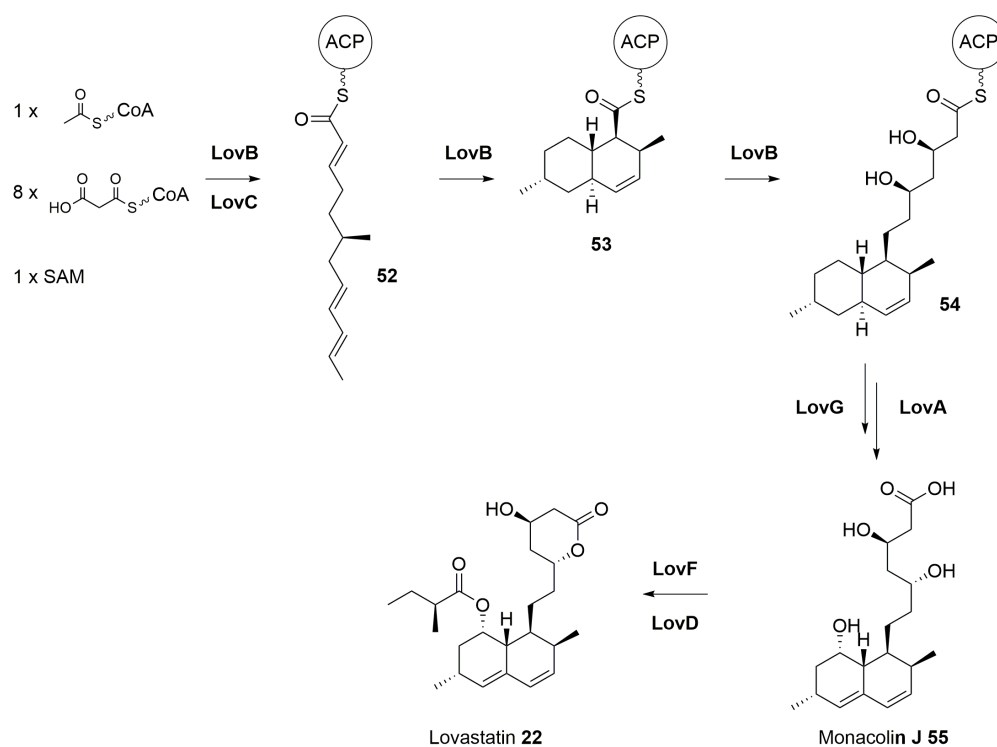
Scheme 1.7: Biosynthesis of lovastatin **22**.

Type II PKS are composed of discrete proteins that transitionally assemble to perform the chain initiation, elongation and release. These complexes are thought to be the ancestors of type I PKS, developing before gene fusion would eventually occur to result in *cis*-interaction.[43] Typically, type II PKS produce aromatic compounds such as the antibiotic actinorhodin **18** (Figure 1.8).[55, 56]
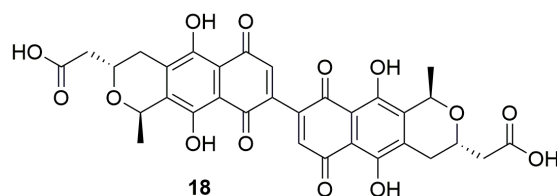
Scheme 1.8: Type II PKS derived actinorhodin **18**.

Type II PKS generally lack the malonyl-CoA AT domain (MAT) and it has been shown that their ACP domains can undergo self-acylation from malonyl-CoA.[57] The assembly of type II derived polyketides happens in an iterative fashion, similar to iterative type I PKS, although type II systems are simpler, with minimal domain composition of a discrete KS dimer and an ACP.[43]

Type III PKS are the simplest kind, constituted by a single protein with only a KS domain dimer. They lack the AT and ACP domains and can process malonyl-CoA directly: loading; chain extension; ring closing and/or aromatisation of the product occur in a single active site.[58] Type III systems typically produce unreduced polyketides; they accept a wide range of building blocks, such as p-coumaril-CoA, feruloyl-CoA and even unnatural CoA-thiolesters, making them ideal candidates in directed biosynthesis of new biologically active compounds.[59] They commonly yield phenylpropanoids in plants, but they have also been reported in bacteria and fungi.[43,60,61] Examples of type III polyketides are naringenin **56** and p-hydroxybenzalacetone **57** from grapefruit and raspberry respectively (Figure 1.9).[62,63]
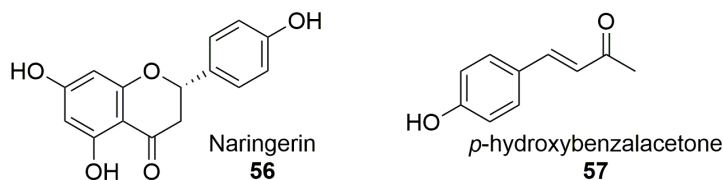
Scheme 1.9: Type III PKS derived naringenin **56** and p-hydroxybenzalacetone **57**.

1.2 Non-Ribosomal Peptide Biosynthesis

Non-ribosomal peptides are compounds constituted by proteogenic and non-proteogenic amino acids assembled by Non-Ribosomal Peptide Synthetases (NRPS).[64] NRPS normally display a modular organization that follow a similar in-line logic found in polyketide biosynthesis for the selection, activation, loading and extension of amino acid building blocks. Nonetheless, iterative NRPS and mixed modular/iterative systems have also

been reported.[65] The products differ for their amino acid constitution, and they are often cyclised and post-assembly modified.

Each module of NRPS enzymes have three core catalytic domains required for amino acid selection, loading and chain extension: the Adenylation domain (A); the Thiolation or Peptidyl Carrier Protein (T or PCP); and the Condensation domain (C). A fourth terminal releasing thiolesterase domain (TE) releases the growing chain from the enzyme (Scheme 1.10).[64,66] The A domain selects and activates the starter amino acid by adenylation, consuming ATP (**58** to **59**, Scheme 1.10). The activated starting unit is then transferred to the PCP (**60** and **60.a**, Scheme 1.10), which acts as anchor of the substrate during each elongation step. The C domain carries the formation of the peptidic bond between the activated amino acid specifically selected by the A domain of each module (**60/60.a** to **61**, Scheme 1.10). The chain is then released by the TE domain usually by intramolecular cyclisation, or by addition of a molecule of water yielding a linear peptide. Reductive release of the product is also possible by a Reductive domain (R) to give a peptide aldehyde.[67] Other known NRPS feature epimerase, cyclase, methyltransferase and formylase domains.[68] One example of NRPS product is penicillin G **62** (Scheme 1.10).[69,70]
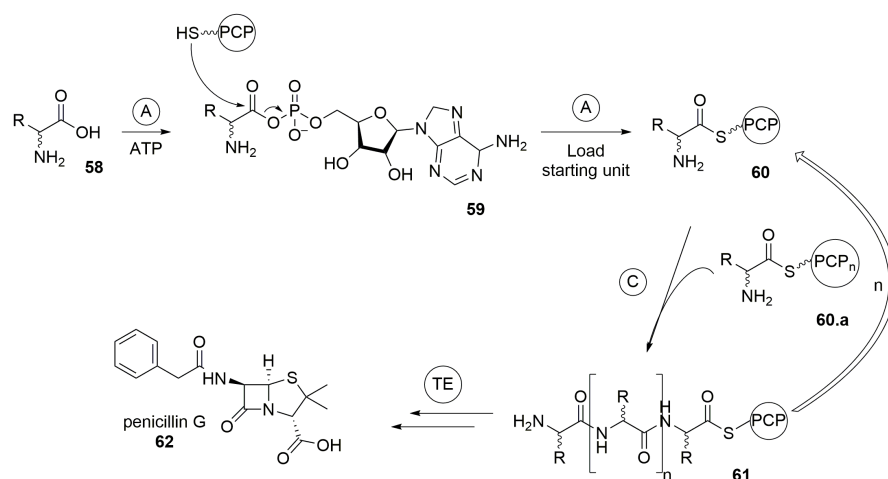
Scheme 1.10: Minimal NRPS enzyme and the NRPS product **62**.

1.3 PKS-NRPS Hybrid Systems

Gene fusion and the compatibility of in-line assembly mechanisms of both PKS and NRPS, allowed the evolution of hybrid PKS-NRPS machinaries.[71–73] Representative examples of fungal iterative PKS-NRPS enzymes are CcsA and FusA, responsible for the biosynthesis of pyrrolidinones cytochalasins **63** and fusarin **64**, respectively.[74–76] Another well known PK-NRP product is the immunosuppressant cyclosporin A **65**.[77,78]
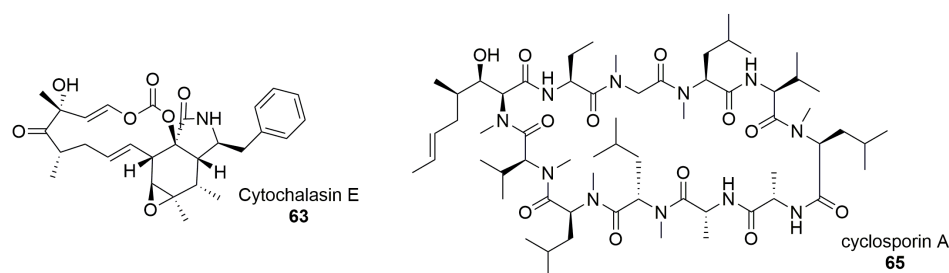
Figure 1.6: Examples of PK-NRP products.

It is interesting to notice that the PKS module can be up or downstream of the NRPS. In case of PKS-NRPS, the condensation domain will form a peptidic bond between the polyketide and the activated amino acid using the nitrogen atom as nucleophile (**66** and **67** to **68**, Scheme 1.11 **A**). In case of NRPS-PKS systems, the KS domain must recognise the thiolester of the PCP-peptidyl as electrophile, forming a C-C bond (**69** and **70** to **71**, Scheme 1.11 **B**).[73]
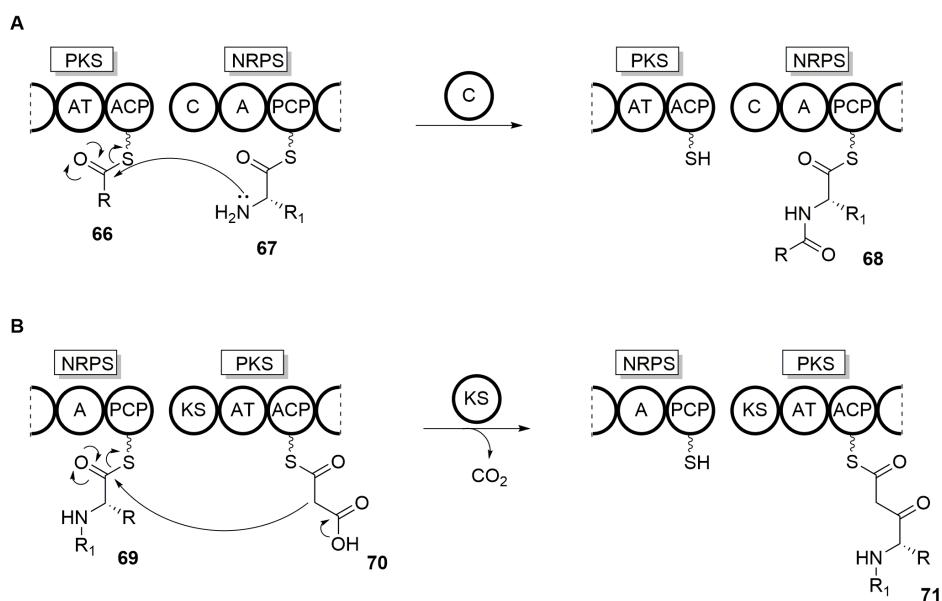
Scheme 1.11: Different architecture of PKS-NRPS (**A**) and NRPS-PKS (**B**) leads to the formation peptidic bond or C-C bond respectively.

Fusarin C **64** is a pyrrolidinone-containing type I iterative PKS-NRPS product. Acyl-pyrrolidinone systems generally derive from fusion between amino acids and polyketide precursors, and feeding experiments confirmed that the pyrrolidinone ring of fusarin C is derived by condensation of a polyketide and, most likely, homoserine or aspartic acid.[79] During the biosynthesis, the polyketide moiety is held by the ACP domain (**75**, Scheme

1.12), homoserine is then adenylated and transferred to the PCP (**72** to **74**). The condensation domain catalyses the nucleophilic attack by the homoserine nitrogen atom on the ACP-bound polyketide forming the intermediate amide **76**. The final biosynthetic steps involve reduction of the PCP-bound thiolester by the NRPS reductive release domain to yield the aldehyde (**76** to **77**), and cyclisation to form the pyrrolidone 5-membered ring of prefusarin C (**78**). Further oxidation downstream, such as epoxidation, hydroxylation and *O*-methylation are carried out by tailoring enzymes to yield the final product fusarin C **64** (Scheme 1.12). The mechanism of formation of the pyrrolidone ring involves the nucleophilic attack of the aldehyde and Knoevenagel condensation with loss of water (**77** to **78**). A detailed discussion about pyrrolidinones and tetramic acids formation is addressed in section 2.4.4.
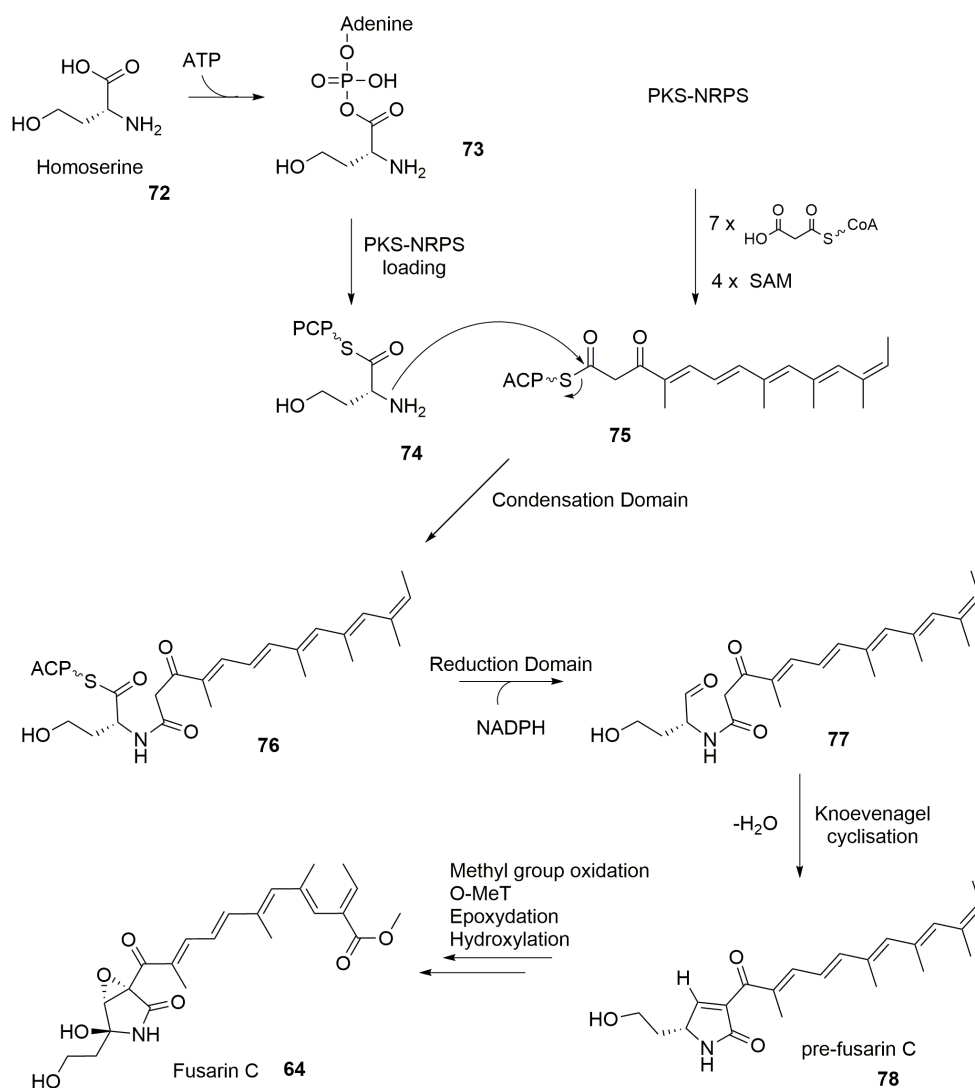
Scheme 1.12: Proposed biosynthesis of fusarin C.[79]

14

1.4 Alkaloid Biosynthesis

Alkaloids are compounds produced in plants and fungi, distinct for possessing at least one basic nitrogen.[80,81] Representative examples of alkaloids are psilocybin **79**, nicotine **80**, quinine **81**, ephedrine **82**, strychnine **83** and morphine **84** (Table 1.2).

psilocybin **79**   nicotine **80**   quinine **81**

ephedrine **82**   strychnine **83**   morphine **84**
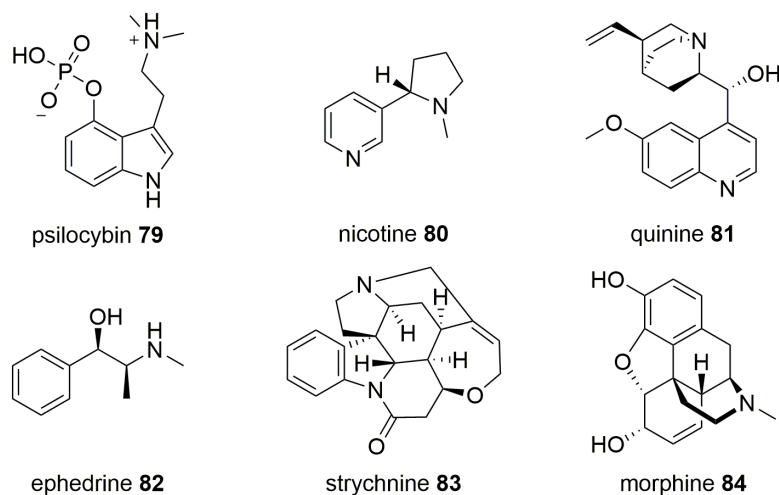
Molecule	Producer organism	Biological activity
Psilocybin	*Psilocybe spp*	Antidepressant[82]
Nicotine	*Nicotiana tabacum*	Stimulant[83]
Quinine	*Cinchona officinalis*	Antimalarial[84]
Morphine	*Papaver somniferum*	Pain killer[85]
Strychnine	*Strychnos spp*	Heart poison[86]
Ephedrine	*Ephedra sinica*	Blood pressure enhancer[87]

Table 1.2: Alkaloids with strong biological activities.

Alkaloids stem mostly from amino acid and nucleic acids such as purines and pyrimidines (**85**, **86**), with a vast array of scaffolds (Figure 1.8). The principal proteogenic amino acids involved in their biosynthesis are histidine **87**, lysine **88**, phenylalanine **89**, tyrosine **90** and tryptophan **91**, plus the non-proteogenic anthranilic acid **92** and ornithine **93** (Figure 1.7).[88]

Figure 1.7: Amino acids and nucleotides involved in alkaloid biosynthesis.

Alkaloids present a great variety of scaffolds, depending on their biosynthetic origin. Indole alkloids are derived from tryptophan, piperidine alkloids stem from lysine, pyrrolidines from ornithine and imidazole alkaloids are derived from histidine (Figure 1.8). Alkaloids do not have a standard assembly line like polyketides and peptides, and their biosynthesis can cross over with other pathways. For example, the formation of tropane rings found in cocaine **94** stems from ornithine and a diketide (Scheme 1.14).

Figure 1.8: Alkaloid scaffolds.

A common, early biosynthetic step in amino acid-derived alkaloid biosynthesis is the enzymatic decarboxylation of the amino acids to the corresponding amine (Scheme 1.13).[89–91] The pyridoxal phosphate (PLP) cofactor **95** plays a fundamental role, tethering the amino group of the substrate and acting as electron recipient. PLP is the activated form of the purine-derived vitamin B_6 **96**, stabilized as an imine in the active site of the decarboxylase by a lysine of the enzyme (**95**, Scheme 1.13). An amino acid can form a Schiff base in conjunction with PLP, and decarboxylation can occur *via* dihydropyridine formation (**98** to **100**, Scheme 1.13). The pyridinium ring can be restored allowing the release of

the decarboxylated amino acid **101** and the restoring of the cofactor (**100** to **95**, Scheme 1.13).
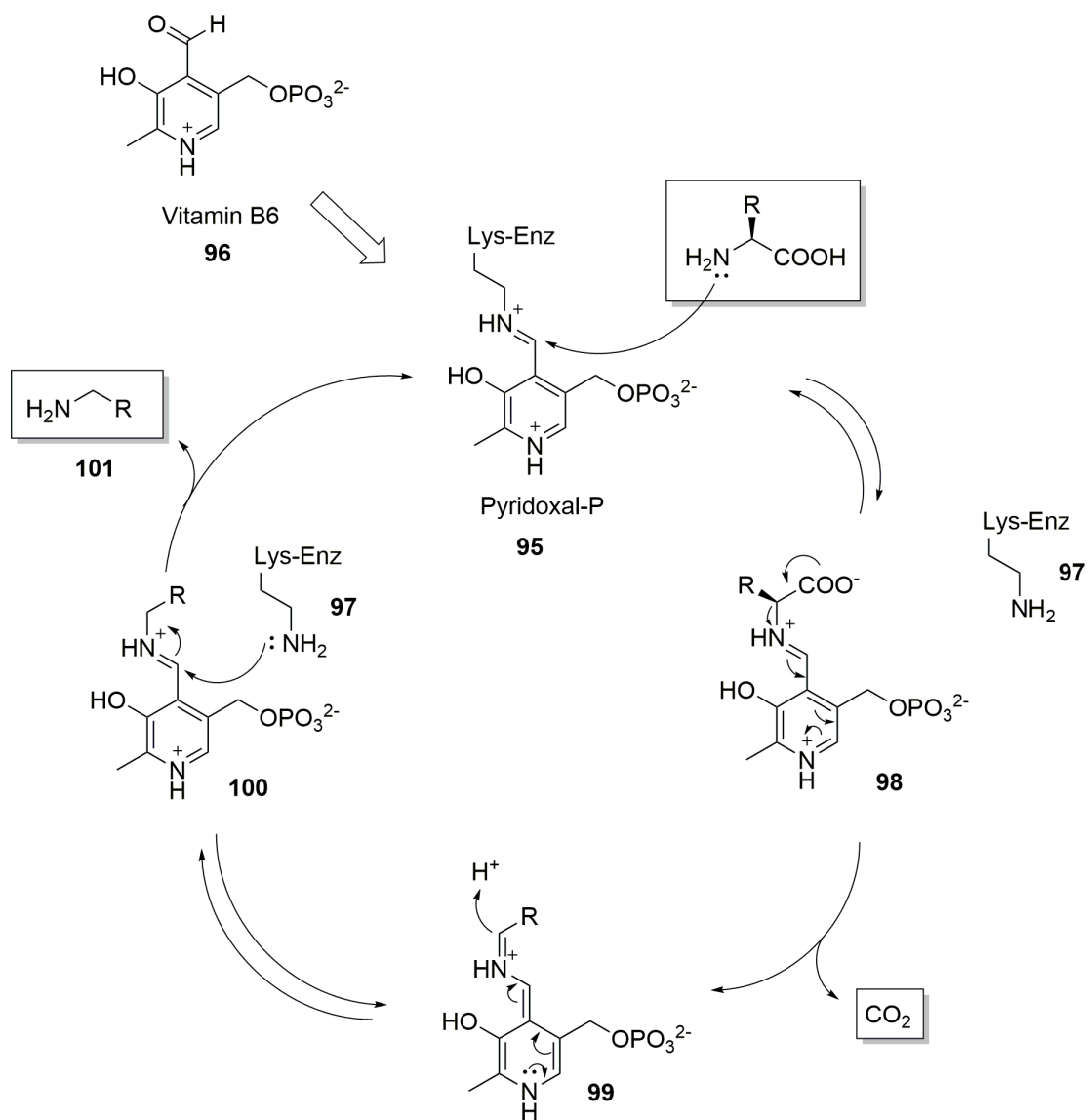
Scheme 1.13: PLP-mediated decarboxylation in alkaloid biosynthesis.

Cocaine **94** is an analgesic and stimulant alkaloid extracted from *Erythroxylum coca* (*Erythroxylaceae*), with a millennial usage in South America to overcome pain and fatigue by chewing the coca plant leaves or by drinking infusions.[92] It is produced from decarboxylation of *L*-ornithine **93** to yield **102** and oxidation to *N*-methylbutanal, which can cyclise (**103** to **104**) and be elongated by two acetate units (**104** to **105**, Scheme 1.14).

Another amine oxidation (**106**) sets the electrophile to allow nucleophilic attack to form the tropane ring (**107**), which is finally reduced (**108**) and condensed to benzoic acid to yield **94**.[93–96]
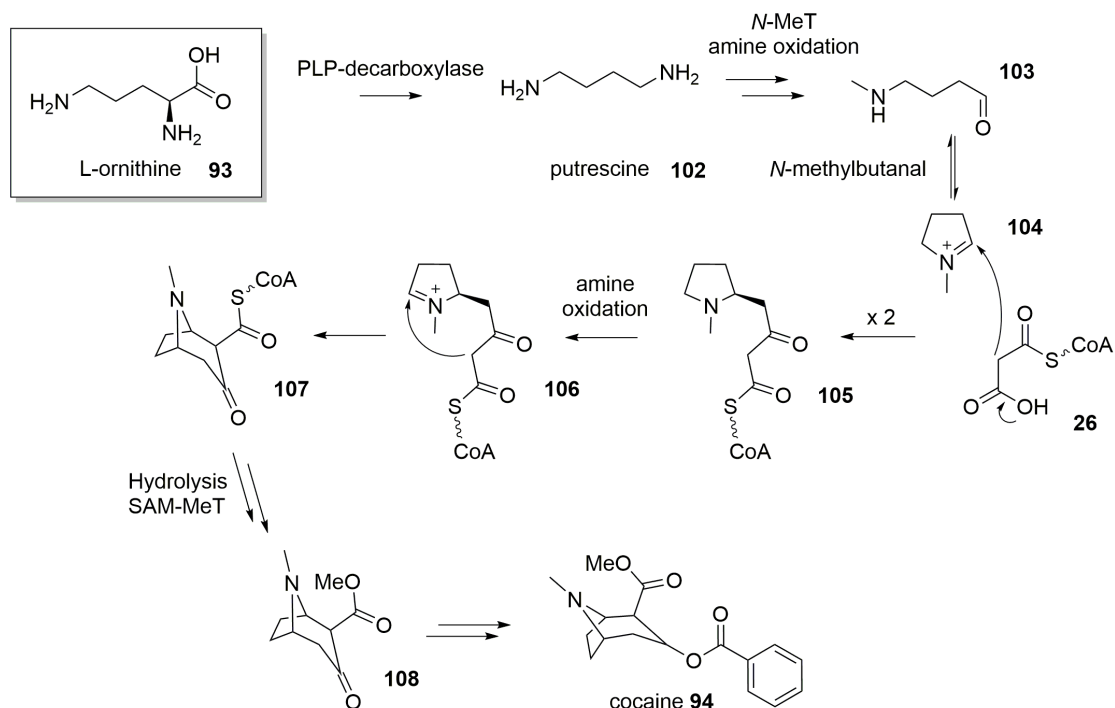
Scheme 1.14: Biosynthesis of cocaine.

1.5 Terpenes

Terpenes are the most abundant natural products in nature, with more that 30.000 known members.[97] They are produced principally in plants as primary constituents of essential oils in flowers. All terpenes are formed through the heterolytic reactions between isopentenyl-pyrophosphate (IPP) **109** and dimethylallyl-pyrophosphate (DMAPP) **110** based on the formation of transient carbocations and carbanions, following a "head to tail" polymerization to form linear chains (Scheme 1.15).[98] Terpenes posses carbons in multiple of five. Geranyl-PP **111** yields monoterpenes (n = 2), farnesyl-PP **112** forms sesquiterpenes (n = 3), geranylgeranyl-PP **113** forms diterpenes (n = 4), farnesylfarnesyl-PP **114** yields triterpenes (n = 6) and so on. IPP and DMAPP are both derived from primary metabolism, and specifically, from the mevalonate and deoxyxylulose pathways.[99,100]
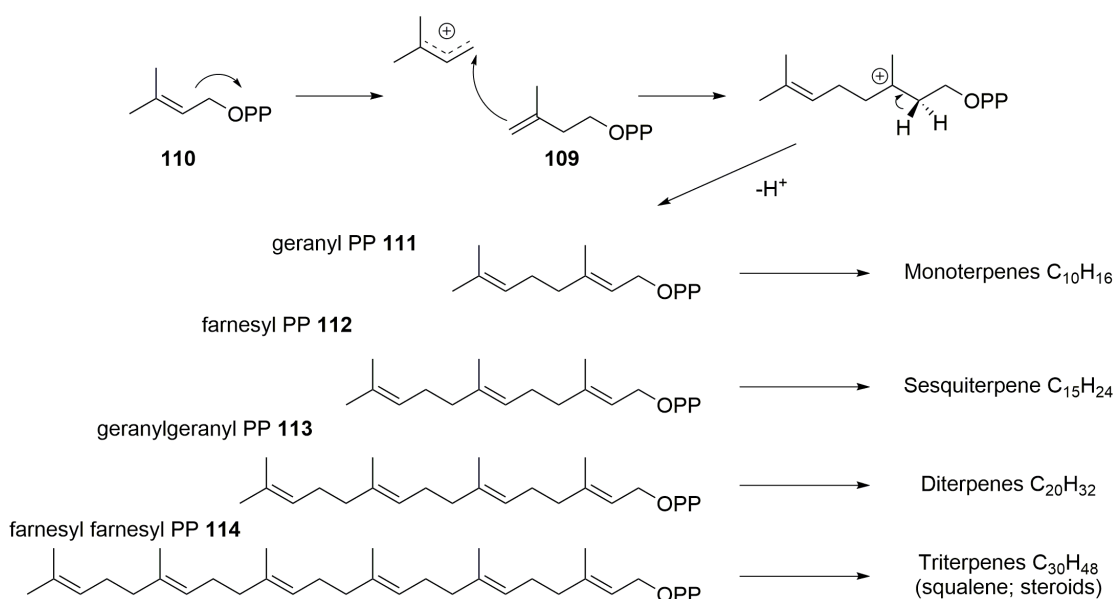
Scheme 1.15: Polymerisation of IPP and DMAP.

Terpene chains can undergo enzyme-mediated cyclizations to yield different ring cores: a simple geranyl-PP can yield up to seven unique ring scaffolds **115-121**, by controlling the reactive carbocations (Scheme 1.16).[101, 102]
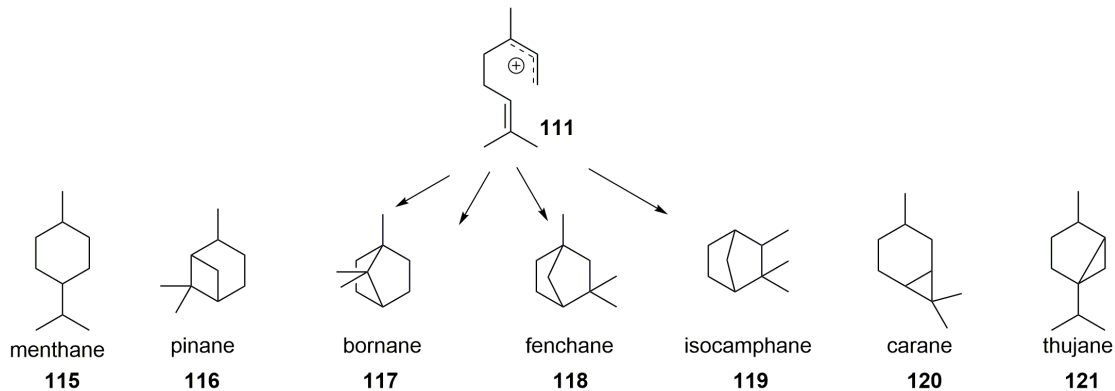
Scheme 1.16: Different cyclisations of geranyl-PP.

Rings or open-chains are often further oxidised and processed by tailoring enzymes, such as cytochromes P450, non-heme iron dependent oxidases and flavoproteins. Decorated scaffolds are generally named terpenoids or isoprenoids. Iconic examples of terpenoids can be found in menthol **122** and camphor **123** (Figure 1.9). Hybrid terpenoid-PKS compounds are known as meroterpenoids, as xenovulene A **124**[103, 104] and tetrahydro-

cannabinol (THC) **125**.[105] Terpenoids take part in alkaloid biogenesis as well, such as in the lysergic acid **126** pathway.[106–108]
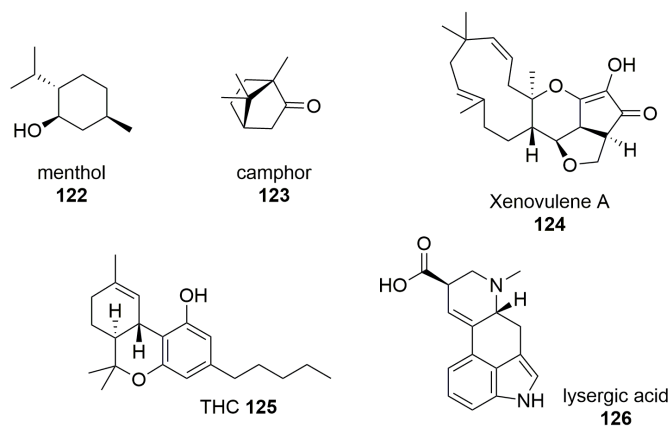
Figure 1.9: Examples of terpenoid or hybrid terpenoid compounds.

1.6 Tailoring Enzymes and Cofactors

Although the core enzymes (PKS, NRPS, Terpene Synthase) lead to the synthesis of the principal scaffolds, the biosynthesis continues downstream with oxidoreductive decorations to give the final bioactive compounds.[77] Oxidative post-assembly modifications are more common, with insertion of one or two oxygen atoms from atmospheric O_2; or electron transfer resulting in oxidation of alcohols to aldehydes and ketones; dehydrogenation and dehydration with consequent olefin formation; and methyl/acyl-transfer events by tailoring enzymes, with S-adenosyl methionine (SAM) as the main methyl donor, and actor in C-C bond formation *via* a concerted radical mechanism (Section 1.6.2).[109,110]

1.6.1 Oxidases and Oxygenases

Atmospheric oxygen, as a ground state triplet ($^3O_2$, Scheme 1.17 **A**), is not a spin paired species, therefore it is inert towards spin paired organic compounds. This is not true for singlet oxygen ($^1O_2$, Scheme 1.17 **B**), the excited state with paired electrons, that can react easily with organic compounds. However, $^1O_2$ has an energy of 22 Kcal mol^{-1} above $^3O_2$, therefore it is not accessible by biological systems.[111,112]
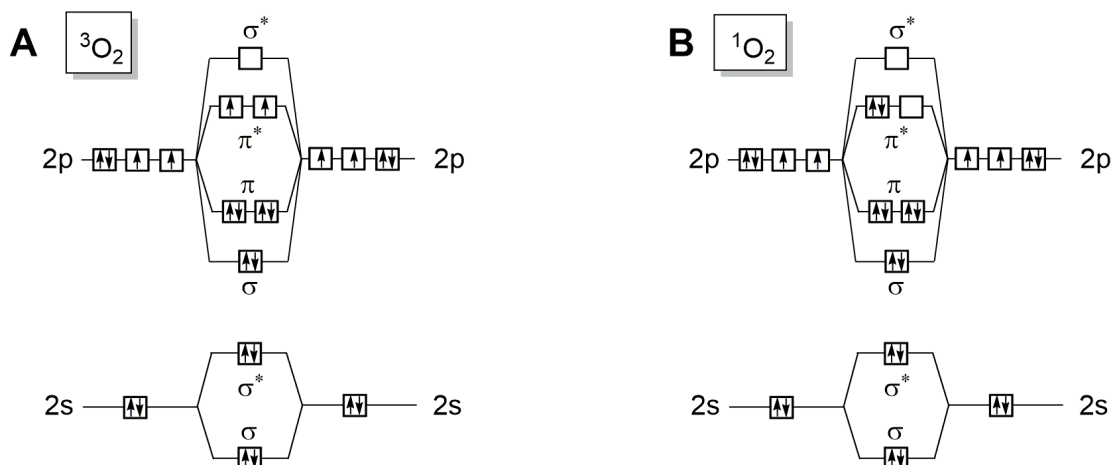
Scheme 1.17: Molecular orbital diagram of ground state oxygen $^3O_2$ (**A**) and excited singlet oxygen $^1O_2$ (**B**).

Organisms have evolved two parallel strategies to reduce O_2 by one/two-electrons transfer: one route involves redox active transition metals, principally iron and copper; the other strategy exploits organic cofactors such as nicotinamide adenine dinucleotide (phosphate) $(NAD(P)^+/NAD(P)H)$ and flavin adenine dinucleotide $(FAD/FADH_2)$ that act as electron sinks. O_2 is not the only acceptor of electrons: feeder enzymes containing Fe/S clusters are also involved in one electron transfer and can donate or accept electrons to/from the active site of monooxygenases/oxidases (Figure 1.10).

[2Fe-2S]   [3Fe-4S]   [4Fe-4S]
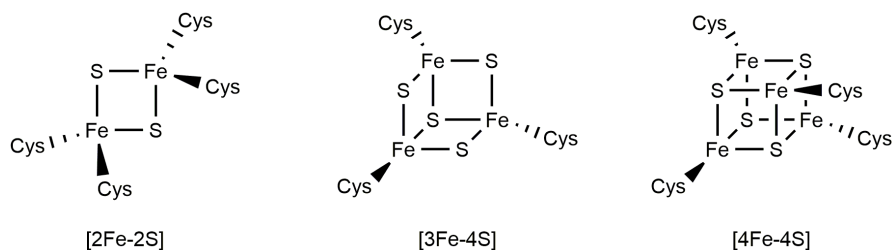
Figure 1.10: Different architectures of Fe/S clusters to accept/donate electrons in redox reactions.

Iron-dependent oxygenases

Iron dependent oxygenases can be subdivided into heme and non-heme oxidases.[113] They differ based on how the Fe(II) atom is coordinated in the catalytic center: it can lay in the equatorial plane of the heme, as the case of cytochrome P450 oxidases (Figure 1.11 left);[114] or it can be coordinated by the side chains of two *His* and one *Glu/Asp* residues

(Figure 1.11 right). In this last family, often α-ketoglutarate occupies the fourth and fifth coordination sites of the iron atom.[113, 115, 116]
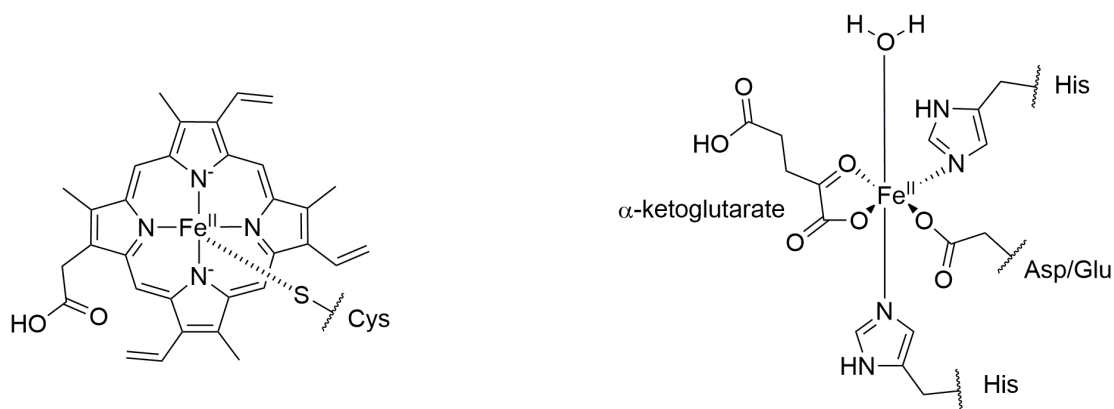
Figure 1.11: Heme-dependent oxygenase coordinate the iron atom in the equatorial plane of the heme (left); non-heme oxygenase coordinate the iron atom with residues in their catalytic site plus α-ketoglutarate (right).

The Fe(II) of P450 oxygenases binds O_2 at the top axial plane of the heme ring (Scheme 1.18 **A**). The Fe(II)-O_2 complex **128** is in resonance with the Fe(III)-superoxide **129**, where one e^- has been passed from the Fe(II) to the oxygen. Transfer of one electron from a feeder Fe/S cluster protein or from $FADH_2$, gives the Fe(III)-peroxide anion **130**, which is often used in Baeyer-Villiger ring expansions.[117] Cleavage of the weak O-O single bond with loss of water from **131** generates the highly valent oxo-iron cation radical $[Fe(IV)=O]^{\bullet+}$ **132**, which is in resonance with Fe(V)=O **133** oxo species. The radical cation is stabilised by the conjugated heme tetrapyrrole macrocycle. This high valent species is a strong oxidising agent, which can abstract hydrogen atoms ($H^{\bullet}$) from inactivated sp^3 carbons of the substrate, producing a carbon radical and Fe(IV)-OH (**133** to **136**, Scheme 1.18 **B**). Hydroxylation happens by $OH^{\bullet}$ rebound, or the carbon radical can be quenched by intramolecular rearrangement (**136** to **135**, Scheme 1.18 **B**).[118, 119]

Non-heme oxygenases follow a similar logic to P450 systems, with the generation of an oxo-iron Fe(IV)=O species (**137**, Scheme 1.19) by decarboxylation of the α-ketoglutarate to succinate to transfer electrons. **137** is also very reactive and can abstract one hydrogen atom from the substrate to generate $C^{\bullet}$, which can be quenched by $OH^{\bullet}$ (rebound) or by intramolecular rearrangements (**138** to **139**, Scheme 1.19).[115, 120, 121]
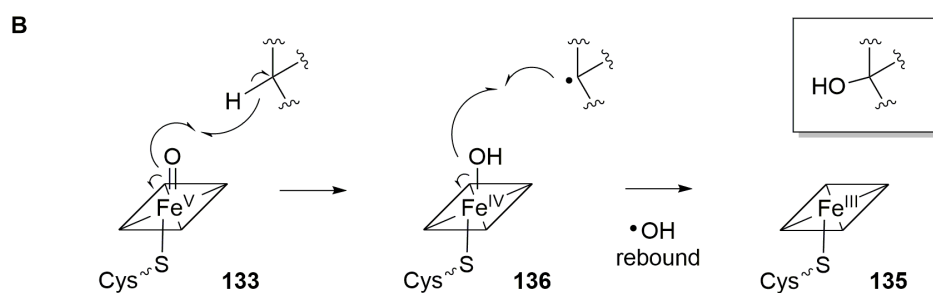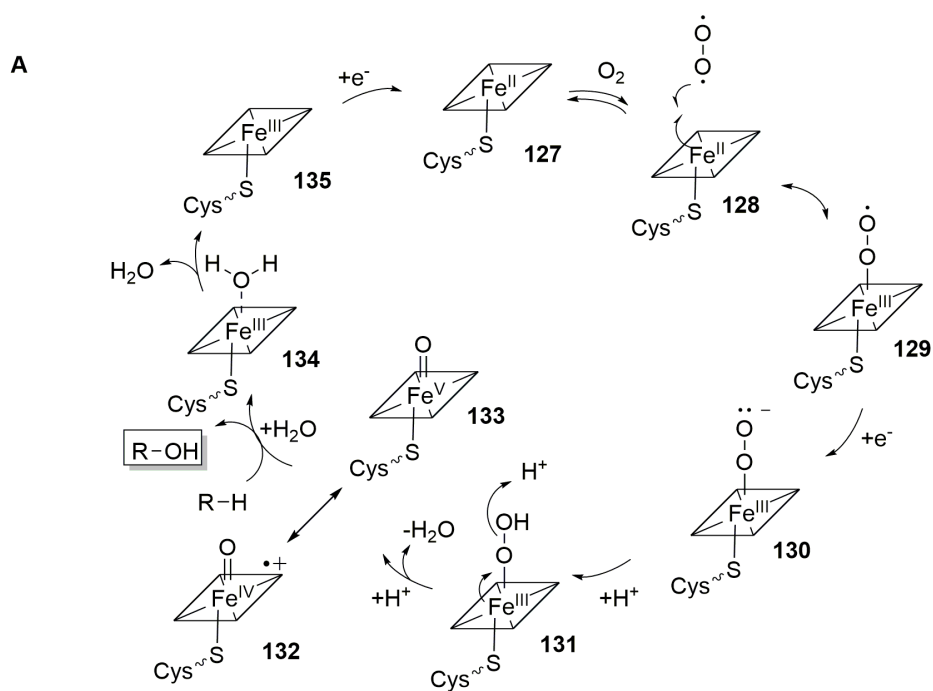
Scheme 1.18: Typical cytochrome P450 reaction cycle (**A**). Homolytic mechanism and OH$^\bullet$ rebound in heme-dependent oxygenases (**B**).
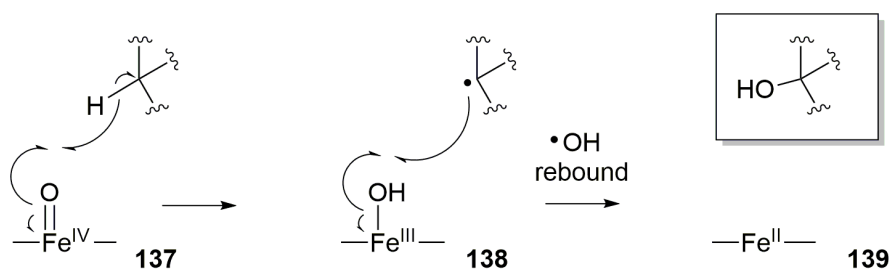
Scheme 1.19: Non-heme oxygenases mechanism.

Nicotinamide adenine dinucleotide and flavin oxidases/oxygenases

Another class of redox enzymes uses flavin (flavin adenine dinucleotide FAD and flavin mononucleotide FMN) and $NAD(P)^+/NAD(P)H$ as cofactors (**140**/**141**, **142**/**143** Scheme 1.20). Often flavoproteins also have a binding pocket for NADPH, as the latter is used to reduce FAD to $FADH_2$ (Scheme 1.21). In contrast to flavins, the nicotinamide cofactors NADH and NADPH can only transfer two electrons simultaneously because of the high energy of pyridinyl radicals (Scheme 1.20).[122] NADH and NADPH are diffusible coenzymes, while FAD and FMN are always enzyme bound.
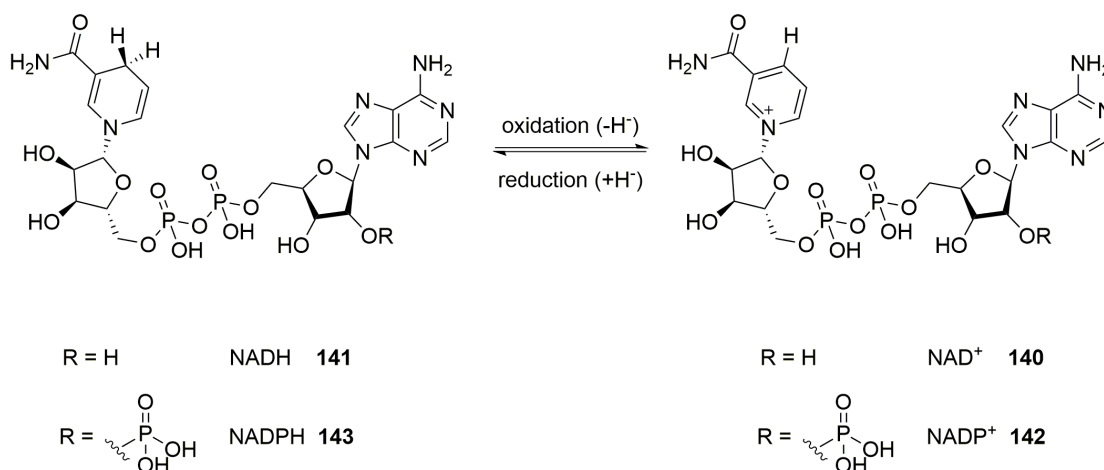
Scheme 1.20: A $NAD(P)H$ is the reduced form of $NAD(P)^+$. $NAD(P)H$ is the reducing agent *via* two electrons transfer by hydride (H^-).

A different strategy is used by vitamin B2-based flavin coenzymes as one-electron transfer platforms, notably FAD and rarely FMN (Scheme 1.21).[123, 124] The tricyclic isoalloxazine of oxidised FAD **144** can be fully reduced to $FADH_2$ **145** by transfer of two electrons from NADPH. **145** can donate one electron to molecular oxygen, yielding the stable semiquinone **146** and the anion O_2^- superoxide. **146** and O_2^- rapidly react to form FAD-4a-OO$^-$ **147**, which is in equilibrium with **148** by proton transfer. **148** is the hydroxy-donor cofactor that acts on nucleophilic centers generating hydroxylated substrate and FAD-4a-OH **149**. **149** spontaneously eliminates water to regenerate the fully oxidised FAD.
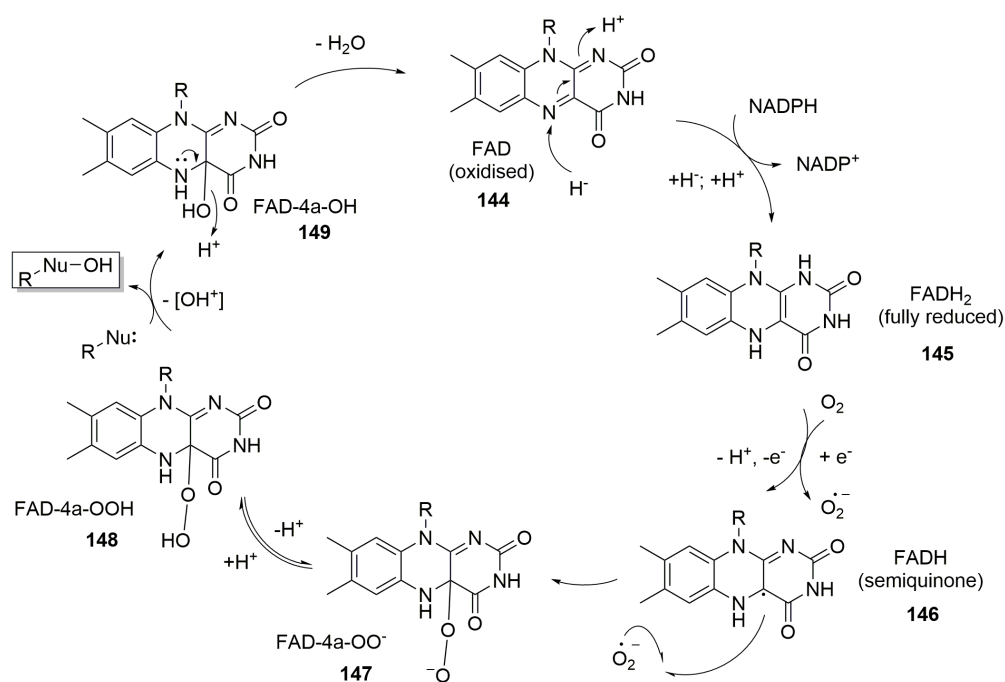
Scheme 1.21: Cycle of FAD/FADH$_2$ and hydroxylation of nucleophilic substrate.

1.6.2 *S*-Adenosyl Methionine

S-adenosyl methionine (SAM) **150** is a methyl group donor that plays a major role in primary and secondary metabolism.[125] It is the cofactor of specific methyltransferases (MeT) that can act upon nucleophilic centers of macromolecules and small molecules as well. An example of SAM in primary metabolism is found in the epigenetic code on histones tails: SAM-depented MeT act on lysines and arginines of nucleosome histones in order to compact the chromatin or to recruit DNA-binding proteins to suppress or activate gene expression.[126–128] There are numerous tailoring MeT that take part during the biosynthesis of natural products.[129–132] For exmple, *N*-methylation of putrescine **102** is a core step in the biosynthesis of tropane ring in alkaloid production (Scheme 1.14, Section 1.4).[93] Given the importance of SAM in biological processes, its usage and regeneration is conserved among all organisms (Scheme 1.22).[133] Nature has evolved two routes for SAM-methyl transfer: nucleophilic attack and radical mechanism.[134, 135]

SAM synthetase generates SAM by bringing together methionine and adenosine with consumption of ATP, forming the sulphonium cation (**151** to **150**, Scheme 1.22).[134] The electrophile sulphonium can be attacked by a nucleophile center to form a C-C bond in the catalytic site of a methyltransferase (**152** to **153**, Scheme 1.22). The consumption of SAM generates *S*-adenosylhomocysteine (SAH **154**), which is converted by SAH hydrolase to **155** and **156**.[133] Methionine is then formed with consumption of

5-methyltetrahydrofolate by methionine synthase (from **156** to **151**) and SAM is regenerated with consumption of ATP (**151** to **150**).[133]
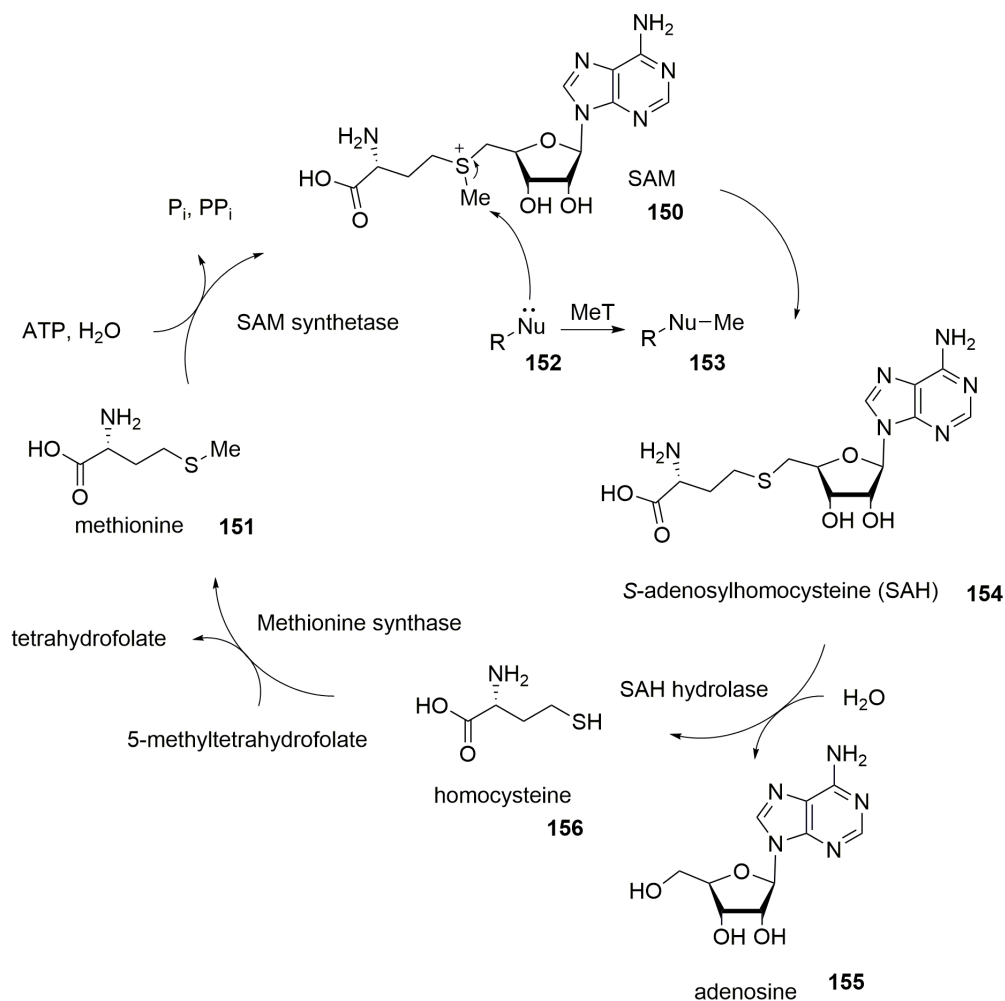
Scheme 1.22: Utilisation and regeneration of SAM. Methyl transfer involves a nucleophile center as methylation site.

Radical SAM is exploited in the methylation of inactivated centers such as aliphatic chains. The enzymes that perform transfer of $CH_3^{\bullet}$ groups have a conserved motif of three cysteines CX_3CX_2C.[135–137] The *Cys* motif coordinates a 4Fe/4S cluster, while SAM binds the fourth coordination site with the sulfur acting like a cysteine (**157**, Scheme 1.23).[135,137] The iron/sulfur inorganic clusters donate one electron to cleave the C-S bond between the methionyl moiety of SAM and its adenosinyl part, generating methionine and 5'-deoxyadenosyl radical (**157** to **159** and **158**). The consequent 5'-dA$^{\bullet}$ **158** can abstract a hydrogen atom (H$^{\bullet}$) from the substrate generating a radical carbon.[135,137]
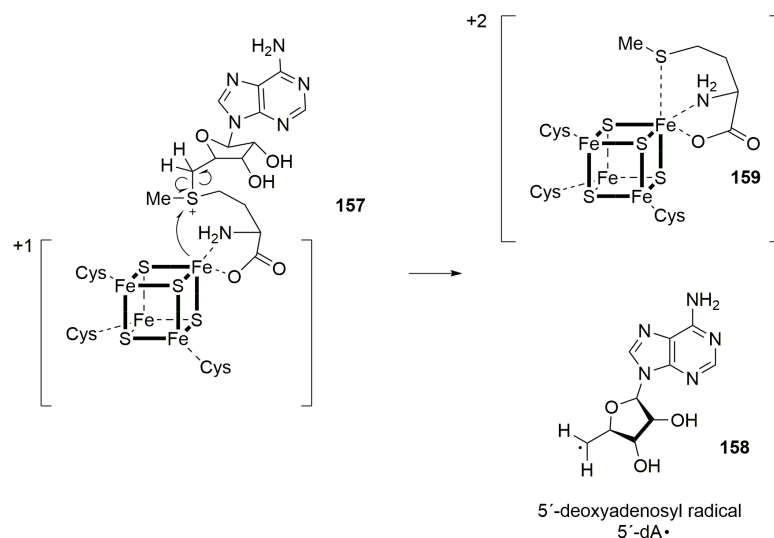
Scheme 1.23: Homolytic cleavage of SAM.

Transfer of $CH_3^\bullet$ requires consumption of two SAM molecules: one to generate the 5'-dA$^\bullet$ (**158**, Schemes 1.23 and 1.24) to abstract one H$^\bullet$ from the substrate generating **160** and **161**, and a second SAM as donor of $CH_3^\bullet$ (Scheme 1.24). Consumption of two SAM results in the formation of a methylated substrate **162**, adenine **161** and SAH **154**.[135,137]

First SAM: substrate radical generation
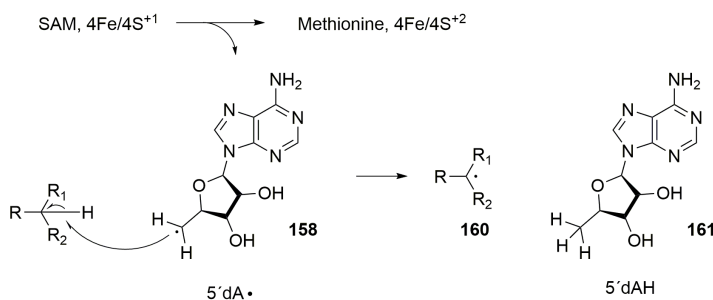
Second SAM: CH$_3^\bullet$ transfer
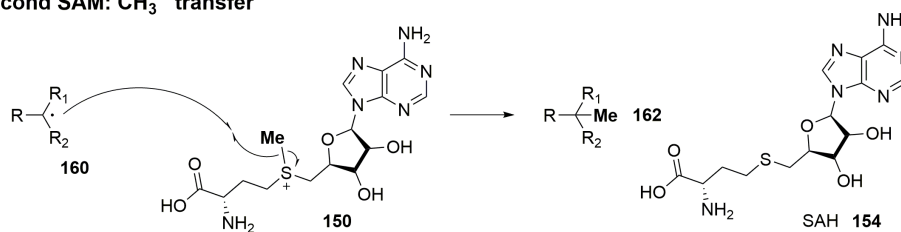
Scheme 1.24: SAM-dependent radical methylation.

1.7 Methods of Investigation

Briefly, the techniques and modern tools in the research of natural product biosynthesis are described.

1.7.1 NMR and isotopic feeding experiments

Nuclear Magnetic Resonance spectroscopy (NMR) is a powerful technique to determine the structure of small molecules. Exploiting different local environments of nuclear spin within a molecule in a strong electromagnetic field, NMR reveals atom connectivity and configuration. Simple 1D spectra reveal the presence of functional groups and their surroundings, while 2D HSQC, HMBC allow the determination of connectivity at short and long distance (up to 5-6 bonds). The environment of a nucleus can also be detected through space by Nuclear Overhauser Effect Spectroscopy (NOESY).

Feeding heavy precursors, such as radioactive $[^{14}C]$-acetate, to cell culture has been exploited in the past. Radiolabels are easy to detect, but determining the site of incorporation within the product was an elaborate route, that included chemical degradation and further radiodetection. Thanks to the advent of NMR, feeding building blocks labeled with spin-active isotopes gives deeper structural information with much less effort. The natural abundance of ^{13}C on Earth is 1.1%, meaning that a 1.1% incorporation of the ^{13}C-labeled precursor doubles the intensity of the relative signal. $^{18}O_2$ incorporation into keto- and hydroxygroups cause a slight shift of few decimal of ppm of the tethered C, by the effect of the greater mass of the oxygen isotope. Multiple labels such as $[1, 2 - {}^{13}C_2]$-acetate results in the splitting of the signals in a pair of doublets with identical coupling constant J, rendering easy to understand atom connectivity. Indeed, acetate is converted to malonyl-CoA by primary metabolism, therefore its incorporation in a polyketide chain will lead to two adjacent ^{13}C-rich positions. ^{13}C has a spin quantum number $I = \frac{1}{2}$, superimposed on the uncoupled natural abundance signals.
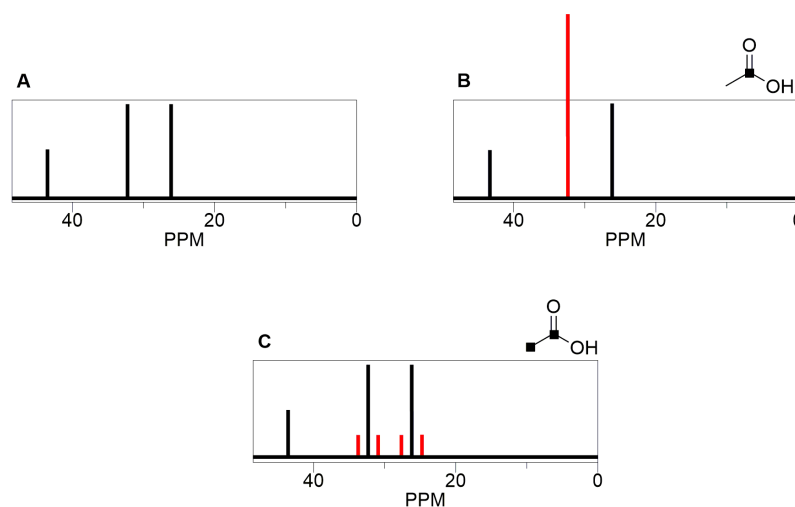
Figure 1.12: (**A**) ^{13}C-NMR spectra of a hypothetical acetate-derived compound. (**B**) Incorporation of a single-labeled building block leads to the enhancement of the corresponding carbon, (**C**) while double labeled acetate leads to C-C coupling superimposed to the natural abundance.

1.7.2 Genome sequencing and annotation

The rapid technological expansion in next generation sequencing (NGS) is making genomic information easier to obtain and more accurate with exponential advancement.[138–141] A common NGS system is represented by Illumina technology.[142] Illumina sequencing exploits the incorporation of fluorescent dNTPs polymerized into a template strand. At each cycle of oligonucleotide elongation the incorporated dNTP is recognized by fluorophore excitation with millions of PCR-amplified gDNA fragments sequenced in parallel.[143] Illumina operates in Paired-End (PE) sequencing. PE involves sequencing both ends of the DNA fragments in a library and aligning the forward and reverse as read pairs. Because the distance between each paired read is known, alignment algorithms can use this information to map the reads over repetitive regions more precisely. This allows more accurate read alignment and the ability to detect insertions and deletion events, which is not possible with single-read data. Analysis of differential read-pair spacing also allows removal of PCR duplicates, a common artifact resulting from PCR amplification. Moreover, PE sequencing facilitates detection of genomic rearrangements such as insertions, deletions, inversions, and gene fusions.[144–146]

Genome sequencing produces massive amounts of raw data that need to be assembled. DNA fragments are aligned and assembled to form a *contig* (contiguous reads, ∼20-50 KBp). Analogously, contiguous reads are joint together to form a scaffold (∼0.1-1 MBp). The fewer and larger the scaffolds to cover a whole genome, the better the quality of the sequencing. An indicative value to determine the quality of *de novo* sequencing is the

N_{50} scaffold value: the median contig size of the assembly. A small value of N_{50} scaffold means that few contigs of significant size were generated.[147,148]

A tool for genome annotation specific for secondary metabolism is AntiSMASH.[149,150] Normally, in bacteria and fungi the genes encoding the enzymes necessary for the biosynthesis of natural products are grouped into biosynthetic gene clusters (BGC).[151–155] This is not true for plants, that generally have genes for secondary metabolism scattered throughout the whole genome,[151,156] although gene clusters in plants have been reported.[157]

In the work presented, fungal metabolites have been investigated and their related BGC have been deduced using bioinformatical tools, including AntiSMASH (Chapters 2, 3).

1.7.3 Genome manipulation in fungi

Knock out (KO) is a process that lead to the loss of functionality of a gene. Targeted KO is a route to determine the role of a single gene during the biosynthesis of a given compound. Indeed, the loss of a tailoring enzyme interrupts the synthetic pathway, often resulting in the accumulation of the related substrate. This shows *in vivo* potentially each step of the chemistry involved in the formation of the product. Although this technique is rather clean and indicative, it can also happen that intermediates are rapidly shunted or degraded by the cell, frustrating any result, or making their interpretation difficult.

An efficient technique to edit fungal genomic DNA has been described by Nielsson and co-workers: the bipartite recombination.[158] It is a precise site-directed tool that exploits the natural occurring recombinases present in filamentous fungi and other higher eukaryotes, that allows gene deletions, promoter replacements, in-frame GFP fusions and specific point mutations. Transformation of two overlapping PCR products carrying a disrupted selection marker flanked by two target homology regions, leads to three distinct events of recombination that restore the selection marker into the *locus* selected by the target homology regions (Scheme 1.25). Splitting the selection marker in two parts reduces the number of false positive. Indeed, the KO construct, transformed as a whole, might be inserted randomly into the genome, inferring selectivity. By increasing the number of needed recombination events to three, random selection is drastically diminished.
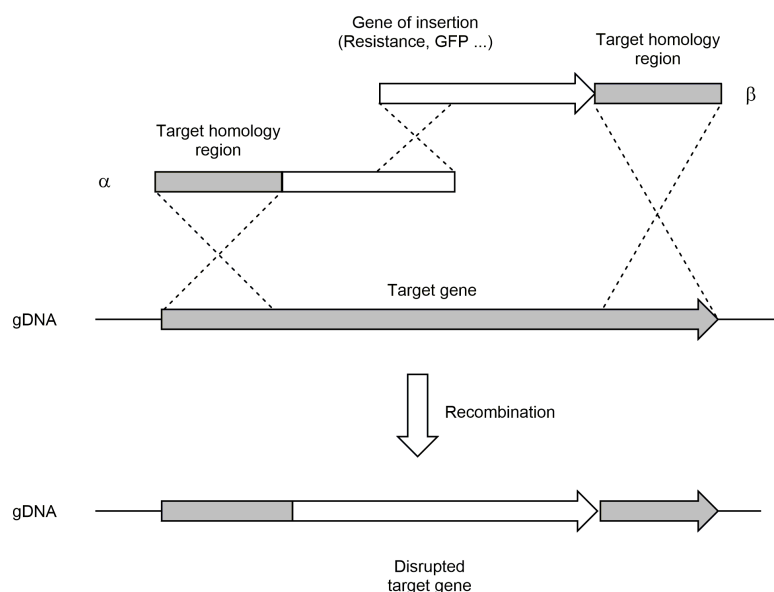
Scheme 1.25: Recombination at homology regions to disrupt the target gene.

Other methods worth mentioning to introduce exogenous DNA into a host are transformation by *Agrobacterium tumefaciens*[159, 160] and CRISPR/Cas directed genome editing.[161, 162] *A. tumefaciens* is a bacterial plant pathogen that cause tumors in the host by introducing and integrating a segment of DNA (namely T-DNA). CRISPR/Cas exploits the bacterial immune system against virus infections to target and edit specific DNA sequences by using RNA guided endonucleases. Examples of successful fungal genome editing by *A. tumefaciens* and CRISPR/Cas have been reported.[163–168]

1.7.4 Fungal transformation

Transformation is the insertion of exogenous DNA into a host. Bacteria are relatively simple to transform, while fungi and plants need to be pretreated in order to remove their cell wall.

The fungal cell wall consists of glycoproteins and polysaccharides; mainly mannoproteins, β-1,3-glucan and chitin (Figure 1.13). The glycoproteins of the cell wall are post-translationally modified with both N- and O-linked carbohydrates such as mannose and often glycosylphosphatidylinositol (GPI) anchors. The composition of fungal glucan cell walls varies among species, presenting also alternative linkages, such as β-1,6-glucan. Chitin is found as chains of β-1,4-linked N-acetylglucosamine residues and is normally the minor entity of the wall. The glycoprotein, glucan and chitin components are extensively cross-linked together to form a complex network.[169, 170] Such a composition of the cell wall differs from the one found in bacteria and plants, being unique to the kingdom of fungi.
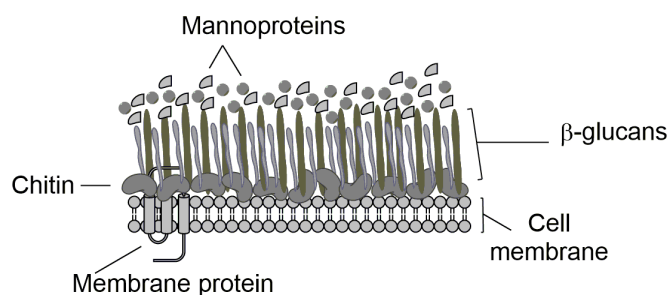
Figure 1.13: Fungal cell wall.

The wall is a dynamic structure, which needs to be constantly modeled by the fungus during its growth and in response to stress factors. Most of the fungal cell wall hydrolases known today have cellulase, chitinase and protease activities with some showing transglycosylase effects.[171] To transform fungal cells, the cell wall must be digested with hydrolytic enzymes to allow the penetration of exogenous DNA. Concentrated mixtures of lysing enzymes extracted from *Trichoderma harzianum* and basidiomycota are effective in the total digestion of the cell wall. The resulting digested single cell, namely a protoplast, is particularly sensitive to osmotic pressure, thus needs to be preserved with saline buffer whose concentration ranges accordingly to the specific fungal species. On the other hand, protoplasts are rather easy to transform with exogenous DNA.

1.7.5 Heterologous expression

Heterologous expression allows the reconstruction of partial or entire metabolic pathways in particular hosts (bacteria, fungi, plants, animal cells) by transforming the host with a set of genes under the control of constitutive or inducible promotors.[172] *E. coli* is a widely used host for protein production and purification, but it is not ideal for expressing fungal or animal proteins, as it cannot process eukaryotic introns, needs codon optimisation and show difficulty in folding fungal polypeptides.[173,174] *Aspergillus oryzae*, *Aspergillus nidulans* and *Saccharomyces cerevisiae* are fungal hosts more suitable for heterologous expression of fungal pathways.[175–177] Usually, auxotrophic strains are transformed with vectors carrying the genes of interest plus the genes to reintroduce autotrophy, allowing an efficient selection of the transformed cells. For example, the biosynthesis of tenellin **197** was reconstituted by expressing four genes (*tenS, tenA, tenB, tenC*) responsible for the production of tenellin in arginine auxotrophic *Aspergillus oryzae* M-2-3 strain.[173] Reconstitution of larger biosynthetic pathways have been achieved in *A. oryzae* NSAR1, an auxotrophic strain for arginine, adenine, methionine and nitrate, by simultaneous expression of up to 20 genes.[178,179] Heterologous expression not only allow the verification and elucidation of entire secondary metabolite pathways, but it is also a tool to enhance the production *in vivo* of interesting compounds otherwise difficult to obtain in large amounts in the natural producer or by totals synthesis (Scheme 1.26).
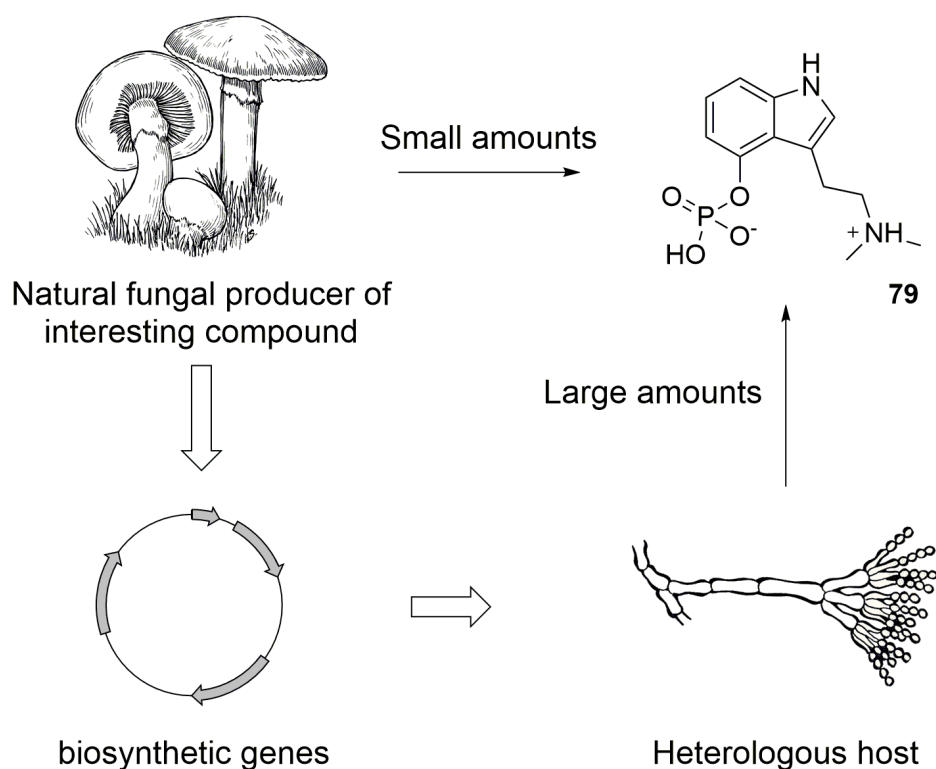
Scheme 1.26: Production of interesting compounds in optimised host. Credit to P. S. Foresman and E. Owen for mushroom and mycelium drawings respectively.

1.8 Fungal metabolites

Today more than 100.000 fungal species are known, although far more than 1.5 million are expected.[180,181] This great diversity of the fungal kingdom reflects the enormous potential of fungal secondary metabolites to improve human life. Fungi are extraordinary producers of natural products including antibiotics, mycotoxins and pharmaceuticals.[3,182] They have existed on Earth for at least one billion years,[183] evolving secondary metabolism for the production of bioactive compounds (Figure 1.14). These compounds include deadly toxins such as phalloidin **163** and α-amanitin **164** produced by the death cap *Amanita phalloides*; psychoactive compounds such as muscimol **165** (*Amanita muscaria*), psilocybin **79** (*Psilocybe cubensis*), lysergic acid **126** (*Claviceps purpurea*) and xenovulene A **124** (*Acremonium strictum*); pharmaceuticals such as the β-lactams **62**, **166** and the statin lovastatin **22**. Moreover, fungi are between the few organisms that produces antimycotic compounds such as strobilurin A **167**.[184–186] These compounds illustrate the utility and diversity of chemical structures produced by fungi, and the diverse biosynthetic potential of these organisms.
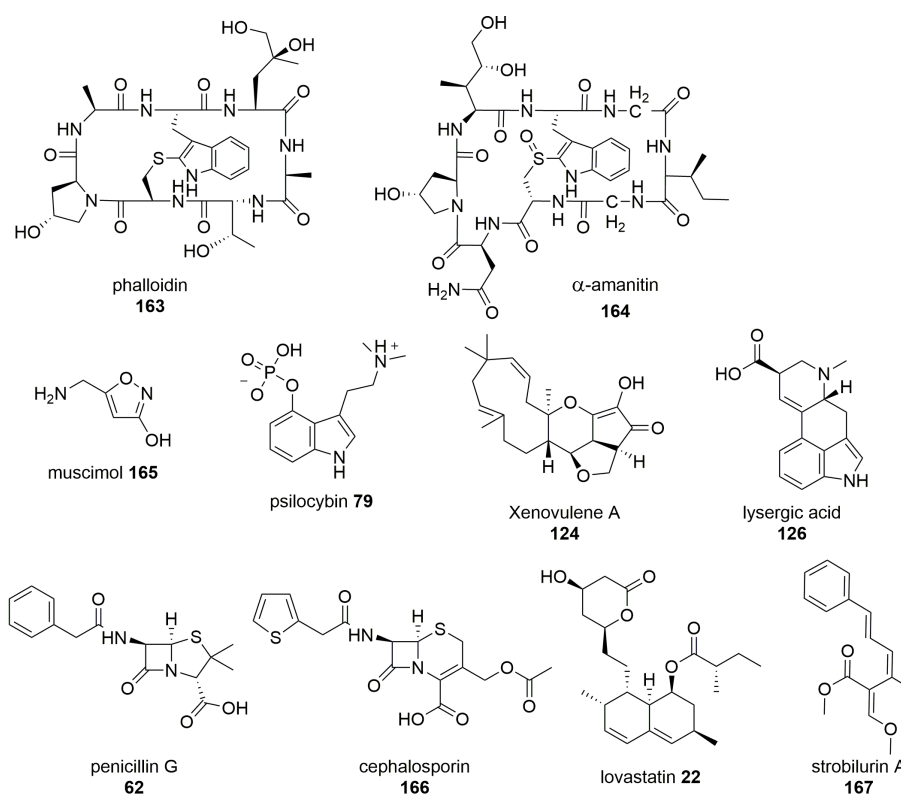
Figure 1.14: Examples of fungal metabolites with impact on human life.

In a world where antibiotic resistance is exponentially rising;[187–189] where new tools for an agriculture revolution are required;[190] and the need of new therapies make the study of fungal natural compounds a key field for human survival and evolution.

1.9 General aims of the thesis

In the work presented we focused the two natural compounds phyllostictine A **168** and Sch-642305 **169** (Figure 1.15) isolated from filamentous fungi.[191,192] Phyllostictine A **168** is a phytotoxin isolated from the plant pathogen *Phyllosticta cirsii* with a unique tricyclic structure. It cannot be assigned to any natural compound class, which might imply a new biosynthetic strategy in fungi, with the formation of a β-lactam connected to a dihydrofuranol ring, both fused to a 11-membered ring. A structural analysis by full NMR and isotopic labeling is necessary to confirm such an unprecedented system, and to reveal which building blocks take part during its biosynthesis. Genome mining of the producer might reveal the enzymatic arsenal needed for the biosynthesis, with the possibility of discovering new enzymes able to perform uncommon chemistry.

Sch-642305 **169** is a decalactone isolated from *P. verrucosum* and the endophyte *Phomopsis CMU-LMA*. It has been investigated in the past by numerous chemists because of its potent cytotoxicity and antiviral bioactivity against HIV. Although its total synthesis has been achieved following different routes, the details regarding its biosynthesis have been completely overlooked, besides a biomimetic synthetic approach by Snider and collaborators.[193] We aim to reveal the molecular mechanism of each chemical step *in vivo*, with specific focus on the formation of the 6-membered ring, that has been proposed to be formed by Michael cyclisation.[194] The biosynthetic gene cluster for **169** production is unknown, so we plan to find it by *in silico* prediction and confirm it by targeted knock out (KO).
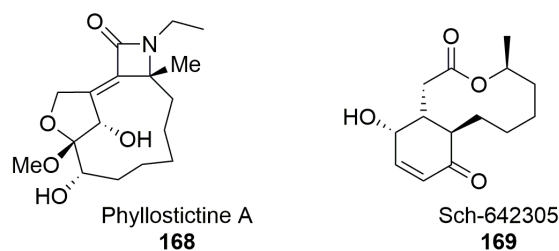
Figure 1.15: Phyllostictine A **168** and Sch-642305 **169**.

The understanding of the molecular basis for **168** and **169** production sets the possibilities for metabolic engineering, either by KO or heterologous expression, in order to produce new analogues whose biological activity can be tested.

2 Structural Revision and Biosynthesis of the Fungal Phytotoxins Phyllostictine A and B from *Phyllosticta cirsii*

2.1 Introduction

Cirsium arvense, known as Canada thistle, is a perennial weed that colonises many habitats such as fields, pastures, rangeland and gardens, causing losses in agriculture all around the world.[195,196] Originally from Europe, Western Asia and Northern Africa, this weed has spread worldwide to become a major pest in agriculture: in cereal crops for example, Canada thistle causes up to 30% loss in grain yield, with a global annual loss estimated at 320 million US $.[197,198] *Cirsium arvense* diffuses by spreading seeds carried by the wind and by clonal propagation *via* its root system. The latter is its most important and effective method of colonization, with the formation of dense patches up to a hundred shoots per square meter, supported by root systems of several hundred meters in length.[199] Roots survive the winter and can restart the cycle even from small parts.[200] In large scale farming, herbicides are commonly used for *Cirsium arvense* control, but they all fail to kill the roots, forcing a continuous usage of herbicide rendering this procedure polluting and expensive.[201] Modern remedies include the chlorinated compounds 2-methyl-4-chlorophenoxyacetic acid (MCPA) **170** and picolinates such as clopyralyd **171** and fusaric acid **172** (Figure 2.1) to control weed pests with reduced harm for valuable plants such as corn and wheat.[202]
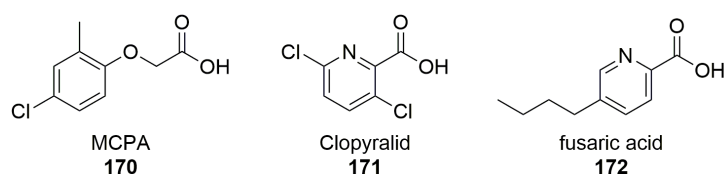
Figure 2.1: Pesticides used in modern agriculture.

MCPA is a synthetic compound first obtained by Synerholme and Zimmerman in 1945 and by Templeman and Foster in 1946 in the attempt to target the auxin plant hormone.[203] MCPA is effective against broad leaf plants including thistle and dock, but it also harms a broad range of deciduous trees and clovers.[203] Although MCPA is scarcely

toxic to mammals (rat LD_{50} ranges from 0.7 g/Kg to 1.3 g/Kg),[204] it can form complexes with Pb^{2+}, Cd^{2+}, and Cu^{2+}, enhancing its half-life in the soil.[205] Clopyralid is one of the few effective pesticides for the control of Canada thistle; it belongs to the picolinate family, which includes **172**. Clopyralid has a broad spectrum of targets: it is not only toxic to weeds, but also to edible plants such as peas, tomatoes and sunflowers; moreover it accumulates in dead plants and compost to phytotoxic levels and spoils potatoes, lettuce, and spinach.[206]

Because of the mild selectivity of these compounds against the weed pests and their prolonged retention in the ground, many countries have restrained MCPA and clopyralid to a limited use, meaning that research towards new solutions is required. Biological control is an alternative strategy that has been proposed to neutralize Canada thistle. The definition of biocontrol agents is broad and ranges from usage of whole or partial organisms or purified compounds. Advantages of this approach include lower persistence in the environment and a narrower range of bioactivity.[207] Specific insect and microorganism pathogens of the plant have been researched. For example, *Aceria anthocoptes* (Figure 2.2) is a gall mite belonging to the *Eriophyidae* family, which could act as control agent of *Cirsium arvense*, as it damages the epidermal cells and the deeper mesophyll layer of its leaves.[208] Microorganisms however, are the most promising candidates as bioagents. Fungal phytotoxins provide new basis in herbicide discovery: pathogens of *Cirsium arvense* may be the source of specific and transient compounds to eradicate the noxious weed without harming crops and the environment. A total of ten fungal species and one bacterium have been investigated as possible agents to target *Cirsium arvense* (Table 2.1).[195]
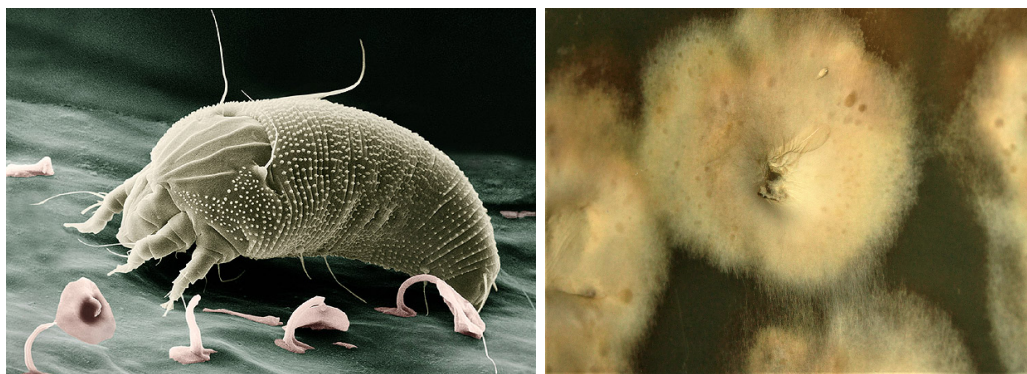
Figure 2.2: proposed biocontrol agents against Canada thistle: *Aceria anthocoptes* (image by E. Erbe; digital colorization by Chris Pooley) and *Phyllosticta cirsii* (original picture).

Pathogen	Target organ	Effectivity	Selectivity
Puccinia punctiformis	leaves, shoots	limited, local	very high[209–211]
Phomopsis cirsii	stems, leaves, roots	high	high[212,213]
Sclerotinia sclerotiorum	stems and leaves	limited, local	very low[214,215]
Alternaria cirsinoxia	leaves	limited	low[216,217]
Phoma destructiva	dead and living plant	high	unclear[218,219]
Phoma exigua	leaves	inconsistent	very low[220,221]
Stagonospora cirsii	leaves	high, with restrictions	low[222,223]
Septoria cirsii	leaves	high	very high[224]
Phyllosticta cirsii	unknown	unknown	unknown[191,225,226]
Fusarium sp.	seeds, seedlings, leaves, roots	inconsistent	low[227]
Pseudomonas syringae pv. tagetis	leaves, shoots	high	low[228,229]

Table 2.1: *Cirsum arvense*'s pathogens proposed for biocontrol.[195]

The fungus *Phyllosticta cirsii* has recently been proposed by Evidente and Vurro as a source of mycoherbicide compounds against Canada thistle. An unprecedented class of oxazatricycloalkenones known as phyllostictines A-D **168**, **173** - **174** have been isolated from this fungus and their structure and bioactivity have been characterized by Evidente *et al* (Figure 2.3).[191,226]
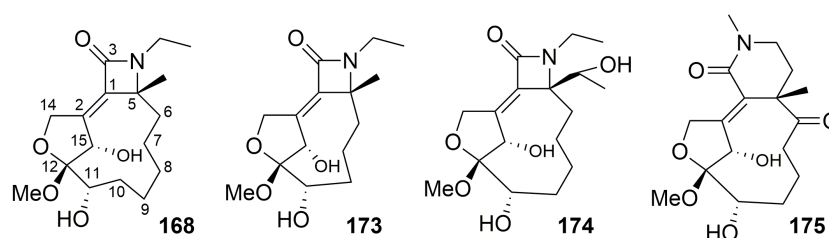
Figure 2.3: Phyllostictine A-D isolated from *P. cirsii.*

Phyllostictines are phytotoxic metabolites which are produced during the infection of the host: leaf-puncture assays (20 μL/droplet) on Canada thistle have showed different activity across the various phyllostictines at concentration of 6 mM. In particular, phyllostictine A is highly phytotoxic, causing rapid insurgence of large necrotic spots. Phyllostictines B and D present milder toxicity compared to the main metabolite, whereas phyllostictine C is almost non toxic (Table 2.2).[226]

Compound	∅ necrosis /mm
168	> 5
173	< 5
174	0
175	< 5

Table 2.2: Phytotoxicity of phyllostictines in puncture assay on thistle leaves.[191]

Phyllostictine A **168** induced high mortality in thistle protoplasts compared to fusaric acid and the herbicide glyphosate.[226] At the highest concentration (1 mM) **168** kills almost all protoplasts after one hour incubation. At 0.5 mM the toxicity is time dependent, killing around 50% of the population after one hour and reaching 100% after six hours, while lower concentrations do not affect protoplast viability. Fusaric acid **172** has a lower efficiency, being time dependent already at 1 mM with 60% mortality after one hour and reaching 100% in six hours. The glyphosate control was tested at 4.2 mM (commercialized concentration) killing the totality of protoplasts three hours after the treatment, while serial dilutions 1:10; 1:100 reduce its potency by 40%. Similar results have been achieved using tobacco protoplasts, suggesting that phyllostictine A has a broad toxicity, instead of a specific target.[226] Antimicrobial and zootoxic activities have been assayed for phyllostictines A and B as they could be easily purified in reasonable amounts. Antimicrobial assays showed Phyllostictines A and B to have no fungicidal activity against *Geotrichum candidum* up to 100 mg/disk. Against Gram-positive *Lactobacillus sp.* **168** is active at 5 mg/disk, but none of the compounds have an effect against Gram-negative *E. coli.* Tested on brine shrimp (*Artemia salina*), **168** causes noticeable larvae mortality at 1 mM.[226] These results are based only on a limited set of organisms and cannot be generalised, and until further data become available phyllostictines cannot be regarded as suitable biocontrol agents of Canada thistle.

Partial synthesis of **168** has been proposed by Coe and Shipman using a ring closing metathesis reaction (**176** to **177** Scheme 2.1). A simplified analogue displayed phytotoxic activity against *Chlamydomonas reinhardtii*, suggesting that the α-methylene β-lactam subunit is responsible, at least in part, for the herbicidal activity of phyllostictine A.[230] However, no completed synthesis of these compounds has been reported.
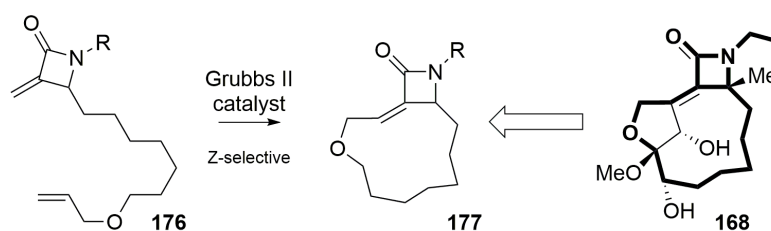
Scheme 2.1: Synthetic strategy towards the core structure of **168** by Coe and Shipman.[230]

The tricyclic phyllostictines belong to an unique category of secondary metabolites. Nat-

urally occurring β-lactams can be subdivided into four main families: the NRPS products penicillins/cephalosporins **62**, **166**; the clavams **178** synthesised by β-lactam synthetase; the carbapenems **179**; and the monocyclic β-lactams **180** whose origin is still obscure (Figure 2.4).[231]
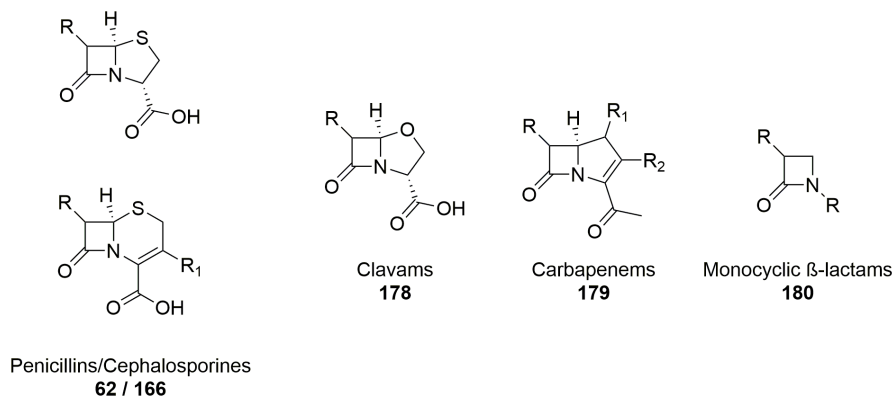
Figure 2.4: Classes of natural occurring β-lactams.

Phyllostictines A-D do not fall within any of these categories. They have an unprecedented β/δ-lactam ring connected through a double bond to a hydroxylated tetrahydrofuran and fused to an aliphatic macrocycle. Judging by the carbon skeleton, they are unlikely to derive either from terpene nor alkaloid biosynthesis, as they do not present either the characteristic isoprene units pattern or contain basic nitrogen atoms. The presence of N-ethylation across phyllostictines A-C is an unique feature for this class of compounds, since no N-alkyl transferase have been reported so far. The nitrogen embedded in the β/δ-lactam of **168** may derive from an amino acid, with C-5 being necessarily the α carbon of the reduced carboxylic acid. Following this hypothesis, the C-5-methyl should not derive from SAM, being also part of an amino acid, along with C-6. If the nitrogen, C-5, C-5-Me and C-6 are derived from alanine, or reduced serine, it could be possible that the remaining carbons of the core backbone derive from a pentaketide, with the exception of C-14, which is a highly unusual branching, as acetate-derived C-1 positions are not normally methylated. In this scenario, the mechanisms that lead to the formation of the β-lactam are hard to imagine. Judging from the structure, the most advisable pathway would be PKS-NRPS biosynthesis, especially because of the amide and the long alkyl chain. Since no known biosynthetic pathway seems to completely fit phyllostictine formation, there is the hypothesis that they might arise from a new type of biosynthetic processes in fungi. The structure of phyllostictines needs to be confirmed, to make sure that these compounds are true natural compounds.

2.2 Aims

Since no biosynthetic investigations have been reported so far for the unique class of macrolactam phyllostictines, we plan to unravel their biosynthesis, with special regard to the major compound phyllostictine A **168**. The methods to understand the biosynthesis involve traditional labeling with stable ^{13}C precursors in order to obtain structural information, and modern genome mining methods aimed to link the product to gene activity. Finally, molecular biology tools will allow knockout experiments, heterologous expression and *in vitro* enzyme assays, to characterise the role and function of the genes and their respective encoded proteins.

Feeding ^{13}C-labeled precursors will reveal the origin of the building blocks needed in the biosynthesis. The nitrogen embedded in the β-lactam may stem from an amino acid, but its origin is an open question. For this reason, we plan to feed [1, ^{13}C]-alanine to test if the C-5-methyl group could be the lateral chain of this amino acid. Feeding [*methyl*, ^{13}C]-methionine should enrich exclusively the methyl ether, but not the C-5-methyl. Supplying ^{13}C-labeled acetate is also source of valuable information, as the alkyl chain is probably the product of a highly reducing PKS activity, possibly with C-C bond rearrangements that could take place during the formation of the tetrahydrofuran moiety.

Total genome sequencing and mining tools coupled to manual annotation will point at the putative BGC linked to phyllostictine production. Software of choice is AntiSMASH,[149] which has proved to be highly efficient in cluster recognition and classification. Manual annotation to corroborate intron prediction, sequence homology, domain analysis and protein alignment, is done by online browsers such as Softberry, NCBI BLAST interface, Artemis Comparison Tool and Augustus and Clustal Omega.[232–238] In the research of the core enzymes involved in the biosynthesis, special regard is held for nitrogen processing enzymes, such as highly reducing PKS-NRPS hybrid system or pyridoxal-phosphate dependent decarboxylases. The likely presence of a SAM-dependent *O*-methyl transferase within the cluster will also be indicative to justify the presence of the methoxy group.

The putative BGC can be confirmed by knock out experiments. Indeed, disruption of the core enzyme should lead to complete abolition of **168** and minor compounds. *In situ* specific knock out of tailoring enzyme-encoding genes, should both prevent phyllostictine biosynthesis and also lead to accumulation of intermediates related to the interrupted chain of reactions, disclosing the function of potentially each enzyme involved in the biosynthesis. In particular, we aim to understand how the α-methylene β-lactam moiety is formed, and the mechanism for the synthesis of the doubly oxidised tetrahydrofuran subunit.

2.3 Results

We obtained the fungus *Phyllosticta cirsii* as gift from Dr. Maurizio Vurro (*Istituto di Scienze delle Produzioni Alimentari*, CNR, Bari, Italy), where it was stored in the mycological collection (ITEM N.8964). Dr. Vurro received in turn the organism from Dr. Alexander Berestetskiy (Department of Phytotoxicology and Biotechnology, All-Russian Research Institute of Plant Protection, Saint-Petersburg, Russia). *P. cirsii* was isolated from infected Canada thistle around 20 years ago, but the details of the original culture have been lost.

2.3.1 ITS characterisation

The sequence of the ITS (Internal Transcribed Spacer) is commonly used to determine species relationship in fungi.[239] Genomic DNA was isolated from *P. cirsii* using the Sigma GenElute Plant Genomic DNA Miniprep Kit (Aldrich).[240] ITS sequences were obtained by PCR using standard ITS1 and ITS4 primers and sequenced. The obtained ITS was blasted into NCBI database using nBLAST (nucleotide Basic Local Alignment Search Tool). Clustal Omega with default settings was used to build a phylogenetic tree from a collection of 20 ITS sequences from different known organisms, including *Phyllosticta* species, the closest organisms found by default NCBI blast search included: *Phoma spp, Phaeosphaeria spp* and *Leptosphaeria spp*. The phylogenetic tree also included the distantly related *Aspergillus oryzae* and *Saccaromyces cerevisiae* as outliers. Surprisingly, the ITS alignment showed a probable misidentification of the plant pathogen. Instead of clustering with other *Phyllosticta spp.* the organism showed phylogenetic proximity to *Phaeosphaeria spp* (Figure 2.5). At this stage of the work we did not characterise the organism precisely, and we continued referring it as *Phyllosticta cirsii*.
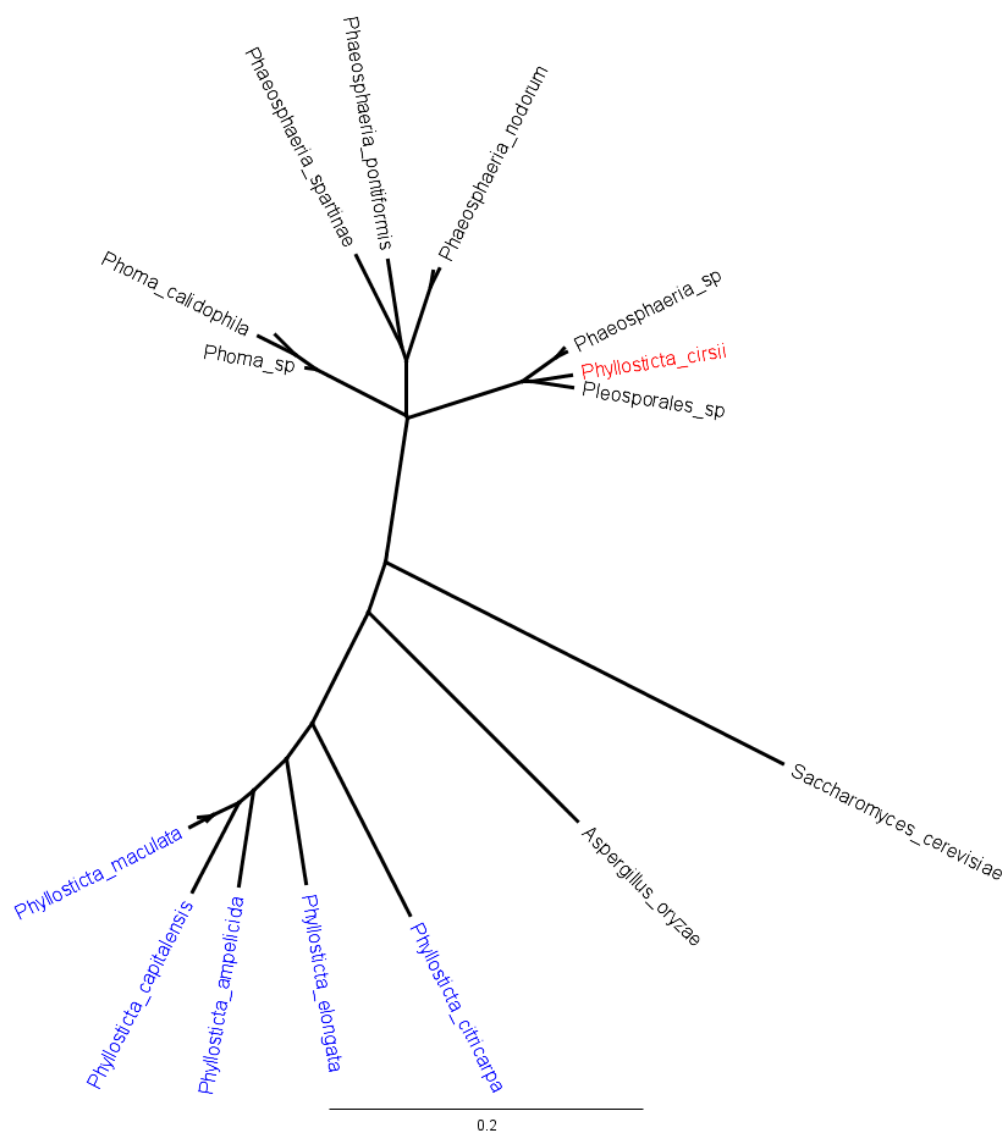
Figure 2.5: Unrooted ITS-based phylogenetic tree. *Phyllosticta spp.* are shown in blue. The ITS sequence of the organism known as *Phyllosticta cirsii* is shown in red. Clustal Omega with default settings was used to calculate the phylogenetic distances.

56 secondary metabolites have been described in the literature as being isolated from *Phaeosphaeria spp* (Reaxys[241]). Focusing on the ones with odd molecular weight, thus containing an odd number of nitrogen atoms, four compounds had a similar molecular weight as **168**: phaeosphaeride A **181** and paraphaeosphaerides A-C **182-184** and **185** (Figure 2.6). Interestingly, none of the found compounds were related to phylostictine.

Phaeospheride A
181

Paraphaeospheride A
182

Paraphaeospheride B
183

Paraphaeospheride C
184

185
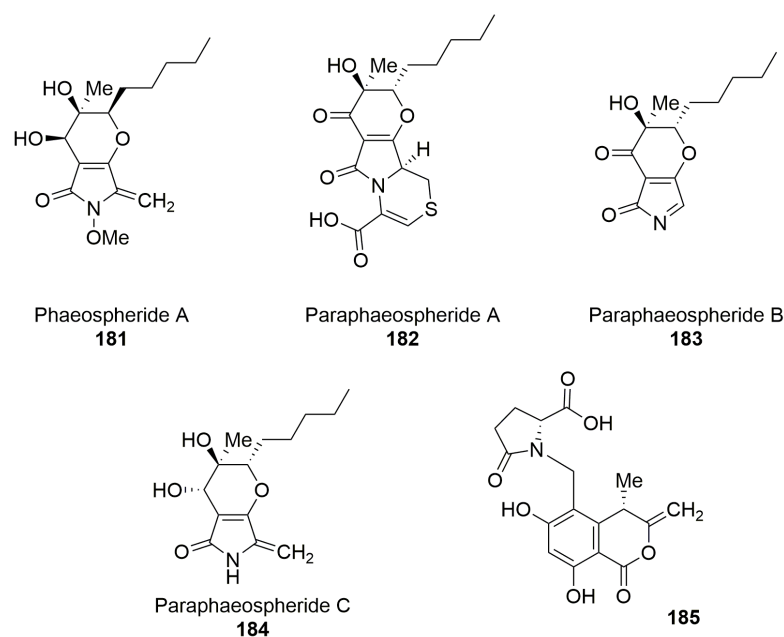
Figure 2.6: Nitrogen-containing compounds produced by *Phaeosphaeria spp* found by Reaxys search.

2.3.2 Phyllostictine A and minor compounds harvesting

Phyllosticta cirsii grew slowly in rich and poor media, with mycelia that varied their pigmentation from yellow to pale pink (Figure 2.8). Literature conditions for phyllostictine A production suggest growth in darkness in static defined M1D medium at 25-28 °C for 28 days.[191] In particular, stationary cultures permitted floating growth of the fungus on the surface of the medium. Solid fermentation on M1D agar plates was also suitable for phyllostictine biosynthesis (Figure 2.8), while shaken PDB medium at 28°C led to production of scytolide **186** (Figure 2.7) after one week and minor accumulation of **168**.

scytoltide
186
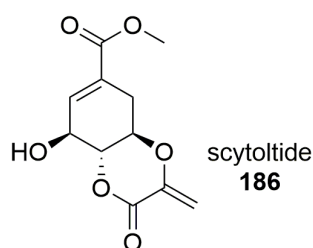
Figure 2.7: Scytolide **186** obtained as major compound in PDB fermentation.

Crude extractions were performed by blending the liquid culture, filtering the cell material and the supernatant treated with ethyl acetate in 1:1.5 volume ratio. Ethyl acetate was dried by MgSO$_4$ and evaporated under vacuum. The crude extract was either stored at -20 °C or dissolved in acetonitrile (10 mg/ml) for LCMS analysis and purification.
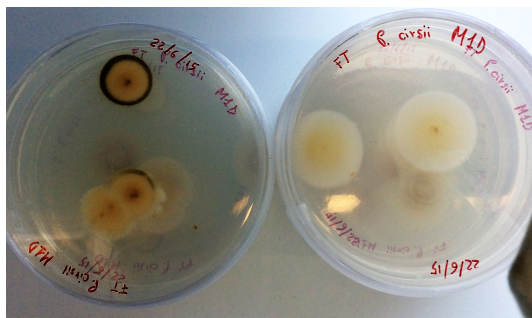
Figure 2.8: *P. cirsii* growing on chemically defined M1D agar medium.

The fungus consistently produced compounds with masses corresponding to the phyllostictines when cultivated in static M1D medium for two weeks at 28 °C, although **173** and **175** were present only in trace amounts (Figure 2.9). Static growth was an important factor in the production, as shaken flasks would not produce as high titres. Time-course analysis showed that **168** and cogeners are produced after approximately 8 days of fermentation in static M1D media, and continue to accumulate for around 20 days (Figure 2.10, 2.11).
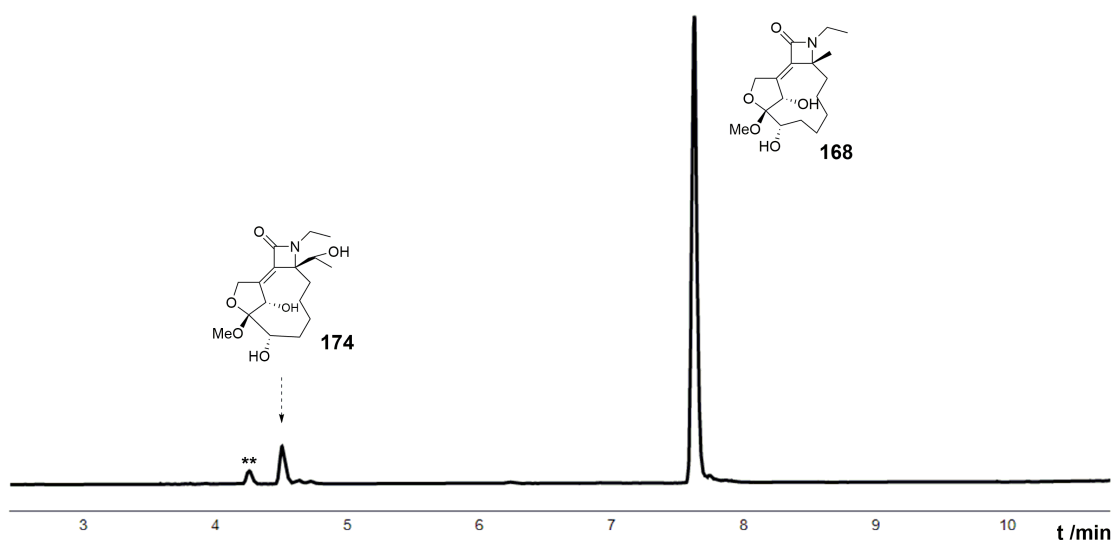
Figure 2.9: ELSD chromatogram of 14 day old *P. cirsii* culture raw extract. ** marks unrelated peak.
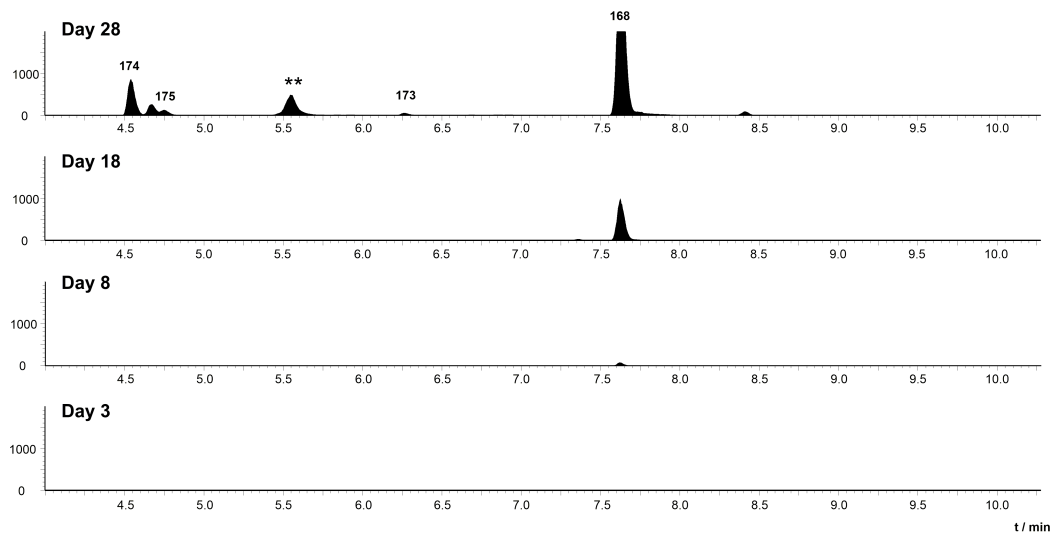
Figure 2.10: Selected ELSD chromatograms up to 28 day old *P. cirsii* culture raw extract. Ordinate axes are linked. ** marks unrelated peak.
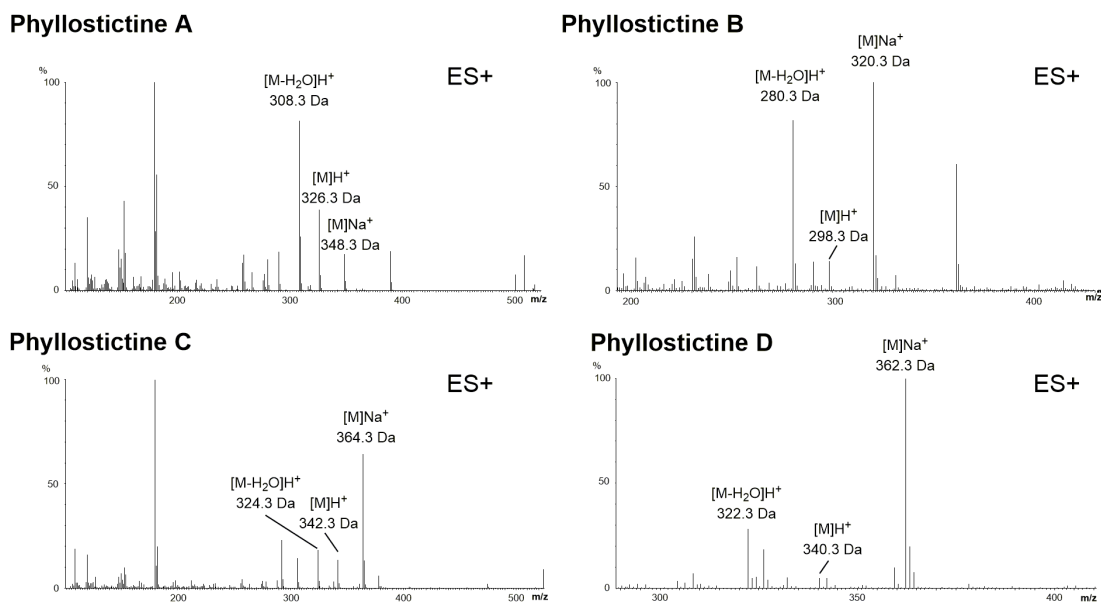
Figure 2.11: ES+ spectra of peaks corresponding to phyllostictines masses.

Purification of phyllostictines was performed by preparative-LCMS using a reverse phase column and a 10% - 90% acetonitrile/water gradient over 20 minutes. Pure phyllostictine A is a yellow oily substance, with production estimated at circa 10 mg $\cdot$ L^{-1}, consistent with literature values.[191]

High Resolution Mass Spectrometry (HRMS) data of pure **168** were compared and confirmed literature values.[191] In particular, the peak corresponding to phyllostictine A (m/z calculated $C_{17}H_{28}NO_5^+$ 326.1967, found 326.1967 [M]H$^+$), coincident with the molecular formula of phyllostictine A $C_{17}H_{27}NO_5$ (not ionized). UV absorption showed λ_{max} = 262.6 nm, value extremely close to the one reported in literature (λ_{max} = 263 nm). Complete HRMS data are hereby reported:

Phyllostictine A (**168**): UV (diode array HPLC, H_2O/CH_3CN) λ_{max} 262.6 nm; LCMS m/z 326.3 [M]H$^+$, 348.3 [M]Na$^+$, 308.3 [M - H_2O]H$^+$, 324.1 [M - H]$^-$; HRESIMS m/z 326.1967 [M]H$^+$ (calcd for $C_{17}H_{28}NO_5^+$, 326.1967).

NMR chemical shifts of 17 mg of pure substance were fully consistent with the data obtained for **168** as described by Evidente and coworkers[191] (Table 2.3, Figure 2.12, 2.13).

| | | | CDCl$_3$, 313 K | |
| | | Measured data | | Literature data |
Position	δ_C/ppm 125 MHz	δ_H/ppm 500 MHz	δ_C/ppm 125 MHz	δ_H/ppm 500 MHz
1	136.8	-	136.3	-
2	156.5	-	156.2	-
3	166.8	-	166.6	-
5	72.2	-	71.8	-
6	22.9	1.30 m (2H)	22.6	1.30 m (2H)
7	29.6	1.30 m (2H)	29.3	1.30 m (2H)
8	29.5	1.30 m, 1.26 m (2H)	29.7	1.30 m, 1.26 m (2H)
9	26.9	1.58 m; 1.37 m (2H)	26.5	1.58 m; 1.37 m (2H)
10	27.8	1.81 m, 1.37 m (2H)	27.5	1.80 m (2H)
11	86.6	4.04 m (1H)	86.3	4.04 m (1H)
12	104.7	-	104.3	-
14	92.8	5.07 d; 5.02 d (2H, 1.6 Hz)	92.7	5.08 d (0.9 Hz)
15	68.9	4.47 br s (1H)	68.4	4.45 br s (1H)
MeO	64.9	3.91 s (3H)	64.5	3.91 s (3H)
Me-C-5	17.0	1.25 s (3H)	17.1	1.27 s (3H)
Me*CH$_2$*N	32.1	1.30 m (2H)	31.8	1.30 m (2H)
*Me*CH$_2$N	14.4	0.89 t (3H, 7.0 Hz)	14.1	0.83 t (3H, 7.1 Hz)

Table 2.3: ^{13}C- and ^{1}H-NMR data of phllostictine A **168**.
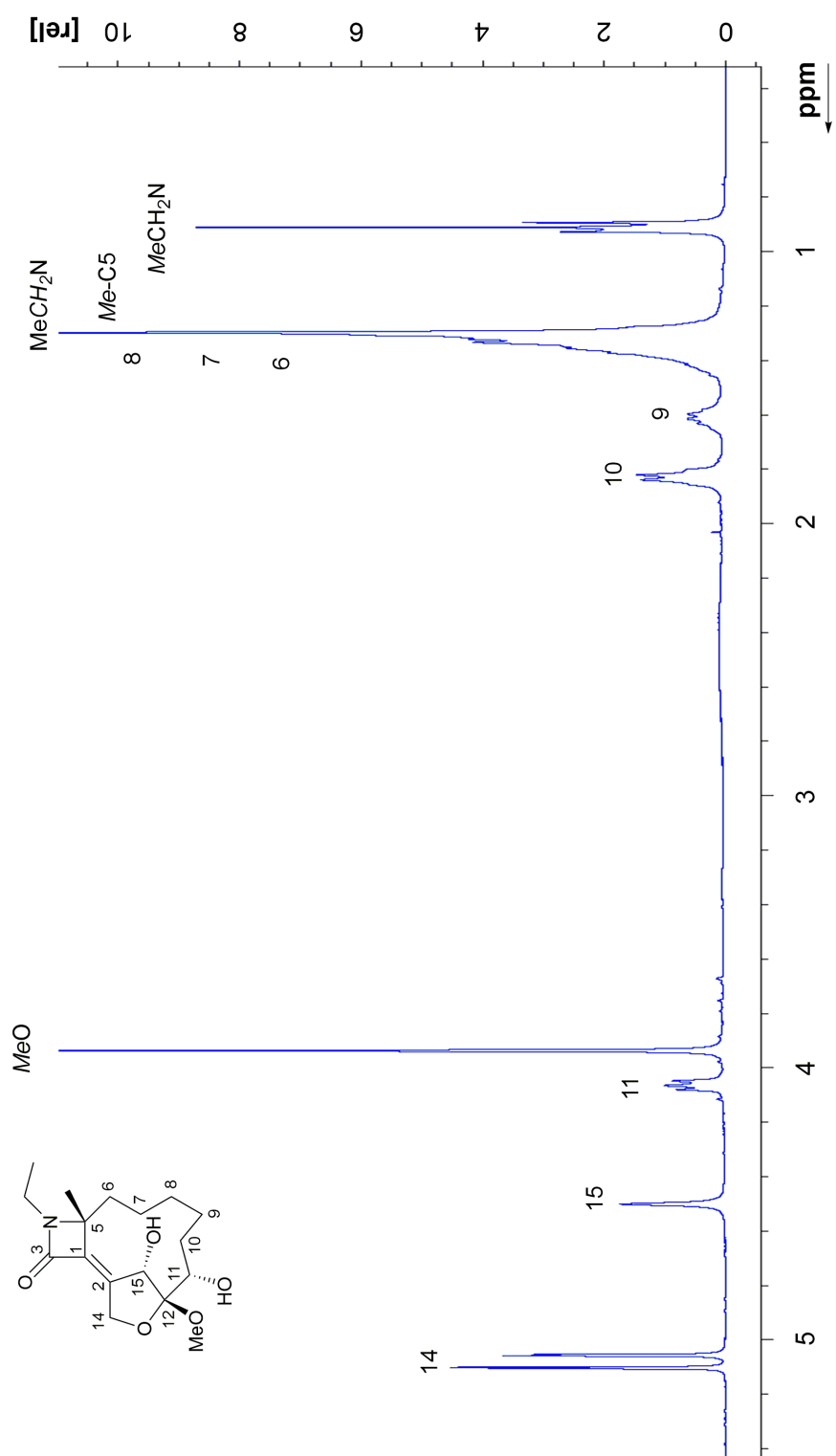
Figure 2.12: Recorded ^{1}H-NMR of **168** in CDCl$_3$.
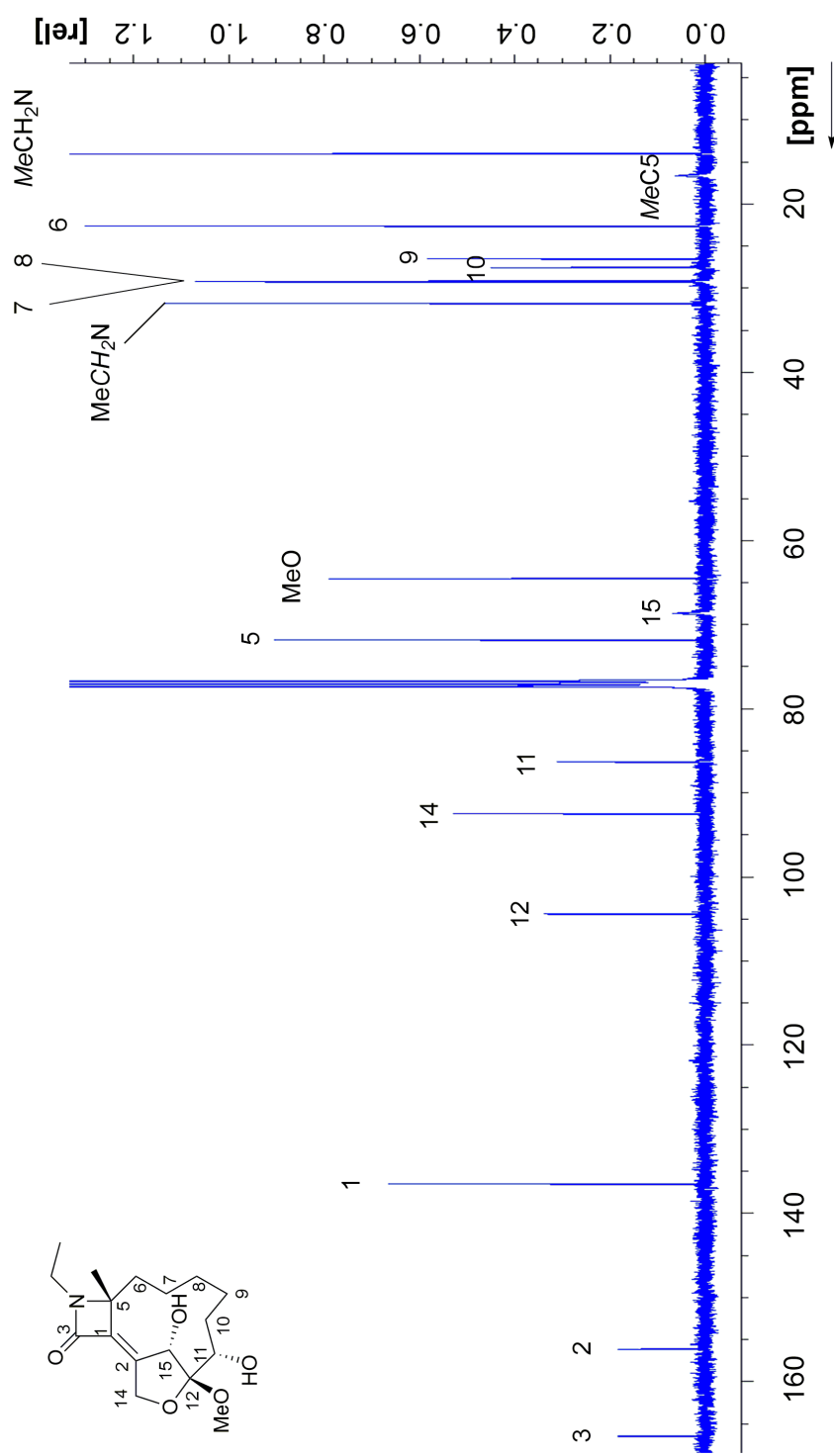
Figure 2.13: Recorded ^{13}C-NMR of **168** in CDCl$_3$.

The ^{13}C-NMR signal with chemical shift 92.8 ppm, assigned to the C-14 ether, is inconsistent with the reported structure which would be expected to have a chemical shift of 60-65 ppm. We speculated a possible misassignment of this carbon resonance and therefore a consequent partial or total erroneous structural proposal. We decided to acquire 2D NMR data (COSY, HMBC and HSQC) to observe H-H and H-C correlations in order to investigate the structure of **168** further. To improve the quality of NMR signal, we measured chemical shifts in DMSO-d$_6$ to avoid proton exchange with the solvent. Full structure elucidation is described in detail in section 2.3.4.

2.3.3 Acetate feeding experiments

Parallel to 2D NMR data acquisition and interpretation, we performed feeding of ^{13}C-labeled acetate. Feeding experiments were done by supplementing producing cultures with either [1-^{13}C]- or [2-^{13}C]-acetate (**187**, **188**) to a final concentration of 14.5 mM. In particular, we expected to see acetate incorporation within the alkylic chain of the 11-member ring in the common alternating fashion. The *N*-ethyl group should not contain any label, and neither C-5, the branching C-14 nor any methyl groups. The labeled building blocks were supplemented at the beginning of **168** biosynthesis, according to time-course evidence, over a time span of 4 days in 400 ml cultures. *P. cirsii* was grown in producing condition; labeled precursors were added on day 11, 12, 13 and 14, and the culture extracted on day 16. Purification of labeled phyllostictine A and ^{13}C-NMR showed incorporation of label up to 6% with a consequent peak intensity enhancement of 7 fold (Figure 2.14, Table 2.4). Peak enhancement (PE) was estimated by calculating the ratio between the normalised peak intensity of each signal in labeled condition (I$_{C, LABELED}$) over the normalised signal intensity of each carbon in unlabeled conditions (I$_{C, UNLABELED}$). Significance was set at a fold increase with threshold ≥ 2. Normalisation of a given carbon signal intensity was performed by dividing its value over the intensity of the methoxy group signal (I$_{MeO}$) of the same dataset acquisition. The methoxy group was used as reference, as it most likely does not derive from acetate.
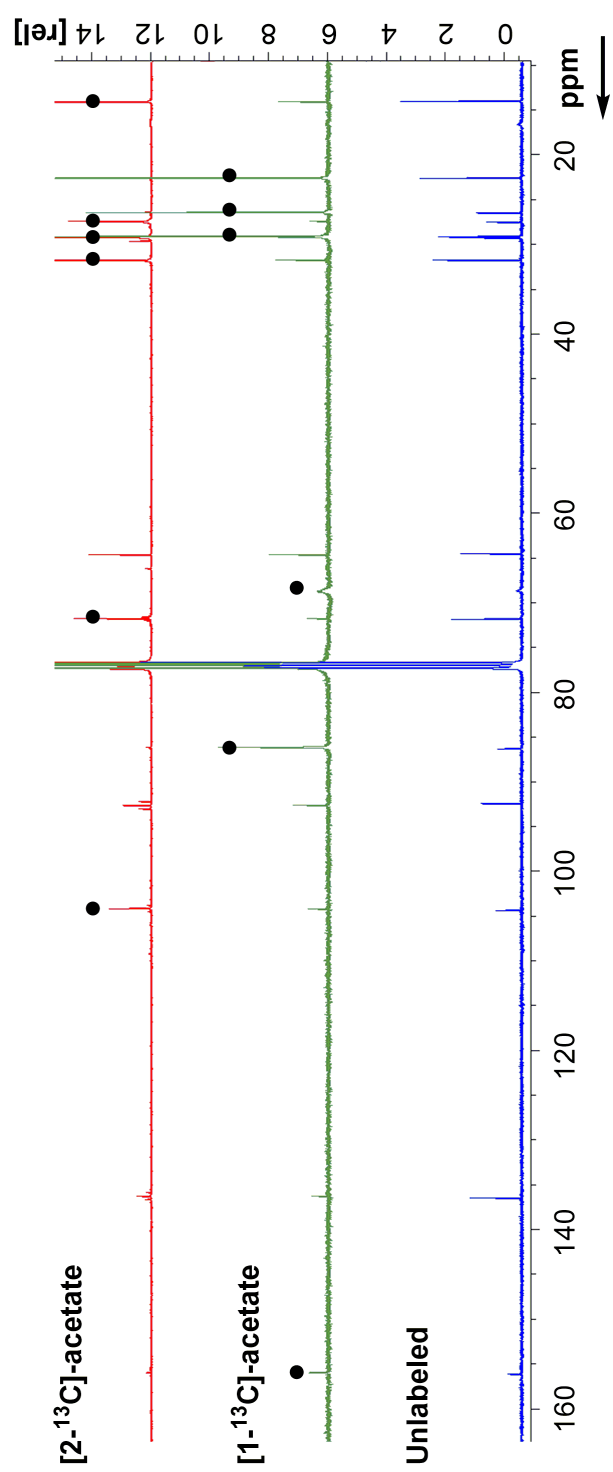
Figure 2.14: NMR signals of enriched phyllostictine A with labeled precursors. Signals with enhanced intensity are marked with ●.

$$\text{I}_\text{C} \text{ normalised} = \frac{I_\text{C}}{I_\text{MeO}} \qquad \text{PE} = \frac{I_\text{C, LABELED, norm}}{I_\text{C, UNLABELED, norm}}$$

C	[1-^{13}C]-acetate PE	[2-^{13}C]-acetate PE
1	1	1
2	1	1
3	5	0.3
5	0.5	3
6	6	0.5
7	1	6
8	5	0.4
9	6	0.2
10	1	5
11	7	1
12	1	2
14	1	1
15	5	1
MeO (reference)	1	1
Me-C-5	1	1
Me*CH$_2$*N	0.7	2
*Me*CH$_2$N	0.4	4

Table 2.4: NMR signal peak enhancement (PE). Significance was set at threshold fold ≥ 2.

Correlation of the labeled carbon resonances with their reported positions in **168** gave surprising results: instead of the regular alternating arrangement of 1- and 2-labeled positions expected for a polyketide, the labeling pattern was scattered randomly across the molecule (Scheme 2.2).
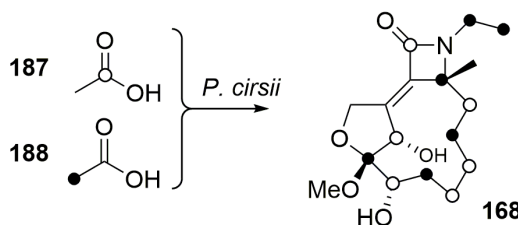
Scheme 2.2: Inconsistent incorporation pattern of acetate units into **168**.

Incongruity was evident at the *N*-ethyl group which appeared to be composed by two acetate C-2 carbons. The β-lactam ammide, labeled as C-1 acetate, next to an unlabeled α carbon was also illogical, along with the two consecutive C-1 carbons on the aliphatic chain of the 11-member ring. This evidence points to the conclusion that the proposed structure **168** is likely to be wrong.

2.3.4 Phyllostictine A structural reassignment

A new set of 1D and 2D NMR data was measured for 17 mg pure phyllostictine A in DMSO-d_6 (Table 2.5, Figures 2.18 - 2.22). HMBC signals arising from C-5-methyl protons showed strong apparent 5- and 7/8-bond correlations with C-15 and C-11 respectively, but no observed 4- and 3-bond correlations to C-1 and C-6 (Figure 2.15). This observation further suggested that the claimed structure **168** was erroneous.
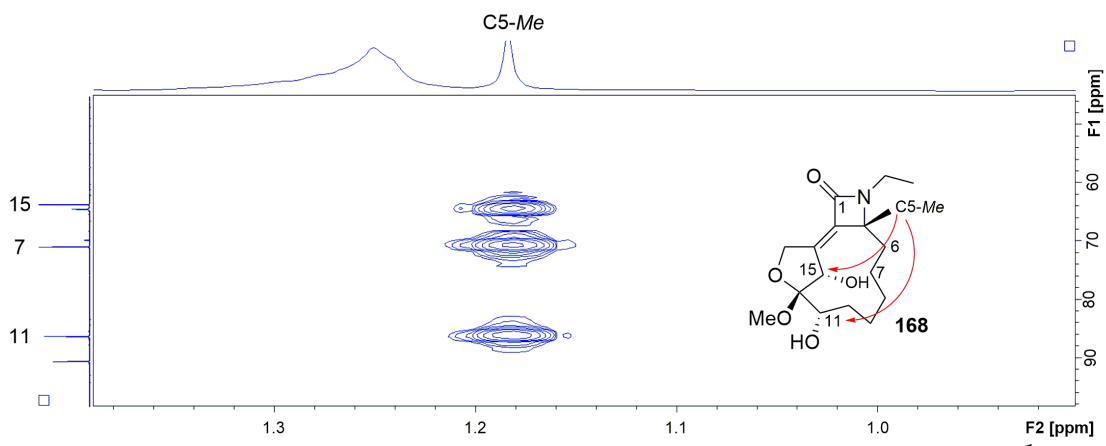
Figure 2.15: Inconsistent proton to carbon HMBC correlations of the claimed structure **168**.

^{1}H-NMR showed a triplet at chemical shift 0.85 ppm with integral of 3, consistent with a terminal methyl group next to a CH_2. The 1-bond HSQC correlation spectrum showed seven methylenes, six with proton chemical shifts between 1.2 - 1.8 ppm, consistent with an aliphatic chain, and one at high chemical shift (4.95 ppm), that could be part of a conjugated double bond. COSY and long range HMBC correlations confirmed the 6 methylenes to form an aliphatic chain connected to the terminal -CH_3 at 0.85 ppm. The signals of the aliphatic -CH_2 showed strong overlap, with the exception of one methylene (nucleus a) that displayed two diasterotopic protons at 1.82 and 1.55 ppm, both correlating in the HSQC to the same carbon (27.5 ppm), therefore we concluded that nucleus a was a -CH_2 connected to a stereocenter (**189**, Figure 2.16).
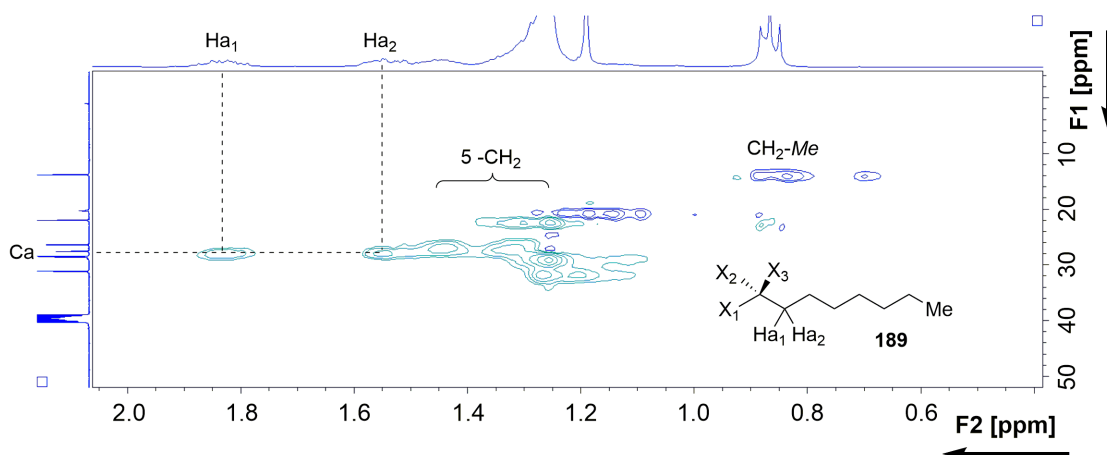
Figure 2.16: HSQC correlation of the 7 member aliphatic chain connected to a stereocenter.

The HSQC spectrum revealed 5 quaternary carbons, four at chemical shifts between 166 - 104 ppm (Cb, Cc, Cd and Ce) corresponding to the ester/amide and olefin region, and one at 70.6 ppm (Ch), characteristic of C-OH signals. COSY coupling proved Ch to be linked to a methyl and to a hydroxy group (1.17 ppm singlet, 4.84 ppm broad singlet respectively). HMBC correlation between Cc-Hi suggested a dihydropyran moiety, as Hi had a chemical shift consistent with a CHOH (4.06 ppm) and Hf displayed HMBC correlations to Ce, Cc, Ch and Ci. Cc, not only correlated with Hi through the ether and Hf, but coupled with Hg as well. Hg was the high shift -CH$_2$ mentioned above, while Hf fell within the hydroxy region (3.86 ppm). Interestingly, COSY correlation showed Ci to be connected to the 7 member aliphatic chain, meaning that Ci was a stereocenter. Cb had a chemical shift of 166.4 ppm, and since there is a nitrogen atom in the structure, we assigned this resonance to an amide group. The absence of -NH_2 signal at high chemical shift, suggested that Cb is a doubly substituted amide or part of a ring. The doublet at 5.34 ppm (integral 1H) showed strong COSY correlation with carbon at 64.4 ppm (Cf), consistent with an -OH attached to Cf. HMBC correlation revealed Ce (104.8 ppm) to couple only with -OHf and Hf, meaning that Ce is probably connected directly to Cb, since Cb also couples with Hf (HMBC). Long distance correlations were also observed between Ch and Hf. HMBC correlation between Cd and Cg revealed the latter to be a terminal -CH$_2$ attached to a quaternary carbon. Cd correlated only with Cg, meaning that there were no protons within a 3-4 bond range, beside the terminal methylene (Figure 2.17).
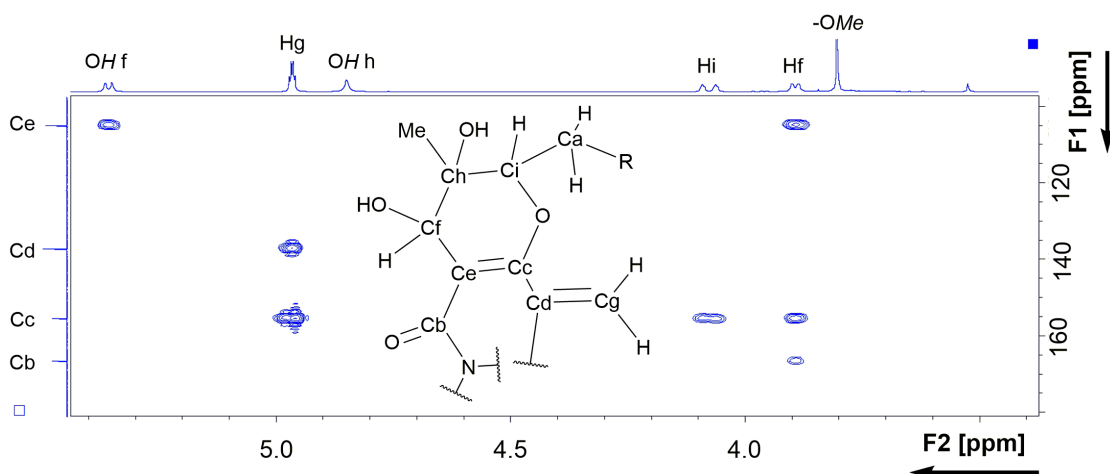
Figure 2.17: HMBC long distance correlations of the quaternary carbons and partial structure of phyllostictine A. R stands for the aliphatic chain.

The partial structure elaborated was sufficient to find related compounds in the literature using the Reaxys database. In particular, we reassigned phyllostictine A structure to the *N*-*O*-methyl 3-methylene tetramic acid **190** based on full NMR data obtained in CDCl$_3$ and DMSO-d$_6$. We also used information from the acetate feeding experiments and comparison to known compounds found by Reaxys search. **190** is related to the phaeosphaeride class (**181**) and it is probably derived from PKS-NRPS biosynthesis, as acetate feeding experiments (Section 2.3.3) showed incorporation of this building block, while the nitrogen probably stems from an amino acid.
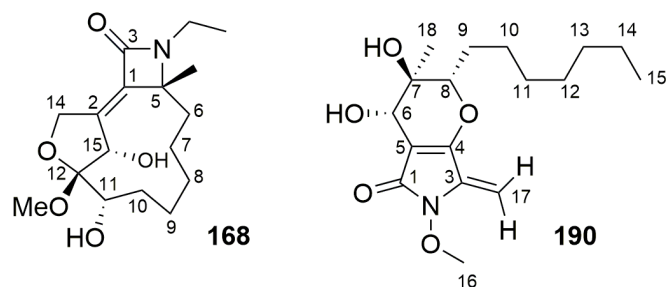
		DMSO-d$_6$, 313 K	
		δ_C/ppm	δ_H/ppm
Literature assignment	Reassignment	100 MHz	400 MHz
3	1	166.4	-
2	4	155.3	-
1	3	137.1	-
12	5	104.8	-
14	17a, 17b	90.5	4.95 d; 4.96 d (2H, 1.5 Hz)
11	8	86.2	4.06 m (1H)
5	7	70.6	-
15	6	64.4	3.86 d (1H, 5.6 Hz)
MeO	16	63.6	3.78 s (3H)
Me*CH$_2$*N	11a, 11b	31.1	1.26 m, 1.19 m (2H)
7	13	28.5	1.25 m (2H)
8	14	28.4	1.30 m, 1.25 m (2H)
10	9a, 9b	27.5	1.82 m, 1.55 m (2H)
9	10a, 10b	26.3	1.44 m; 1.33 m (2H)
6	12	21.9	1.24 m (2H)
Me-C-5	18	20.2	1.17 s (3H)
*Me*CH$_2$N	15	13.8	0.85 t (3H, 6.8 Hz)
-	O*H*-6	-	5.34 d (1H, 5.9 Hz)
-	O*H*-7	-	4.84 s (1H)

Table 2.5: Reassignment of literature ^{13}C- and ^{1}H-NMR values of phyllostictine A.
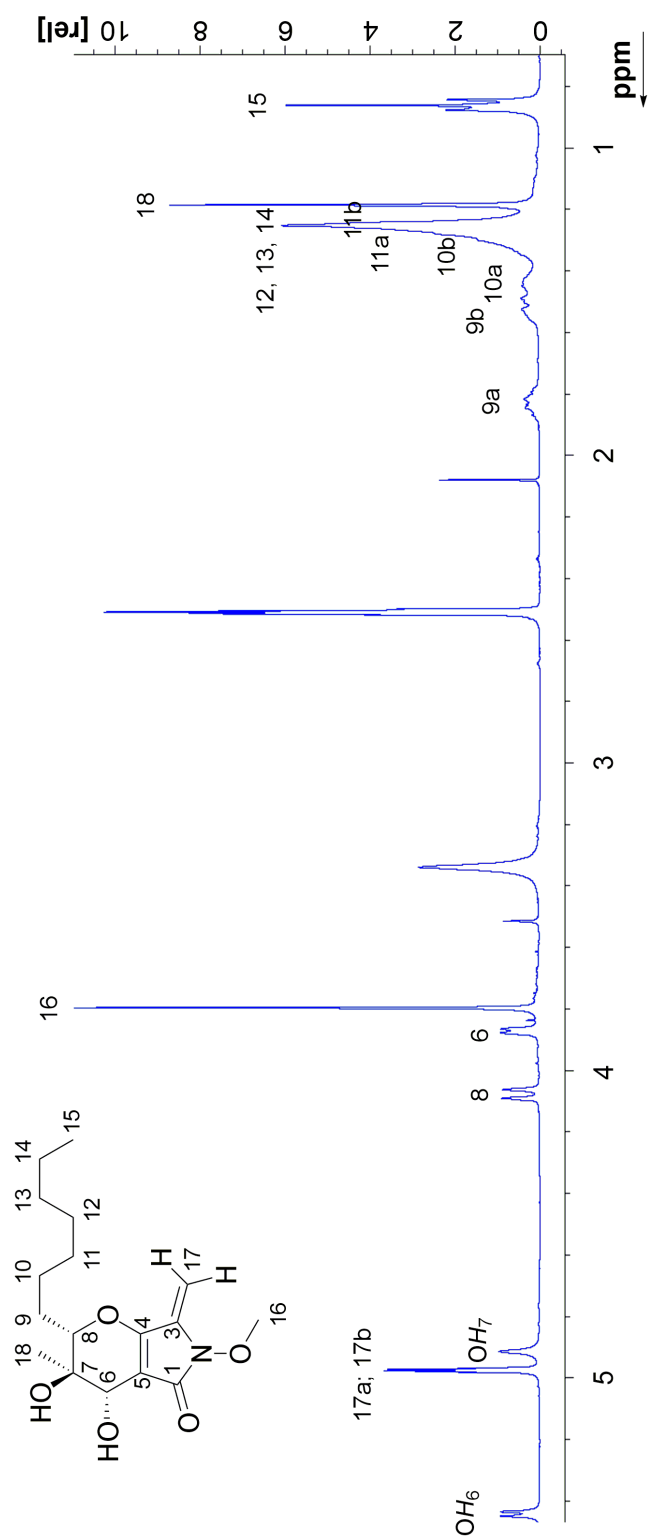
Figure 2.18: Recorded ^{1}H-NMR spectrum of phyllostictine A in DMSO-d$_6$.
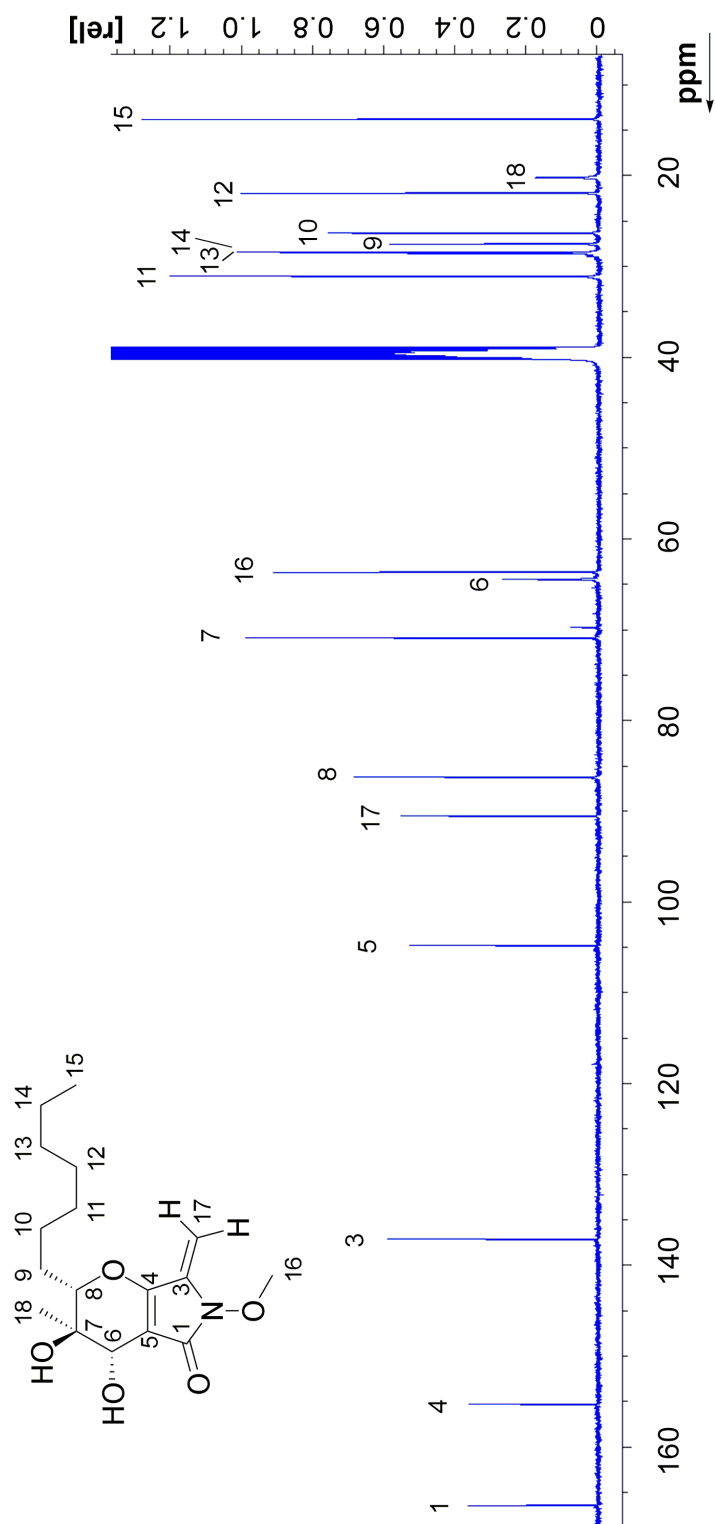
Figure 2.19: Recorded ^{13}C-NMR spectrum of phyllostictine A in DMSO-d$_6$.
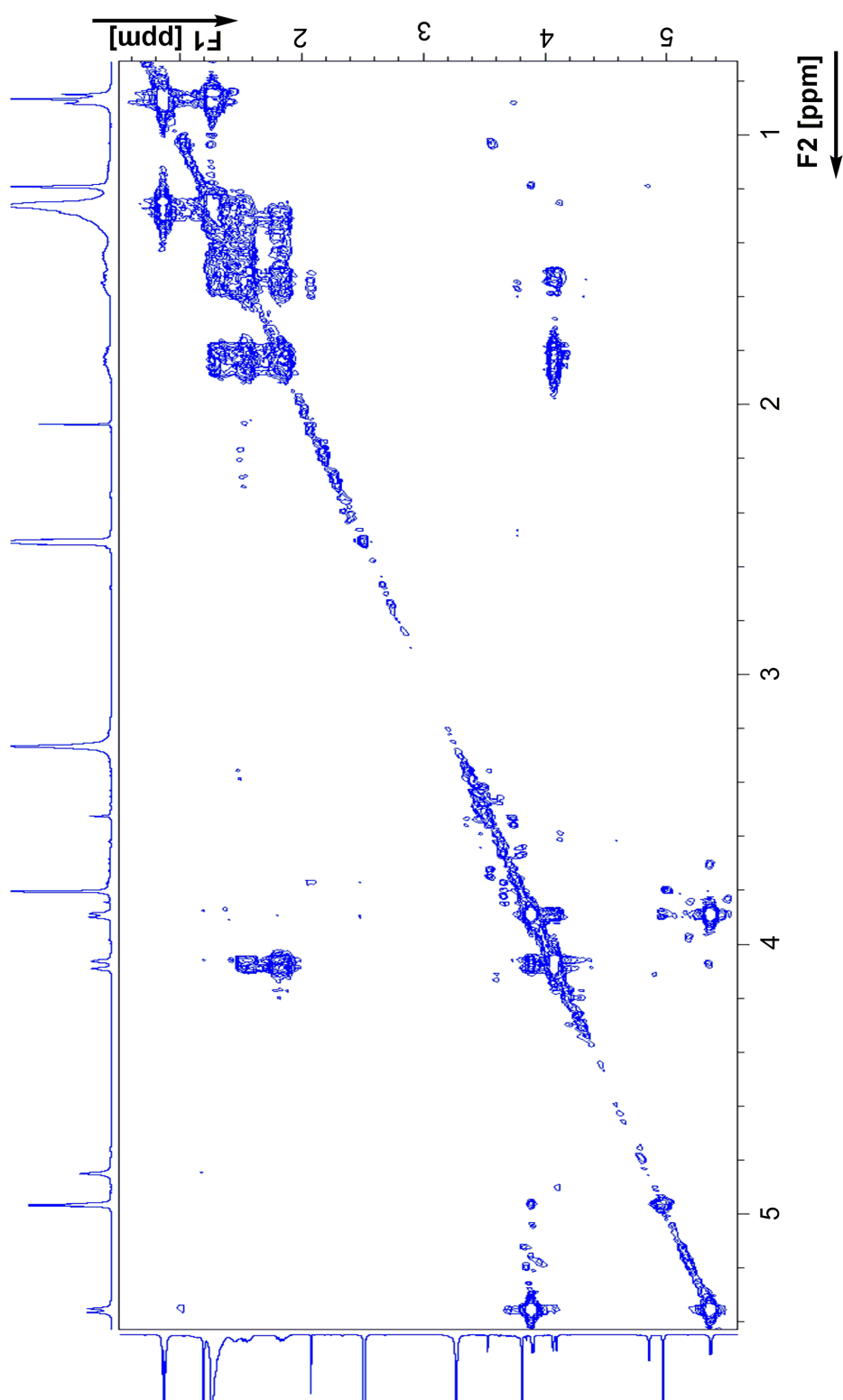
Figure 2.20: COSY correlation of Phyllostictine A in DMSO-d$_6$.
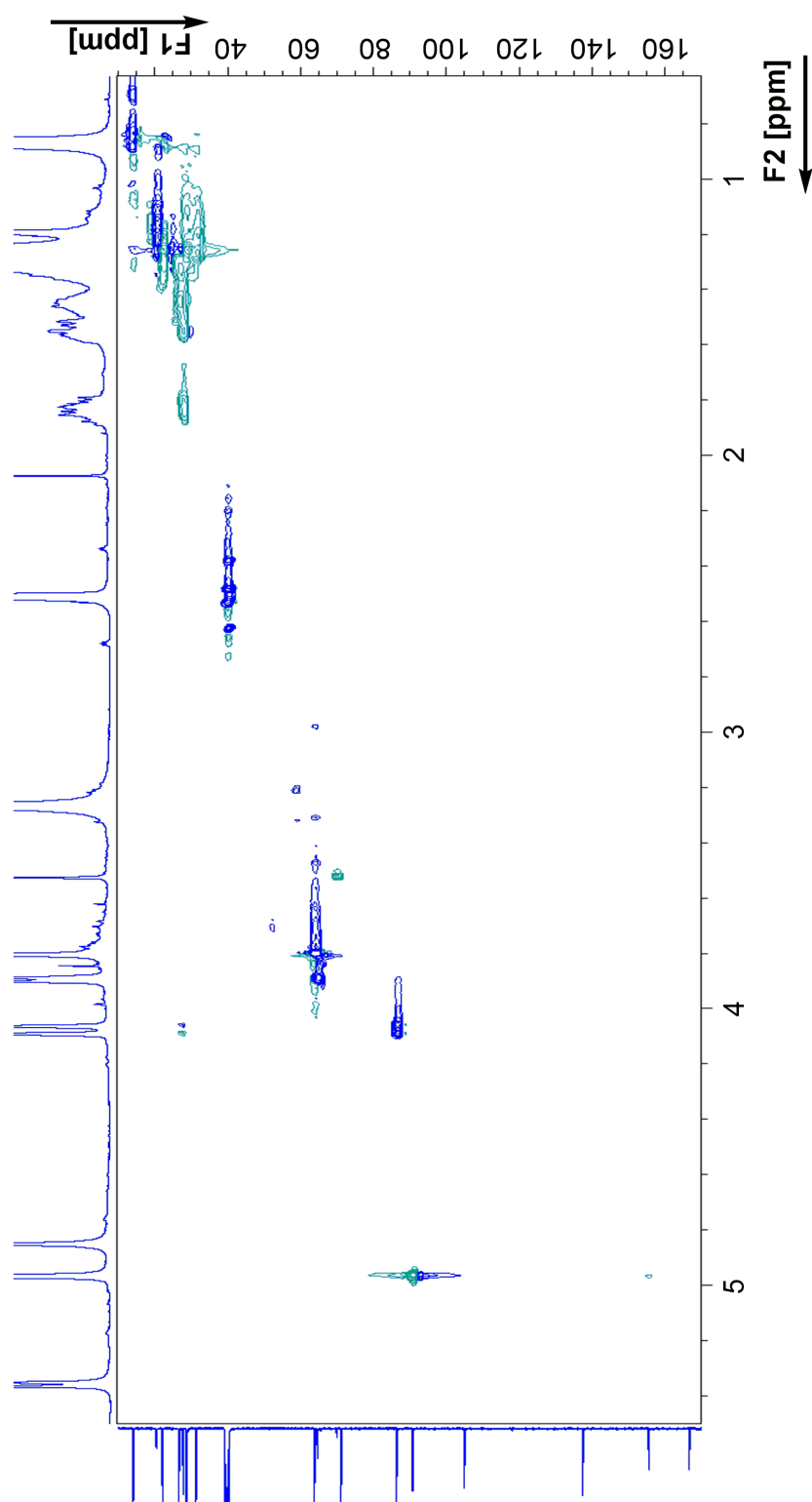
Figure 2.21: HSQC of Phyllostictine A in DMSO-d$_6$.
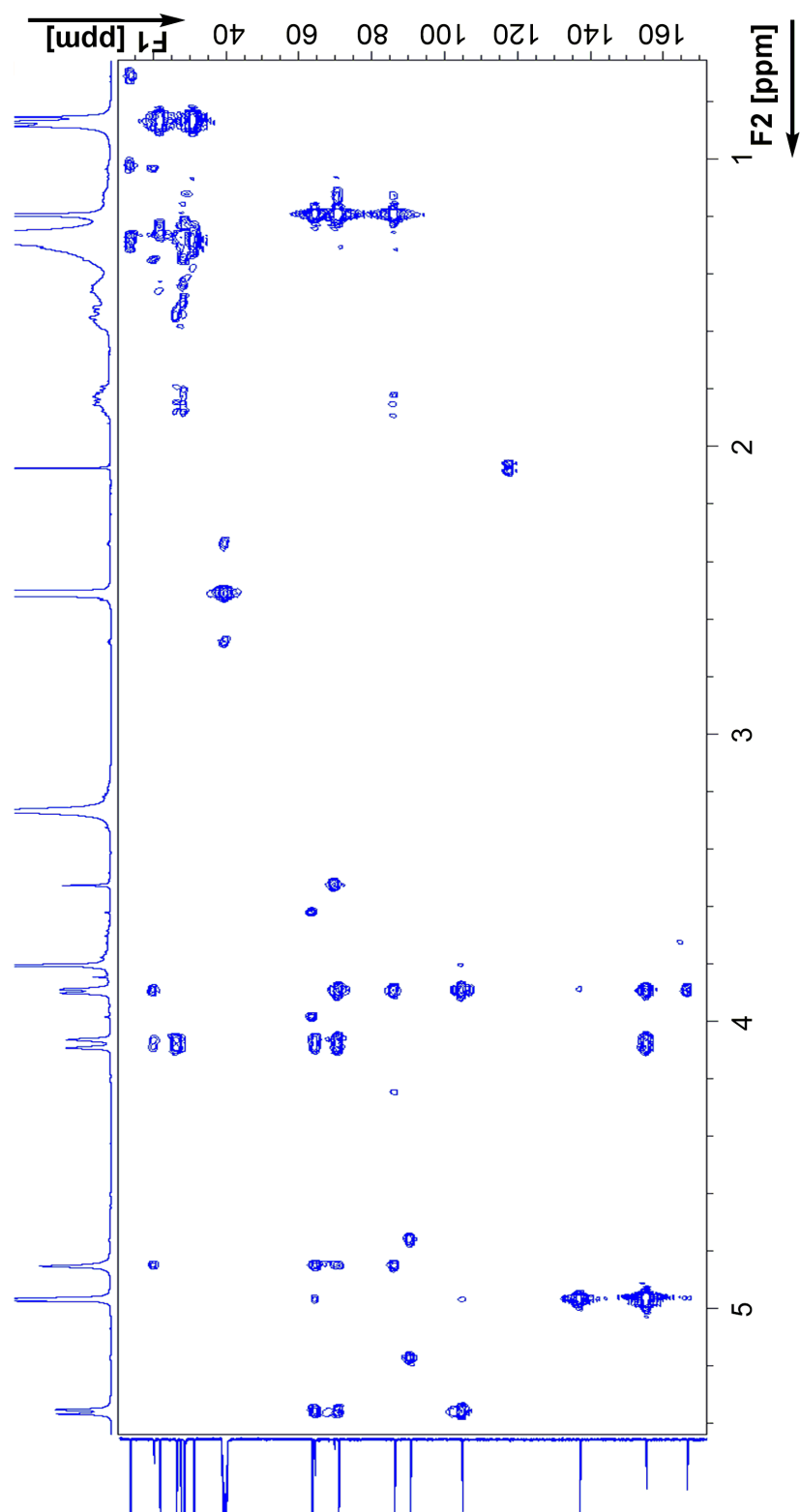
Figure 2.22: HMBC of Phyllostictine A in DMSO-d$_6$.

HMBC correlations suggested a bicyclic system for **190**, with strong 2/3-bond correlations from H-6 to C-1, C-4, C-5, C-7 and C-8. Also the weak 4-bond correlation H-6/C-4 was visible (Figure 2.23). C-16 has no strong HMBC correlation, consistent with it being the *N*-*O*-methyl group.
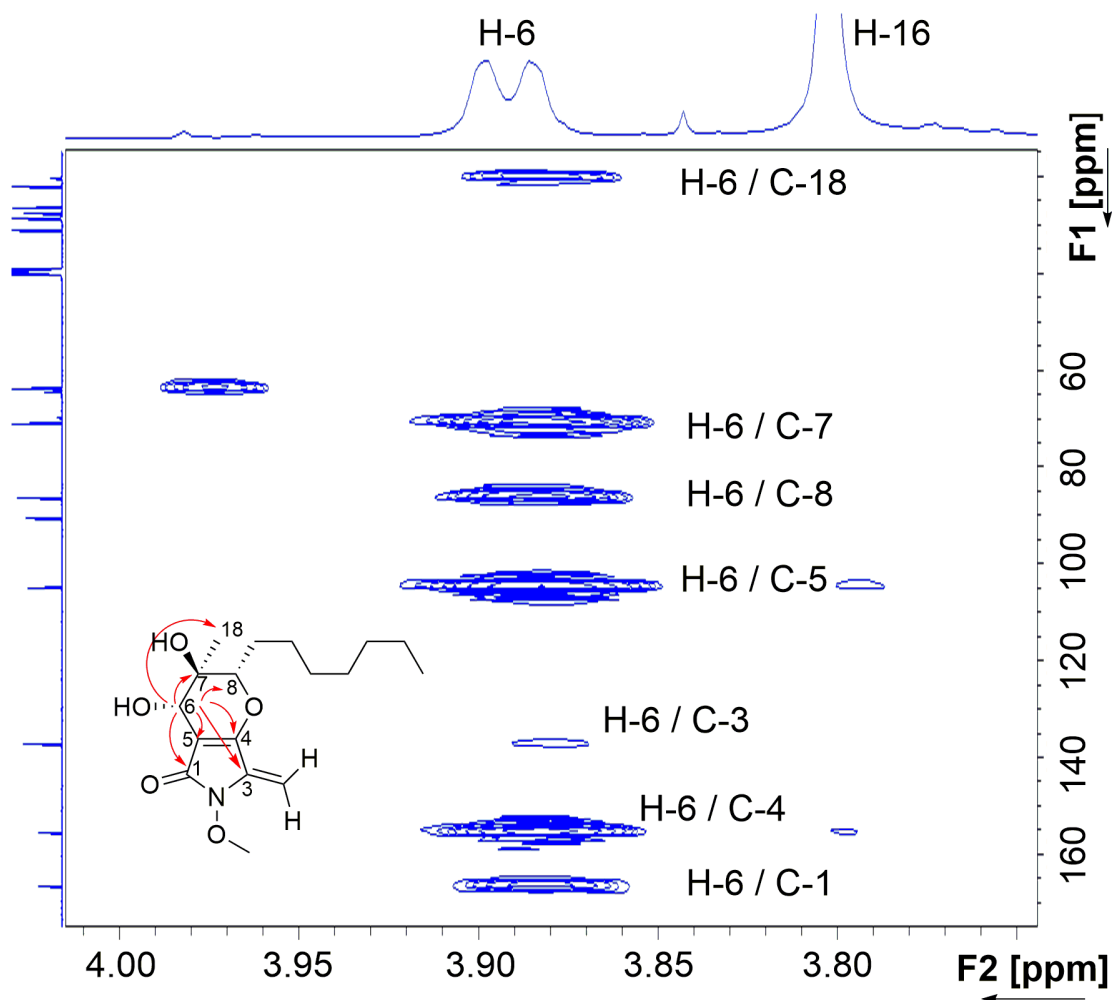
Figure 2.23: HMBC correlation of H-6 to the carbons of the bicyclic dihydropyran-tetramic acid system.

The conjugated double bond system including the terminal CH_2-17 was confirmed by long distance 6-bond COSY correlation between H-17 and H-6 (Figure 2.24).
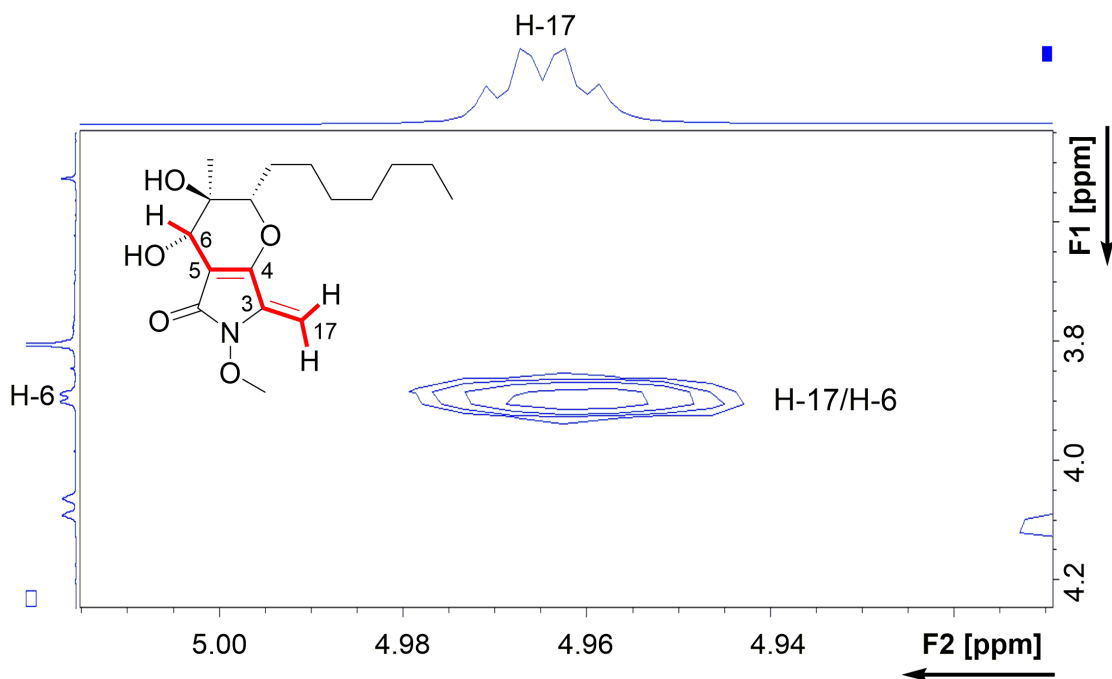
Figure 2.24: Phyllostictine A COSY in DMSO-d$_6$.

HMBC data showed a strong correlation between H-17 and both C-3 and C-4, confirming the tetramic acid moiety (Figure 2.25). Weak correlations were observed for the C(17) nucleus, such as the 5-bond H-17-C-6 and the weaker 4-bond H-17-C-5, through the conjugated system. Correlation between 7-OH, C-7 and C-18 locates both the methyl and -OH as attached to C-7. The 7-hydroxyl group couples to C-6, 3 bonds away and sensible correlation were observed for nuclei (8) and (9) both in COSY and HMBC. Long range correlation data showed 3-bond correlation between H-8 and the methyl C-18, H-8/C-4 through the ether bridge, H-8/C-9 of the acyl chain. Data acquisition in DMSO-d$_6$ showed -O*H*-6 to correlate with C-5, C-6 and C-7.
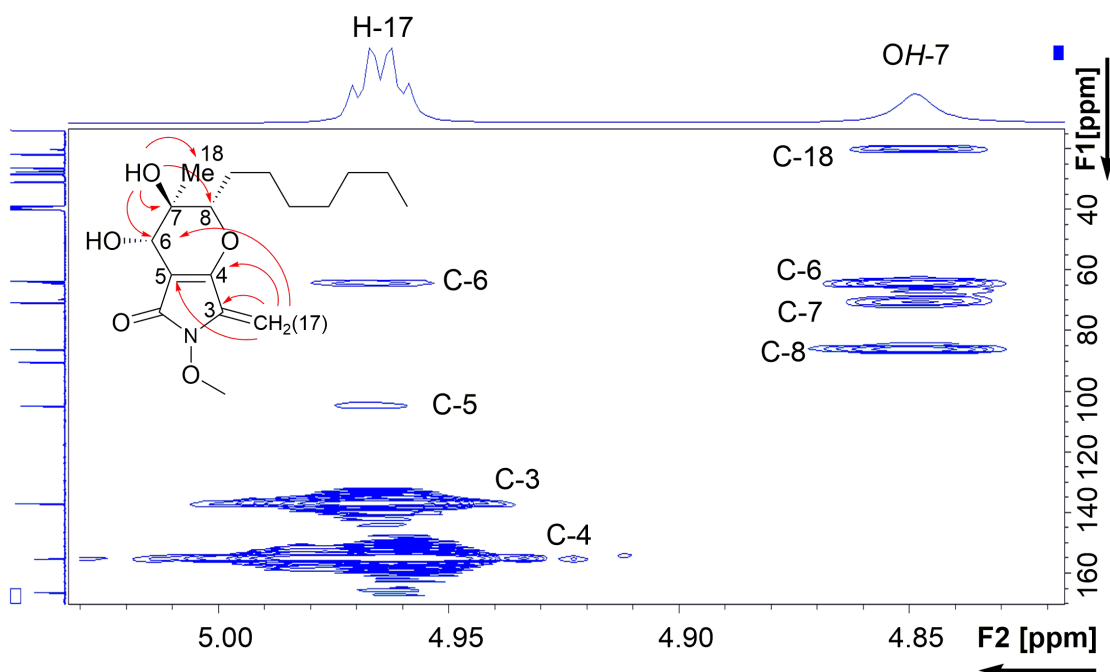
Figure 2.25: HMBC correlation of H-17 to the carbons of the bicyclic system. 7-O*H* correlate with the adjacent C-7 and C-18.

2.3.5 Methionine and Alanine feeding experiments

To determine the origin of methyl groups and the amino acid involved in tetramic acid formation we supplemented [*methyl*, ^{13}C]-methionine and [1-^{13}C]-alanine. Acetate feeding experiments (Section 2.3.3) showed a total of five unlabeled carbons that, located on **190**, correspond to the two methyl groups C-16 and C-18, and to C-17, C-3 and C-4 of the 3-methylene tetramic acid (Scheme 2.3). C-16 and C-18 are most likely derived from SAM by the action of a tailoring *O*-methyl transferase and the C-MeT domain of the PKS respectively. Alanine is the building block of choice because of the three remaining unlabeled carbons next to the nitrogen atom: C-3 is probably the former α-carbon of the amino acid and the terminal methylene the oxidised lateral chain of alanine (-CH$_3$ to -CH$_2$).
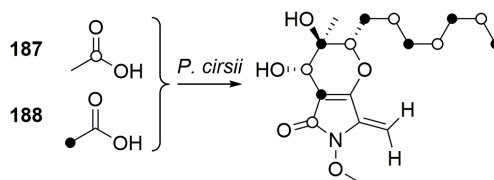
Scheme 2.3: Incorporation of ^{13}C-acetate into **190**.

Analogously to acetate feeding experiments, labeled amino acids were fed in parallel to 400 ml culture growing in producing conditions at day 10 for 4 days in a row, to a final concentration of 2.5 mM. We expected [*methyl*, ^{13}C]-methionine (**17**) feeding to enrich C-16 and C-18 NMR signals and [1-^{13}C]-alanine (**191**) to increase C-4 NMR signal. C-17 was used as reference to normalise each NMR dataset as it lays on a unlabeled position in every feeding experiment. Peak enhancement (PE) was calculated (Section 2.3.3) and two folds increase was regarded as significant incorporation site (Table 2.6, Figure 2.26, Scheme 2.4).

C	[*methyl*-^{13}C]-Met PE	[1-^{13}C]-Ala PE
1	1	1
3	0.5	0.5
4	0.5	6.5
5	0.5	0.5
6	0.5	0.5
7	0.5	0.5
8	0.5	1
9	0.5	1
10	1	1
11	0.5	0.5
12	0.5	0.5
13	0.5	0.5
14	0.5	1
15	0.5	0.5
16	6	1.4
17 (reference)	1	1
18	5	0.5

Table 2.6: NMR signal ratios. Significance was set at fold ≥ 2.

We observed enriched carbons to be located in positions C-16 and C-18 for the methionine feeding, and at C-4 for alanine feeding. This is consistent the hypothesis that **190** is the product of a highly reducing PKS-NRPS hybrid system, decorated by SAM-dependent methylations (Scheme 2.4).
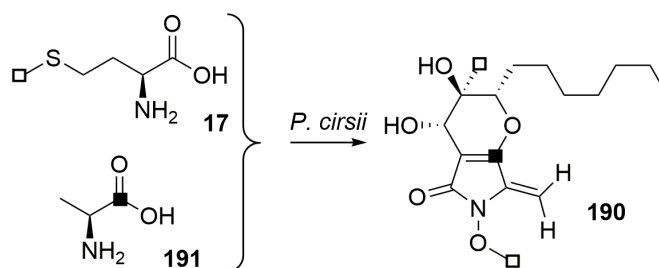
Scheme 2.4: Incorporation of [*methyl*-^{13}C]-methionine and [1-^{13}C]-alanine into **190**.

65
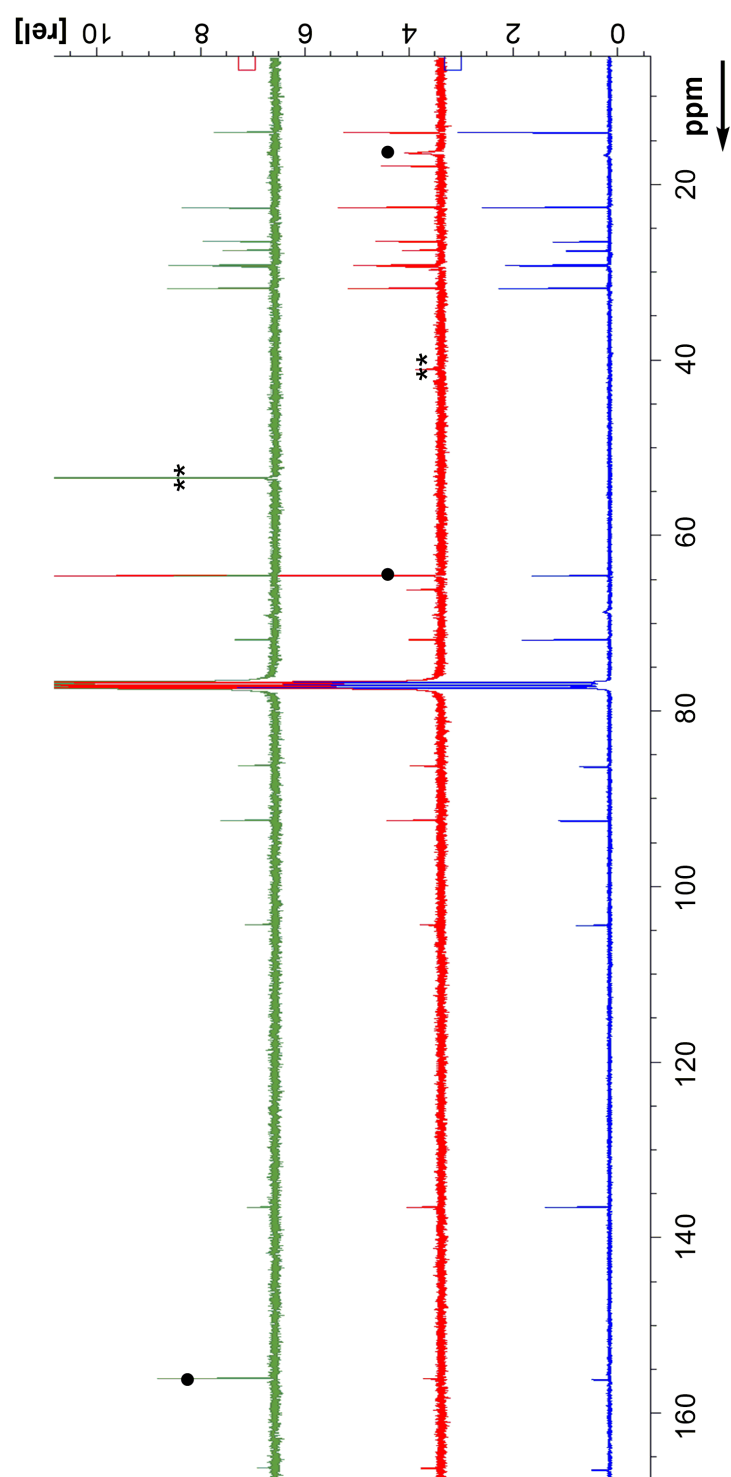
Figure 2.26: NMR signals of enriched **190** with labeled precursors. Signals with enhanced intensity are marked with ●. ** marks unrelated peaks.

2.3.6 Phyllostictine B NMR analysis

Isolation by preparative LCMS and NMR analysis were performed for the minor product phyllostictine B **181** (HRMS calculated $C_{15}H_{24}NO_5^+$ 298.1654, found 298.1654 $[M]H^+$). Phyllostictine B molecular weight differs from **190** by 28 Da, corresponding to two methylenes shorter in the aliphatic chain. NMR data were acquired in DMSO-d_6 and compared to literature value relative to phaeosphaeride A **181** and its diastereomer phaeosphaeride B **192** (Figure 2.27), which were characterised by Clardy and collaborators by NMR in the same solvent (Tables 2.7, 2.8).[242]
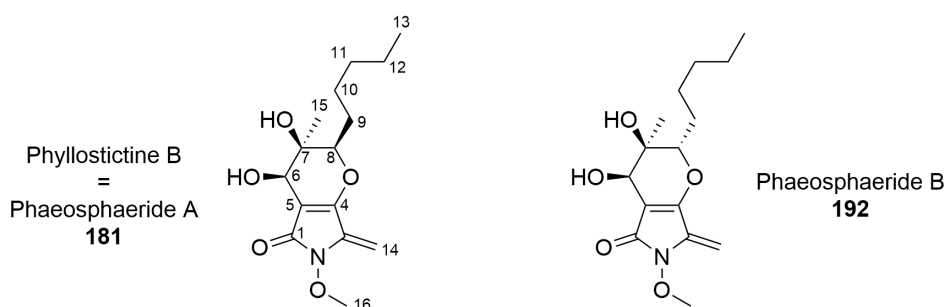
Figure 2.27: Phyllostictine B and phaeosphaeride A-B.

	δ_H/ppm; DMSO-d_6		
	Phaeosphaeride A	Phyllostictine B	Phaeosphaeride B
Position	500 MHz	400 MHz, 313 K	500 MHz
6	3.86 d, 1H, J = 5.5 Hz	3.86 d, 1H, J = 5.6 Hz	3.78 s, 1H
8	4.07 d, 1H, J = 11.5 Hz	4.07 d, 1H, J = 11.5 Hz	3.98 dd, 1H, J = 10.6 Hz
9a	1.82 m, 1H	1.81 m, 1H	1.86 m, 1H
9b	1.51 m, 1H	1.52 m, 1H	1.54 m, 1H
10	1.44 m, 2H	1.44 m, 2H	1.63 m, 2H
11	1.27 m, 2H	1.26 m, 2H	1.32 m, 2H
12	1.28 m, 2H	1.27 m, 2H	1.33 m, 2H
13	0.85 t, 3H, J = 6.7 Hz	0.86 t, 3H, J = 6.8 Hz	0.90 m, 3H
14a	4.97 d, 1H, J = 1.8 Hz	4.97 bs 2H	5.03 d, 1H, J = 2.0 Hz
14b	4.96 d, 1H, J = 1.8 Hz	4.97 bs, 2H	5.02 d, 1H, J = 2.0 Hz
15	1.18 s, 3H	1.18 s, 3H	0.88 s, 3H
16	3.79 s, 3H	3.79 s, 3H	3.75 d, 3H, J = 5.9 Hz
OH-6	5.44 d, 1H, J = 5.5 Hz	5.43 d, 1H, J = 5.8 Hz	5.49 d, 1H, J = 5.9 Hz
OH-7	4.92 s, 1H	4.91 s, 1H	4.59 s, 1H

Table 2.7: ^{1}H-NMR data of phllostictine B **181**, phaeosphaeride A **181** and phaeosphaeride B **192**.

| | δ_C/ppm; DMSO-d$_6$ | | |
Position	Phaeosphaeride A 125 MHz	Phyllostictine B 100 MHz, 313 K	Phaeosphaeride B 125 MHz
1	166.5	166.4	165.3
3	137	137.1	136
4	155.2	155.3	156.8
5	104.7	104.8	105
6	64.1	64.4	63.6
7	70.8	70.7	69.4
8	86.2	86.2	80.3
9	27.5	27.5	26.6
10	25.9	25.9	25.1
11	30.8	30.8	30.8
12	21.9	21.9	21.6
13	13.8	13.7	13.7
14	90.8	90.6	91.4
15	20.5	20.2	17.9
16	63.7	63.6	63.4

Table 2.8: ^{13}C-NMR data of phllostictine B **181**, phaeosphaeride A **181** and phaeosphaeride B **192**.

NMR data for phyllostictine B exactly matched that of phaeosphaeride A, revealing that they must be the same compound. Also the diastereomer **192** showed very close values, with the exception of nuclei at position 8, which reflect the different configuration for that center.

Optical rotation was measured for phyllostictine B, giving an α_D = -99.8, very similar to the reported phaeosphaeride value (α_D = -93.6).[242] Complete HRMS data of phllostictine B and phaeosphaeride A (both **181**) are hereby reported.

Phyllostictine B (**181**): UV (diode array HPLC, H$_2$O/CH$_3$CN) λ_{max} 262.6 nm; LCMS m/z 298.3 [M]H$^+$, 320.2 [M]Na$^+$, 280.2 [M + H - H$_2$O]$^+$; HRESIMS m/z 298.1654 [M]H$^+$ (calcd for C$_{15}$H$_{24}$NO$_5$$^+$, 298.1654), 5 double bond equivalents.

Phaeosphaeride A (**181**): $[\alpha]^{25}_D$ = -93.6 (c 2.0, CH$_2$Cl$_2$)) yellow glass. HRESIMS m/z [M]H$^+$ 298.1656, C$_{15}$H$_{23}$NO$_5$ + H (calcd 298.1654), 5 double bond equivalent.

2.3.7 Genomics

Total genome sequencing has never been described for *P. cirsii*. Genomic DNA was obtained by growing the fungus in 200 ml Potato Dextrose Broth (PDB) medium for 5 days. Mycelia were freeze-dried and then ground with mortar and pestle. gDNA

purification was performed using the GenElute Plant Genomic DNA miniPrep kit (Sigma-Aldrich[240]).

De novo sequencing of *P. cirsii* genomic DNA was performed by Paired End Illumina sequencing in collaboration with Daniel Wibberg and Andreas Schlüter at the Center for Biotechnology, Bielefeld, Germany (CeBiTec).[243] Raw data was processed by an in house software platform.[243] After assembly of all sequence reads by applying the GS /De Novo/ Assembler (version 2.8 software with default settings), the draft genome consisted of 303 scaffolds of average size 112 Kb. The calculated genome size was 34 Mb with a scaffold N_{50} of $\sim$287 KBp (Table 2.9). The annotation of the draft genome was made within the GenDBE platform[244] including AUGUSTUS 3.0.3 for gene prediction.[245]

Aligned Reads (All/Paired end)	5,245,747/2,613,250
Assembled Bases	1,391,131,960 Bp
Pair End Size(s)	728 $\pm$ 245 Bp
Scaffolds (All/True)	303/303
Contigs (Scaffolded/Large(>500 Bp)/All)	503/1,242/3,087
Bases in Scaffolds	33,926,906 Bp
Coverage	41 x
GC content	52.2%
Avg. Scaffold	111,969 Bp
N_{50} Scaffold	287,226 Bp
Largest Scaffold	1,257,330 Bp
Avg. Scaf. Contig	64,912 Bp
Avg. Contig	27,704 Bp
N_{50} Contig	191,349 Bp
Largest Contig	540,905 Bp

Table 2.9: Assembly details of *P. cirsii* genome.

The whole genome sequence of *P. cirsii* was submitted to the Antibiotics & Secondary Metabolite Analysis Shell (antiSMASH) for the automatic annotation of putative gene clusters. In total 32 gene clusters were found, two of which encode a PKS-NRPS hybrid system (Table 2.10).

Gene cluster type	
Type I PKS	10
Type III PKS	1
NRPS	5
PKS-NRPS	2
Terpene	6
Homoserine lactone	1
Other	7
TOT	**32**

Table 2.10: Predicted gene clusters of *P. cirsii* by AntiSMASH.

Polyketide synthases and PKS-NRPS clusters were manually annotated by protein BLAST analysis, using the NCBI non-redundant protein sequence and the Protein Data Bank (PDB) database, in order to focus on the highly reducing PKS. One of the two PKS-NRPS clusters had one highly reducing PKS-NRPS plus an *O*-Methyl transferase encoding gene, possibly involved in the construction of the *N*-*O*-methyl moiety of **190**, therefore it was selected as a feasible BGC involved in phyllostictine biosynthesis. The other PKS-NRPS cluster did not have an *O*-Methyl transferase encoding gene, therefore it was considered less likely to be involved in phyllostictine biosynthesis. We labeled the best candidate BGC *phy* (Table 2.11). 25 KBp flanking the core *phy*S were further investigated using FGENESH2 for exon/intron and open reading frame (ORF) prediction. The reference organism of choice was *Leptosphaeria*, a plant pathogen of the order of the *Pleosporales*, which was used for intron prediction. The borders of the clusters were defined by rational analysis: genes on the right hand side of the core PKS-NRPS were observed to encode for exonucleases involved in DNA repairing, tRNA encoding genes and RNA polymerase, therefore they were not included in the *phy* BGC.
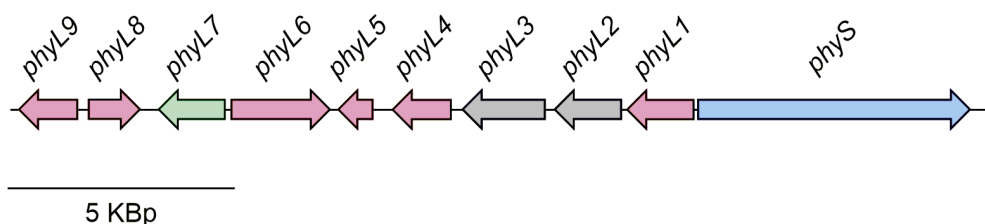
Name	Annotation	Nearest pBLAST hit	Predicted cofactor
*phy*L9	Monooxygenase	*P. griseofulvum* PGRI_02982	NAD/FAD
*phy*L8	Oxygenase	*M. mycetomatis* MMYC01_206825	α-KG; Fe(II)
*phy*L7	*O*-methyl transferase	*P. tritici repentis* PTRG_04253	SAM
*phy*L6	Cytochrome P450	*M. mycetomatis* MMYC01_206826	heme-thiolate
*phy*L5	Very long chain 3-oxoacyl-CoA reductase	*P. attae* AB675_8878	NAD(P)$^+$
*phy*L4	Enoyl reductase	*E. lata* UCREL1_10570	Zn^{2+}; NAD(P)$^+$
*phy*L3	Toxin efflux transporter	*R. necatrix* MFS	
*phy*L2	Transcription factor	*M. mycetomatis* MMYC01_206820	Zn^{2+}
*phy*L1	Oxidoreductase	*A. oryzae* GclD	FAD
*phy*S	PKS-NRPS	*M. mycetomatis* MMYC01_206823	

Table 2.11: Annotation of *phy*. Genes on + strand are annotated as arrows pointing right. Red genes have redox activity, blue is for the core PKS-NRPS, green for *O*-methyltransferase.

The domains of the core PKS-NRPS PhyS were manually analysed by three different browsers: NCBI Conserved Domains Database (CDD); AntiSMASH; and PKS/NRPS ANALYSIS.[246] Noticeably, PhyS contained a broken enoyl reductase domain (ER^0), despite the highly reduced features of **190**, especially in the terminal aliphatic chain. The reductive function is most likely reintroduced by the *trans*-acting enoyl reductase PhyL4. Multiple alignment of PhyS with TenS (tenellin synthetase from *Beauveria bassiana*) and PksD from the human pathogen *Madurella mycetomatis* strain MMYC01_206823 showed the high homology for each domain between each other, especially between *Phyllosticta cirsii* and *M. mycetomatis* (Table 2.12).
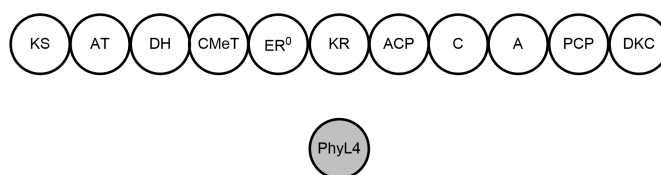
Domain identity matrix

PhyS domain	*vs* TenS	*vs* PksD
KS	58%	75%
AT	53%	70%
DH	33%	54%
MeT	44%	54%
ER^0	-	-
KR	45%	61%
ACP	50%	65%
C	33%	64%
A	38%	76%
PCP	41%	72%
DKC	32%	61%
Global identity	36%	59%

Table 2.12: Comparison of PhyS domains with TenS and PksD.

TenS is responsible in *Beauveria bassiana* for the biosynthesis of the tetramic acid pretenellin A **193** (Scheme 2.5), through condensation of a pentaketide and a tyrosinyl-PCP (**194** and **195** to **196**) and a release mechanism that involves Dieckmann cyclisation to yield the tetramic acid (**196** to **193**). The 5-membered ring undergoes ring expansion to yield the 2-pyridone tenellin **197**.[247,248] It is sensible to compare the two core PKS-NRPS since they probably have similar mechanism for the formation of the hydroxy tetramic acid.
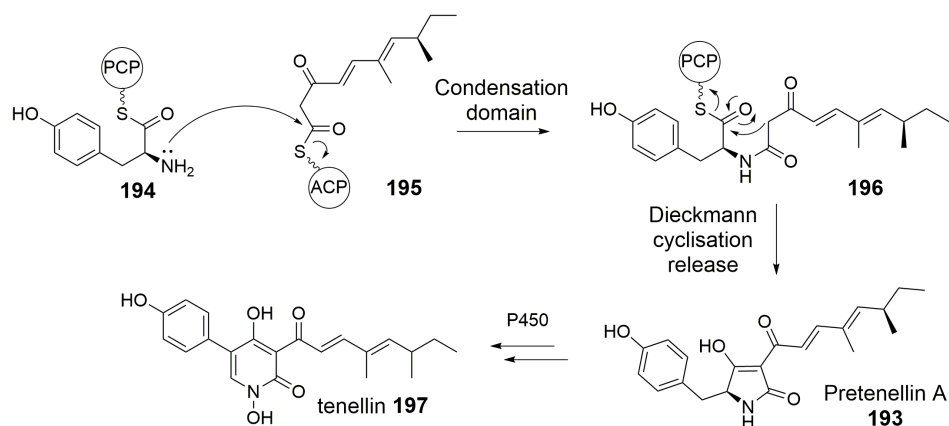
Scheme 2.5: Biosynthesis of pretenellin by fungal PKS-NRPS.

The high identity between PhyS and *M. mycetomatis* PKS-NRPS induced us to perform a homology comparison at cluster level. A ∼50 KBp sequence flanking the MMYC01_206823 gene (PksD PKS-NRPS) was selected and compared to *phy* BGC using the Artemis Comparison Tool. The analysis revealed the two clusters to have high homology (Figure 2.28, Table 2.13). Interestingly, the *trans*-acting enoyl reductase PhyL4 had no correspondent encoding gene in the *M. mycetomatis* cluster and the core PksD PKS-NRPS presented a ER^0 domain, suggesting that the putative product of the *M. mycetomatis* cluster can be only partially reduced to an olefin.
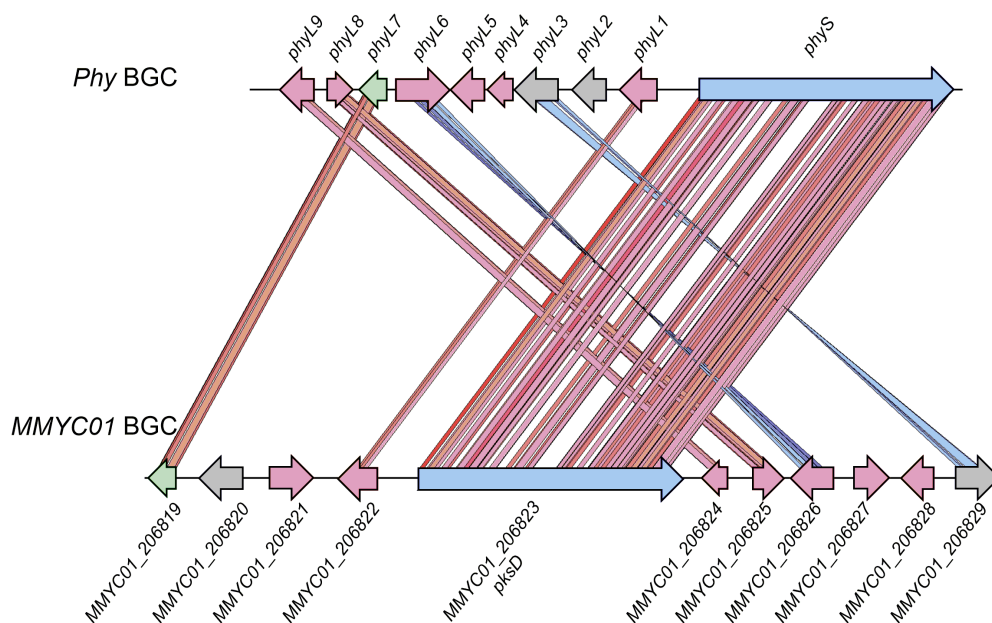
Figure 2.28: *phy* cluster compared to *Madurella mycetomatis* BGC by ACT. Artemis with minimum score cutoff set at 21, and maximum to 515. Threshold was set at 210. Red lines show genes on the same DNA strand, while blue lines show genes on opposite strands. Genes with redox activity are marked in red, blue for the core PKS-NRPS, green for *O*-MeT.

Name	Annotation	Predicted cofactor
MMYC01_206819	*O*-methyltransferase	SAM
MMYC01_206820	Transcriptional regulator	
MMYC01_206821	D-xylose 1-dehydrogenase	NAD(P)$^+$
MMYC01_206822	Bifunctional solanapyrone synthase	
MMYC01_206823	PKS-NRPS	
MMYC01_206824	Salicylate monooxygenase	FAD
MMYC01_206825	1-aminocyclopropane-1-carboxylate oxidase	n/a
MMYC01_206826	Cytochrome P450	Iron, Heme-thiolate
MMYC01_206827	Hydroxy-D-nicotine oxidase	FAD
MMYC01_206828	Bifunctional solanapyrone synthase	
MMYC01_206829	Transport protein	

Table 2.13: Annotation of the *M. mycetomatis phy* BGC homologous.

Particular attention was also dedicated to the enoyl reductase PhyL4, as it probably functions as *trans*-acting ER domain to provide for the broken reductive domain of PhyS (ER^0). The amino acid sequence of PhyL4 was aligned with the two *trans*-acting ER enzymes TenC and LovC, which function during the biosynthesis of tenellin and lovastatin, respectively (Table 2.14). PhyL4 displayed ∼40% identity to TenC and LovC.

Identity matrix		
	TenC	**LovC**
PhyL4	39%	40%

Table 2.14: Comparison between PhyL4 and TenC, LovC.

Phyllostictines and phaeosphaerides belong to the same class of compound, therefore is probable that their producing organisms should contain highly homologous gene clusters. In the original publication, phaeosphaeride A was isolated from strain FA39, which had 97% identity with *Phaeosphaeria avenaria*.[242] MycoCosm[249,250] text research for *Phaeosphaeride spp* gave as equivalent hit *Stagonospora nodorum*, which is its equivalent anamorph (asexual stage).[251] *Stagonospora nodorum* is the closest organism to *P. avenaria* with available genome sequence. PhyS protein was blasted into *S. nodorum* genome (MycoCosm portal) and the highest hit (SNOG_00308.3) used to define a BGC selecting ∼50 KBp flanking region. This sequence (namely SN15) was compared to the whole *phy* BGC using Artemis Comparison Tool, but no close homology was observed beyond the PKS-NRPS and the *trans*-ER system (Figure 2.29, Table 2.15).
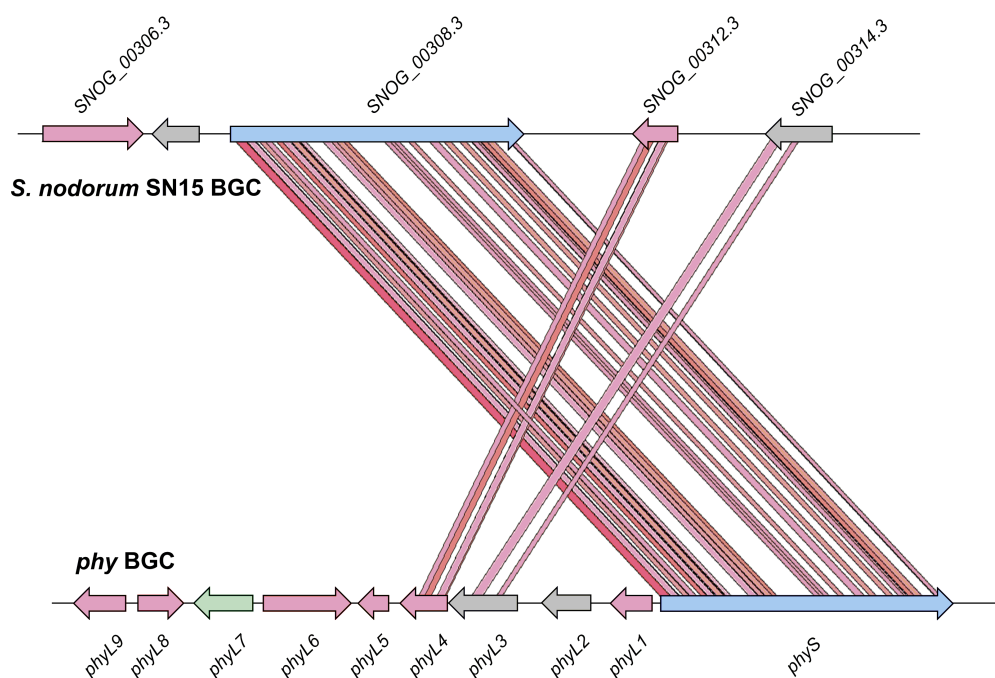
Figure 2.29: *phy* cluster compared to SN15 BGC from *S. nodorum*. Red lines show genes on the same DNA strand. Genes on + strand are represented as arrows pointing right. Red genes encode for redox enzymes, blue genes encode for the core PKS-NRPS, green for *O*-MeT.

SN15 cluster

Name	Annotation	Cofactor
SNOG_00306.3	Cytochrome P450	iron, heme
SNOG_00308.3	PKS-NRPS	
SNOG_00312.3	Short chain alcohol dehydrogenase	NAD(P)$^+$
SNOG_00314.3	Transporter protein	

Table 2.15: Annotation of SN15 cluster from *S. nodorum*.

Phyllostictine A is also structurally similar to pyranonigrin E **198**, whose biosynthetic gene cluster (A1179 BGC) and biosynthetic details have been characterised in *Aspergillus niger* (Figure 2.30).[252,253]
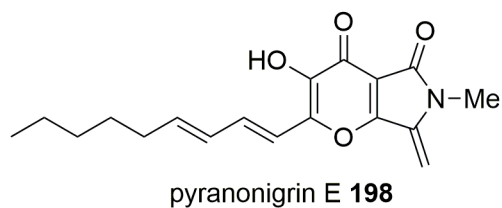
Figure 2.30: Pyranonigrin E produced by *Aspergillus niger*.

The *phy* and A1179 clusters were compared by ACT. Despite the structural similarities between **190** and **198**, genes from the *phy* BGC showed almost no homology to the pyronigrin E BGC (Figure 2.31, Table 2.16)
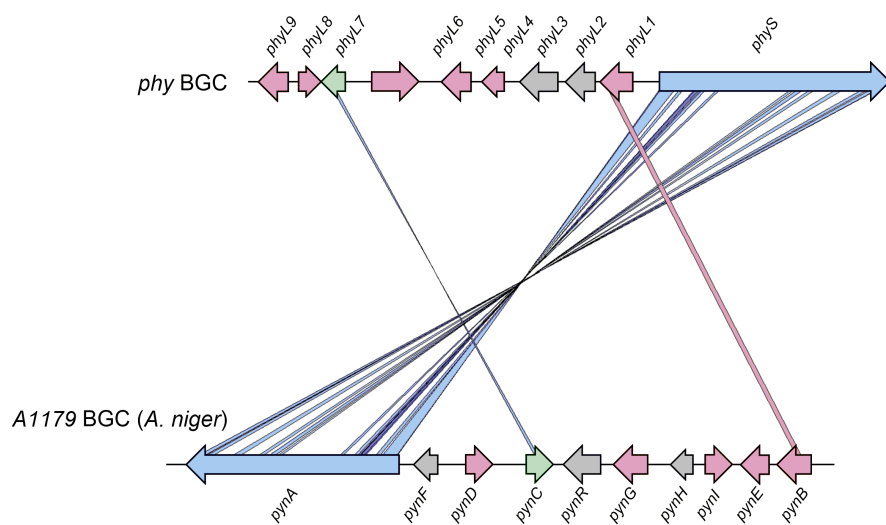
Figure 2.31: *phy* cluster compared to pyronigrin A1179 in *A. niger* by Artemis comparison tool. Red lines show genes on the same DNA strand, while blue lines show genes on opposite strands. Genes on + strand are represented as arrows pointing right. Red genes encode for redox enzymes, blue genes encode for the core PKS-NRPS, green for methyltransferase.

Pyranonigrin A1179 BGC

Name	Annotation	Predicted cofactor
*pyn*A (ANI_1_1528094)	PKS-NRPS	
*pyn*F (ANI_1_1530094)	Transport protein	
*pyn*D (ANI_1_1532094)	Cytochrome P450	iron, heme
*pyn*C (ANI_1_1534094)	Methyltransferase	SAM
*pyn*R (ANI_1_1536094)	Transcription factor	
*pyn*G (ANI_1_1538094)	Oxidase	FAD
*pyn*H (ANI_1_1540094)	Aspartyl protease	
*pyn*I (ANI_1_1542094)	thiolesterase	
*pyn*E (ANI_1_1544094)	Oxidoreductase	NAD(P)$^+$
*pyn*B (ANI_1_1546094)	Oxidase	FAD

Table 2.16: Annotation of the *A. niger* A1179 cluster.

The C-terminal DKC domain of PhyS was compared by multiple alignment to the release domain of various tetramic acid producing PKS-NRPS systems using Clustal Omega with default settings. We included TenS, PynA and EqiS, producers of β-oxidised tetramic acids; FusS and ACE1, producers of pyrrolidinone systems; and the reductase Lys2 involved in lysine biosynthesis (Scheme 2.13, Section 2.4.4). Proteins accession numbers: PhyS (*P. cirsii*, KY682688.1), TenS (*B. bassiana*, A0JJU1), PynA (*A. niger*, ANI_1_1528094), EqiS (*A. fumigatus*, EDP53404), FusS (*F. verticillioides*, *XP_018758499*), ACE1 (*M. oryzae*, Q6ZX14) and Lys2 (*S. cerevisiae*, NP_009673.1).

Protein alignment showed PhyS, PynA, EqiS mutated in the NADPH binding domain (yellow), and TenS and PynA mutated in the catalytic reductive site (red). This is in accord with the observation that the mentioned PKS-NRPS form tetramic acids. Therefore, we concluded that PhyS releases the polyketide chain by *intra*-molecular aldol condensation. FusS, ACE1 and Lys2 showed no mutation in either the two domains, in accord with their pyrrolidinone products (Figure 2.32).

```
PhyS_red   VLVTGATGNLAKRVLPVLEANPLVGRIVCVAVRDK..PQE...A...........
TenS_red   ILLTGATGFLGGQLLERLVQSPRVSTVHCVAVPVD..EQS...L...........
PynA_red   ILLTGIDTFIGLHMLATILSNNHNATVYVIGIHDE..LTA...DHLVEGLTKYKL
EqiS_red   ILYTGATGHSAKYVLEKLVKDERVSKVYCVAVRSP..D.................
FusS_red   VALTGATGFLGVSILETLLEDKRVSKVHCLAVRSS..SNT...............
ACE1_red   VLLTGATGFLGRQLMAFLLRQPSVKRIHCLAVRGG..APP...S...........
Lys2       VFVTGVTGFLGSYILADLLGRSPKNYSFKVFAHVRAKDEEAAFARLQKAGITYG.


PhyS_red   SRSIL......RGDKIVQHAGDLSKPRLGLTVEQIQSLSGEVDVILH.MG.....
TenS_red   LEPFLQQQADGTRRKVRCYIGNLAAPALGLTAADQTALSQTADVIVH.AG.....
PynA_red   LDAHL..STEDVLSRTCAVPGKMTSPRFGLAEEAFRNLADKVRVIFN.IA.....
EqiS_red   .KLSI......VSDKIVQFSGDLLDKRLGLMEEQFDFLANEADVIFH.AA.....
FusS_red   SDPVF......SSSRVACYPGDLSLPRLGLSQEQFDQLANAVDRIIH.NG.....
ACE1_red   SAAPF......SDPRVSIHAGDLNAPHLGLGEAVAELLFAQADVIIH.NG.....
Lys2       ......TWNEKFASNIKVVLGDLSKSQFGLSDEKWMDLANTVDTIIH.NG.....


PhyS_red   .......ALRSFWDNYHMLRLSNVQSIKELVQLAAPRR...IPIHFVSTSGVLPR
TenS_red   .......SMGHCLNTYATLAAPNFASTRHLCSLALSRSP.PIPLAFASSNRVALL
PynA_red   .......ADVSLLKTYVDLKTVNTSAILTLIELATSSHGHLLEIHHLSTWSVPHL
EqiS_red   .......ANRSFWDNYQVMRRANVLPAHTLVALAARRK...VPIHFMSSAAVHLF
FusS_red   .......ADVSFLKTYQSLQRSNVSSSRELARMAITRR...IPMHFVSTGGVVQL
ACE1_red   .......ADVSFLKTYATLRATNVGSTRELARLAAPRR...IPFHFVSSASITQL
Lys2       .......ALVHWVYPYAKLRDPNVISTINVMSLAAVGKP..KFFDFVSSTSTLDT


PhyS_red   DIL.............DSATS.QAPSSAAAYQPPVD...GSNGYVASKWVGERIL
TenS_red   TGS...............TA.PPPASAAAFPPPP....GAQGFTASKWASEAFL
PynA_red   QTWKKTS.......RTRAFAS.NREEDPSHFTPPTA...DEYGYFKSRWAAEMYL
EqiS_red   G..............RSTDE.DSPETPATHPPPAD...GKDGYLATKWAAEKIF
FusS_red   TGQ............DGLDE....VSVIDSTPPND...GSLGYVASKWASEAIL
ACE1_red   TGL............DEFGE....ASMAAWAPPADPRGMSGGYAAAKWASEVLL
Lys2       E..........YYFNLSDKLVSEGKPGILESDDLMNSASGLTGGYGQSKWAAEYTI
```

Figure 2.32: Protein alignment between the release domains of PhyS, TenS, PynA, EqiS, FusS, ACE1, and Lys2. NADH/NADPH binding GXXGXXG domain is highlighted in yellow; the catalytic triad S-Y-K in red. Mutations are marked in black, and in blue the consensus.

2.3.8 Molecular biology and Fungal Transformation

Based on the bioinformatic findings, we designed a knock out cassette to specifically target the core *phy*S gene and other genes encoding the tailoring enzymes. First, a selection method based on antibiotics was devised in order to choose a selection gene. Hygromycin B (HygB) and zeocin were tested on PDB plates at an increasing concentration (25, 50, 100, 150 and 200 μg/ml) and compared to a control growing in absence of antibiotics. The fungus proved to be sensitive to low concentration of both antibiotics as no colony could grow at concentrations above 25 μg/ml (Figure 2.33). Therefore, we chose HygB

as we already possess the vector pTH-GS-eGFP with the hygromycin B resistance gene *hph* under the control of the constitutive promoter P_{gpdA} from *A. oryzae*. The *hph* gene was isolated from *E. coli*[254, 255] and it encodes a kinase that deactivates the toxin by phosphorylation. Additionally, pTH-GS-eGFP contains the enhanced Green Fluorescent Protein (eGFP) as a secondary reporter gene.
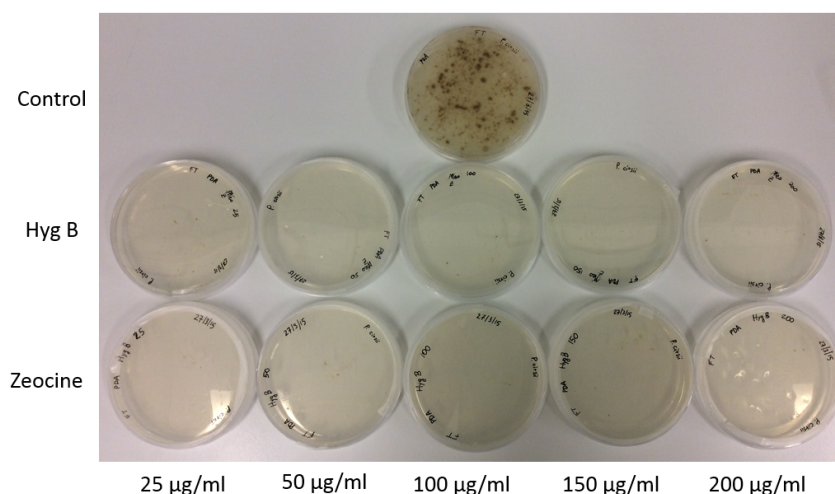
Figure 2.33: Antibiotic testing against *P. cirsii*.

Genetic manipulations of *P. cirsii* have not previously been described, thus we established a protocol for bipartite recombination through protoplast generation. Before trying any targeted KO, we transformed *P. cirsii* with pTH-GS-eGFP to prove the *hph* gene to confer resistance to the transformed fungus. Protoplasts were generated by enzymatic digestion of the cell wall using a mixture of lysing enzymes from *Trichoderma harzianum*, which possess cellulase, chitinase and protease activiteis,[256] and driselase from basiodiomycota, which has additionally laminarinase and xylanase activities.[257] Transformation of ∼10 μg of pTH-GS-eGFP led to single colonies growing on HygB plates (50 μg/ml), that could be propagated on secondary plates. The presence of *hph* gene was confirmed by PCR on genomic DNA. Fluorescence microscopy did not give any insights, as the WT cells also displayed fluorescence (Figure 2.34).
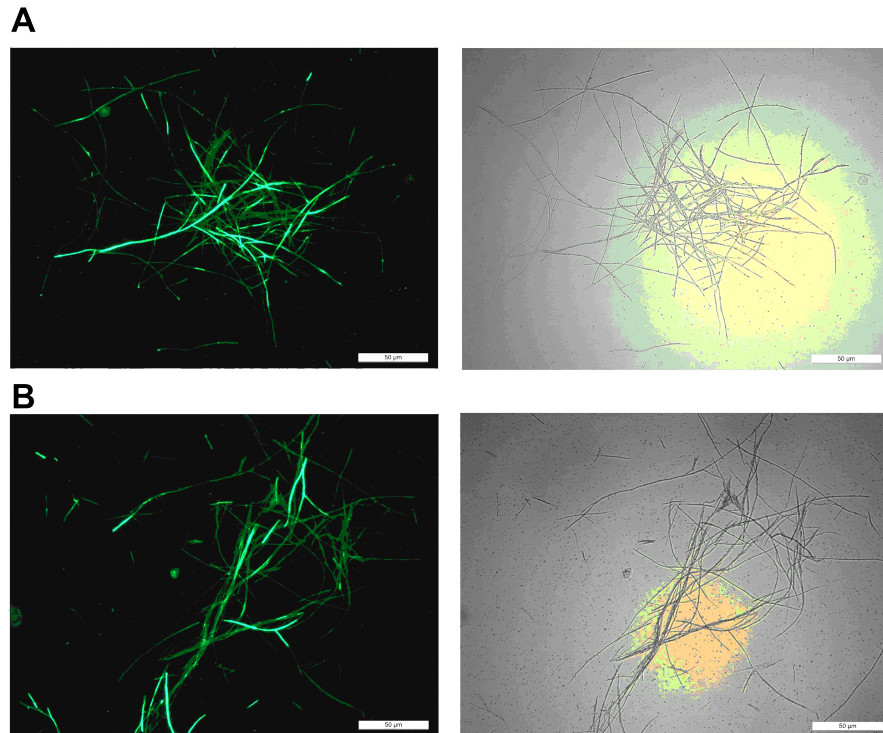
Figure 2.34: A pTH-GS-eGFP transformed *P. cirsii* under UV and white light. **B** WT *P. cirsii* under UV and white light.

The KO cassette was built exploiting yeast recombination. Three fragments were joined together using this technique: the target left sequence (T_L; ∼1 KBp), the hygromycin B resistance (*HygR*; ∼4 KBp) composed by the promoter P_{gpdA} and the resistance gene *hph* and the target right fragment (T_R; ∼1 KBp). T_L and T_R were generated by Taq PCR from gDNA using tailed primers with 30 nt overlapping to the downstream fragment. T_L and T_R are responsible for the site-specific homologous recombination during the KO generation. *HygR* was generated from high fidelity Q5 PCR, also using primers bearing a 30 nt tail on the terminal 3'homologous to the 5'of T_R. High fidelity Q5 polymerase was used to guarantee functionality of the resistance protein. The T_L-HygB-T_R construct was cloned into pE-YA, which carries two selection markers: *ura*, for revertant selection of uracil auxotroph yeast, and *kan* for kanamycin selection during *E. coli* vector amplification (Scheme 2.6).
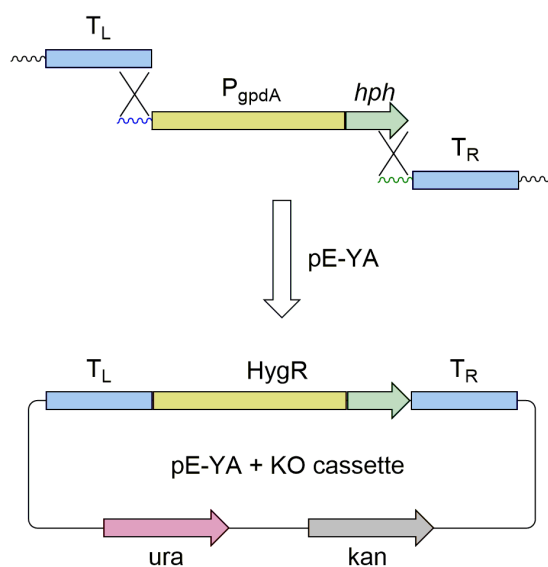
Scheme 2.6: KO cassette construction into pE-YA vector by yeast recombination.

A total of 6 constructs were generated by yeast recombination: namely *phy*S, *phy*L1, *phy*L4, *phy*L6, *phy*L8 and *phy*L9 KO cassettes. Genes with an oxidoreductive activity were prioritized as being more interesting in the discovery of the biosynthetic pathway. The success of the whole process was confirmed by enzymatic digestion.

As the bipartite transformation requires, two PCRs were performed using the KO cassette as template in order to split the antibiotic resistance into two overlapping, non functional fragments α and β, later transformed simultaneously into the fungus (Scheme 2.7). Three events of recombination are required in order to rebuild the resistance and to deplete the target gene, while random insertion of either α or β should not yield selection.
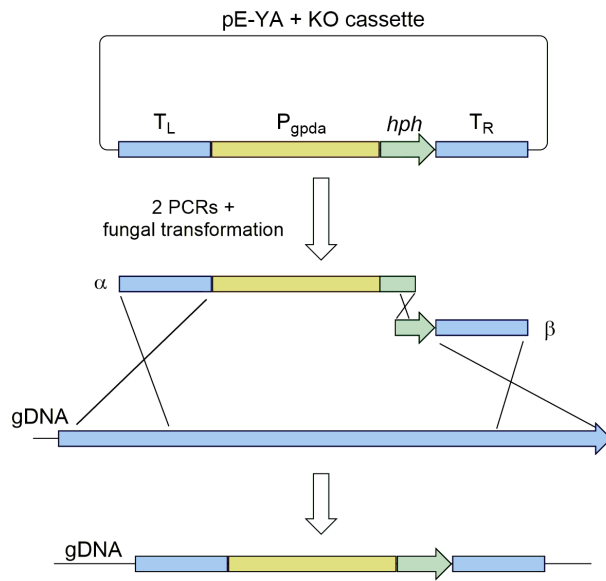
Scheme 2.7: α and β fragments generation and triple homologous recombination into the fungus genome.

Transformed protoplasts were plated in buffered PDB with HygB 100 μg/mL and the growing colonies transferred onto secondary plates with the same concentration of antibiotic. Transformant mycelia were resistant to the antibiotic, whilst WT was incapable of growing (Figure 2.35).
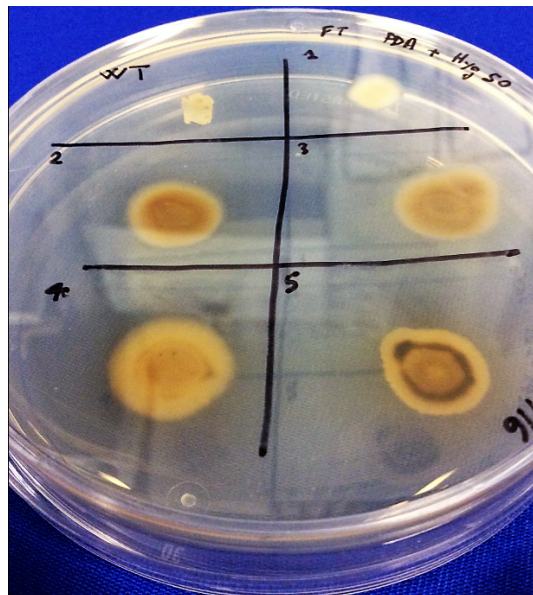
Figure 2.35: WT and transformed *P. cirsii* on secondary plate HygB 100 μg/ml.

Colonies able to grow on secondary plates were screened by PCR of genomic DNA to reveal the exact locus of insertion. In particular two patterns were investigated: *whole* PCR (wPCR) and *external* PCR (extPCR). wPCR spanned the entire target gene (except for the core *phy*S) and revealed whether the WT gene was still in place: indeed, due to the calculated elongation time, the Taq polymerase is incapable to yield the PCR product in case of insertion. The KO cassette is always longer than the target tailoring gene, so no product is observable in a real mutant. Special mention is for the core *phy*S gene, which is much longer ($\sim$6 KBp), so instead of amplifying the whole gene, a sequence of 1 KBp comprised between the T_L and T_R flanking regions was inquired. Indeed, the mentioned sequence must be deleted in a true mutant, and would give a PCR product only in an intact gene. extPCR on the other hand should always give a positive PCR product in a true mutant. Indeed, it was designed to cover the sites of homologous recombination, with the forward primer at the very end of the *HygR* and the reverse primer located outside the gene. A positive extPCR would indicate site-specific integration of the construct (Scheme 2.8).
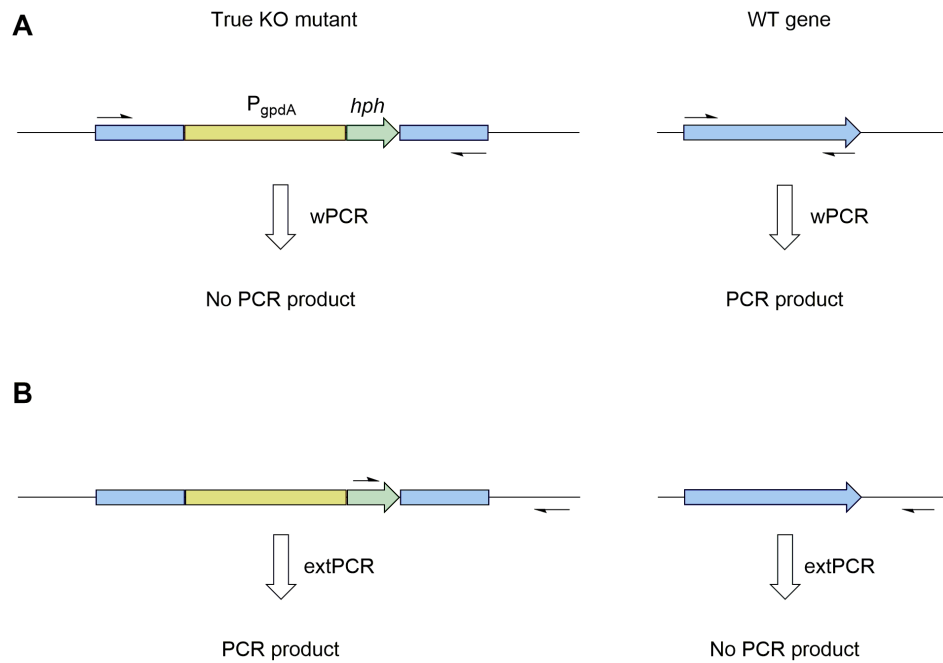
Scheme 2.8: A Whole PCR, from start to end of the gene. **B** External PCR, from the edge of the *HygR* to a region outside the target gene.

The transformation rate was rather low, with no true mutant observed beside ΔphyS and ΔphyL6 (Table 2.17).

Target	Transformations	First selection (colonies)	Second selection (colonies)	Real mutants colonies
*phy*S	10	1	1	1
*phy*L1	48	52	20	0
*phy*L4	45	49	18	0
*phy*L6	11	12	5	3
*phy*L8	51	50	33	0
*phy*L9	58	64	24	0

Table 2.17: Transformants screening.

We were able to achieve solely two true KO mutants of the core *phy*S PKS-NRPS and the P450 cytochrome oxidase *phy*L6. The mutants were fermented in producing conditions and their chemotype investigated.

2.3.9 Δ*phy*S and Δ*phy*L6 fermentation and products characterisation

Δ*phy*S mutant grew in M1D with altered phenotype (Figure 2.36). The loss of yellow colour may be a consequence of the loss of the principal metabolite phyllostictine A **190**, which is yellow itself.
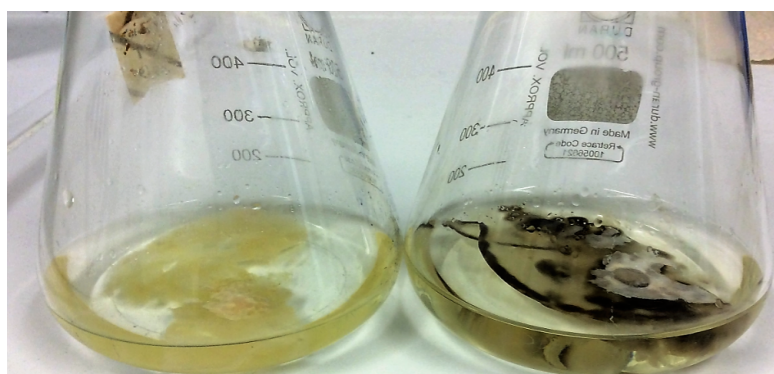
Figure 2.36: WT *P. cirsii* (left flask) and Δ*phy*S mutant (right flask) growing in producing M1D medium.

Δ*phy*S completely lacked the ability to produce **190**, or any of the related phyllostictines B-D (Figure 2.37).

ΔphyS
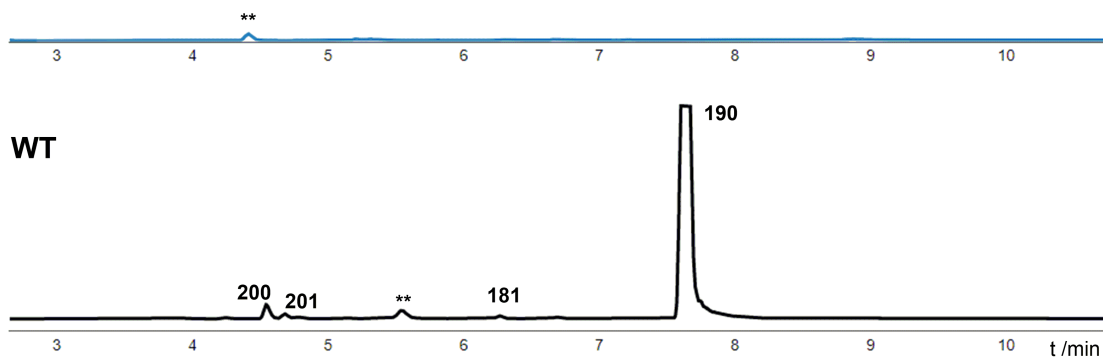
Figure 2.37: ELSD chromatogram of ΔphyS *vs* a WT control. y axes are linked. ** marks unrelated peaks.

Fermentation, raw extraction and LCMS analysis of ΔphyL6 not only led to the abolition of the phyllostictines, but also showed a new peak with mass of 295 Da, 30 units less than **190** (Figures 2.38 and 2.39). We named this compound phyllostictine E **199**.
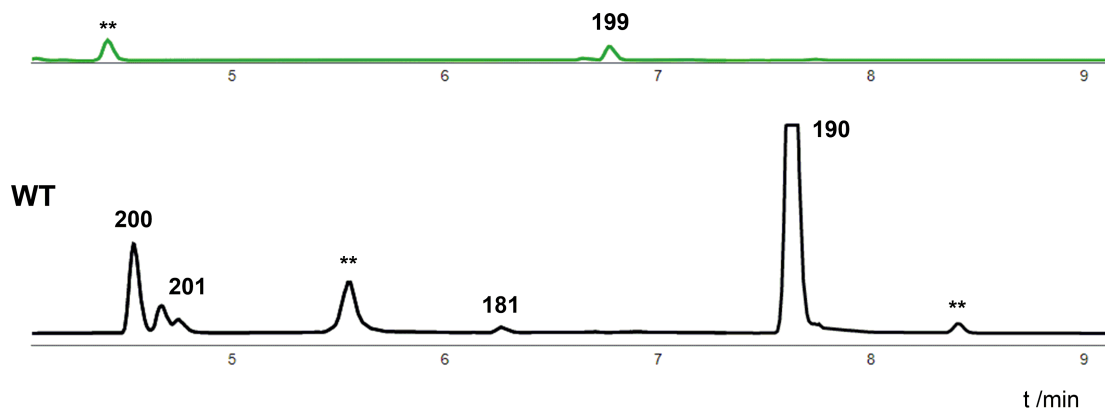
ΔphyL6

Figure 2.38: ELSD chromatogram of ΔphyL6 *vs* a WT control. y axes are linked. ** marks unrelated peaks.
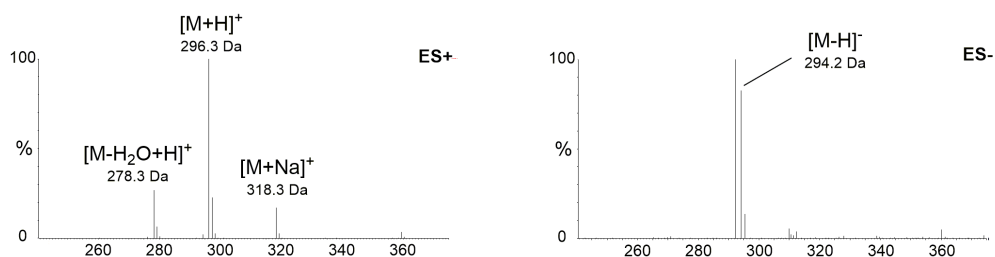
Figure 2.39: ES+ and ES- mass spectra relative to peak **199**.

High resolution MS pointed this compound to have molecular formula $C_{16}H_{25}NO_4$. Complete HRMS data are hereby reported:

Phyllostictine E (**199**): UV (diode array HPLC, H_2O/CH_3CN) $\lambda_{max} = 259.6$ nm; LCMS m/z 296.3 $[M]H^+$, 318.3 $[M]Na^+$, 278.3 $[M + H - H_2O]^+$, 294.2 $[M - H]^-$; HRESIMS m/z 296.1862 $[M]H^+$ (calcd for $C_{16}H_{26}NO_4^+$, 296.1862)

NMR data proved the new compound **199** to be the corresponding -NH tetramic acid of **190**, devoid of the methoxy group.

| | CDCl$_3$, 313 K | |
| | δ_C/ppm | δ_H/ppm |
position	125 MHz	500 MHz
1	170.4	-
4	159.9	-
3	137.1	-
5	107.1	-
17a	93.9	5.01 bs (1H)
17b	93.9	4.86 bs (1H)
8	86.6	4.02 m (1H)
7	72.1	-
6	69.9	4.55 bs (1H)
11	32.2	1.29 m
13	29.7	1.26 m
14	29.5	1.34 m
9	27.9	1.83 m (2H)
10a	26.9	1.61 m (1H)
10b	26.9	1.39 m (1H)
12	23.0	1.31 m
18	16.4	1.25 s (3H)
15	14.4	0.88 t (3H), J = 6.9 Hz
N*H*	-	7.1 bs

Table 2.18: ^{13}C- and ^{1}H-NMR data for **199**. Not all signals could be accurately integrated due to overlap.

The recorded ^{13}C- and ^{1}H-NMR data for the new compound **199** matched the value of **190**, with the exception of the signal relative to position 16, which disappears in **199** spectra (Figure 2.40). ^{1}H-NMR showed -NH at 7.1 ppm as a broad singlet, supporting the structure **199**.
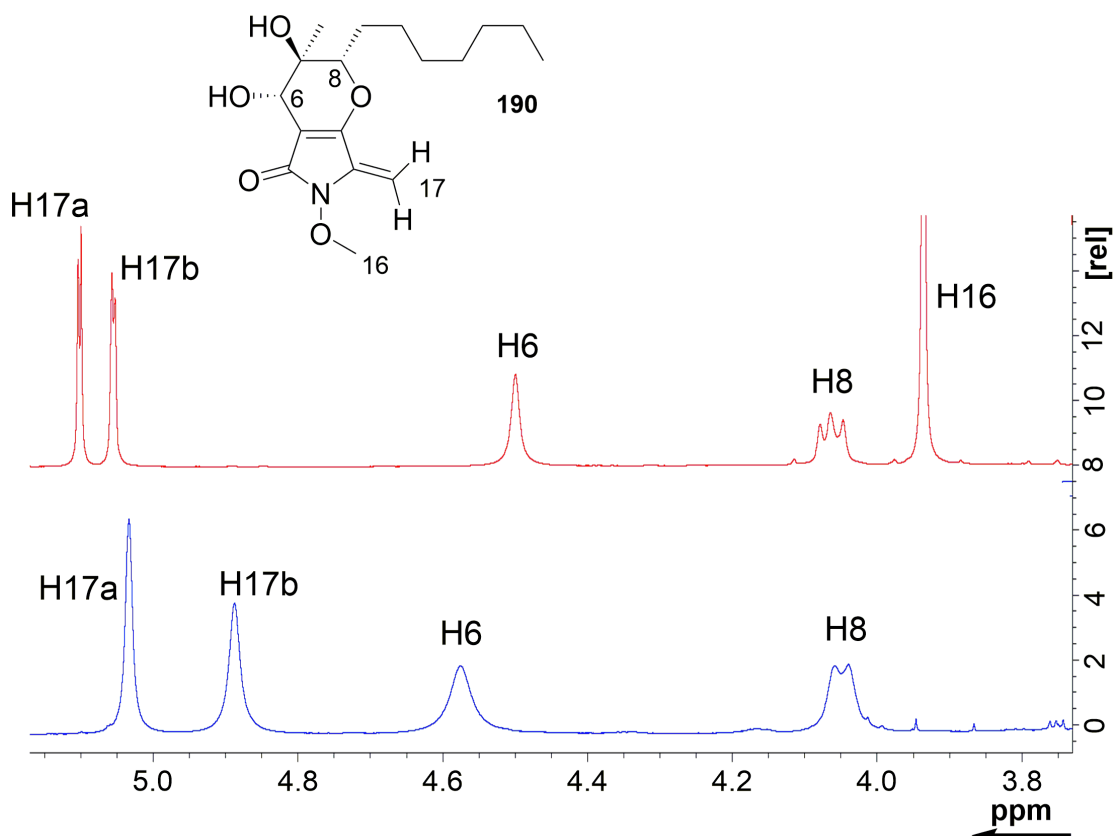
Figure 2.40: ^{1}H-NMR of phyllostictine A (top) compared to phyllostictine E (bottom).

2.3.10 Biological activity

Phyllostictine A was tested on bean leaves by Sona Mohammadi. The goal was to observe a possible effect on the stomata, thus understanding the mechanism of toxicity of the compound. Stomata are epidermal pores controlled by guardian cells that regulate gas exchange in plants. Water, CO_2 and O_2 diffuse through these openings, regulating photosynthesis, respiration and turgor of the plant. Controlling the open-closed state of stomata is essential for plant survival.

In the experiment 10 μg/ml of phyllostictine A were sprayed on the leaf tissue and compared to a control by microscopy pictures after 30 min (Figure 2.41). Cell vitality was confirmed by Neutral Red 2% (not shown).
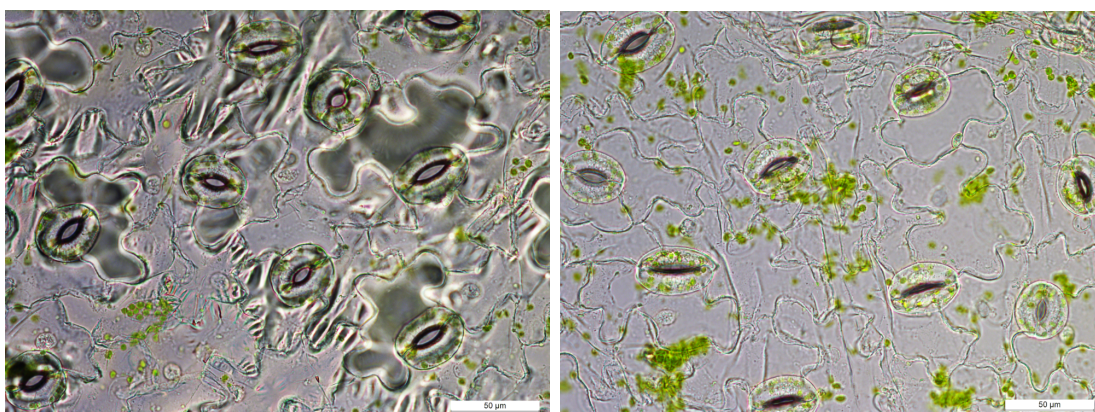
Figure 2.41: Control leaf (left) *vs* phyllostictine A treatment (right), microscopy picture.

The result suggested that **190** forces the stomata tightly closed, unbalancing gas exchange, thus leading to necrosis of the plant tissues. These data are insufficient to understand the mechanism of action of **190**, but further investigations could be done to define the molecular target within the guardian cells.

2.4 Discussion and Conclusions

In our hands, *P. cirsii* produced the class of compounds known as phyllostictines when cultured for 3-4 weeks in the chemically defined medium M1D and static condition. Growing the fungus for 14 days was sufficient to obtain reasonable amounts of the major compound phyllostictine A. Purification, HRMS and NMR data acquisition of the latter led to identical values reported in literature, confirming to deal with the same substance.

2.4.1 Feeding experiments and structure reassignment

Feeding experiments showed that phyllostictine structures reported in literature are erroneous, with random incorporation pattern of labeled acetate building blocks into **168** structures. 2D NMR data further confirmed the literature structures to be wrong, with strong apparent HMBC correlation between nuclei 8 bonds far apart from each other and no HMBC correlation between close 2/3-bond distance. Full NMR data obtained in $CDCl_3$ and DMSO-d_6, allowed revision of the structures of phyllostictines to a series of tetramic acid containing metabolites, related to the (para)phaeosphaeride class of metabolites produced by *(Para)Phaeosphaeride ssp*[258] (Figure 2.42). A detailed discussion about tetramic acid natural products biosynthesis is addressed further on (Section 2.4.4).
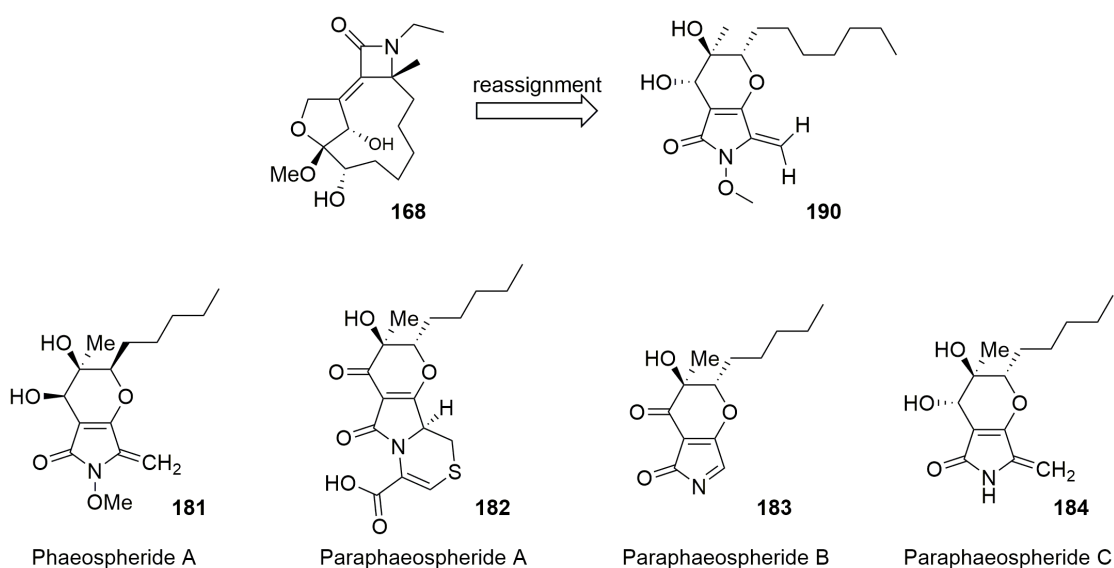
Figure 2.42: Reassignment of phyllostictine A structure and related tetramic acid natural compounds.

The bicyclic tetramic acid-dihydropyran core of phyllostictine A was further corroborated by additional feeding experiments. Correlation with the ^{13}C labelling positions derived from acetate confirmed that **190** is derived from a hexaketide fused to a 3-carbon and nitrogen containing moiety forming the tetramic acid. Various tetramic acid-containing natural products, such as prefusarin **78** and pretenellin A **193** have been reported to be biosynthesised by hybrid PKS-NRPS system,[79, 248] suggesting that also the class of phyllostictines may follow the same assembly line. Not only does the acetate pattern show consistency throughout the hexaketide moiety, but also the incorporation of [1-^{13}C]-L-alanine was congruent within the new structure as it specifically labeled the C-4 of **190** indicating that this amino acid is a precursor in the biosynthesis. Other amino acids may be used as building block by the NRPS module of the core enzyme: serine, for example, might be used as well and may undergo H_2O elimination to form the terminal methylene, but the high rate of incorporation of alanine, estimated at 5.5%, suggested that this amino acid is the real building block used in the biosynthesis. [*Methyl*, ^{13}C]-L-methionine feeding confirmed C-16 and C-18 to be derived from a tailoring SAM-dependent *O*-methyl transferase and from the *C*-MeT domain of PhyS respectively (Scheme 2.9).
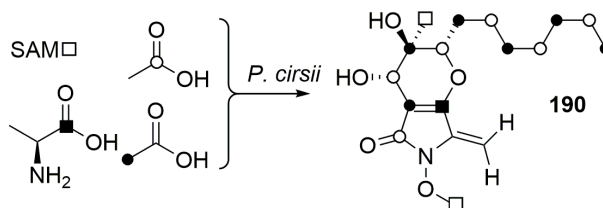
Scheme 2.9: Incorporation of ^{13}C-building blocks throughout **190**.

Reaxys research of natural product isolated from *Phaeosphaeride ssp* and *Stagonospora spp* gave useful hints in determining the real structure of phyllostictines, given the many analogies between these tetramic acids. In particular phyllosictine B and phaeosphaeride A were found to be the same compounds by NMR chemical shift comparison. The absolute configuration of phaeosphaeride A is known, as it was determined by crystallography by Berestetskiy and collaborators,[259] and comparison of the specific rotation between phyllostictine B and phaeosphaeride A led to the conclusion that they present the same absolute configuration. Since phyllostictine A and B are the products of the same pathway, the first deriving from a hexaketide and the latter from a pentaketide, we assumed that also their absolute configuration must be identical.

2.4.2 Bioinformatics

Analysis of the ITS sequence of *Phyllosticta cirsii* suggested that, rather then belonging to the *Phyllosticta* species, it is closely related to *Phaeosphaeride ssp* and *Stagonospora spp*, which also synthesise the phaeosphaeride **181** and the paraphaeosphaerides **182 - 184**. We did not push forward the phylogenetic analysis of *P. cirsii*, but we hereby suggest that more accurate investigation needs to be done in order to reclassify and rename this organism.

Genome sequencing and mining found 32 putative BGC, mainly for type I/III polyketide, terpene and NRP biosynthesis and only two clusters encoding hybrid PKS-NRPS systems. The candidate gene cluster for phyllostictine production (BGC *phy*) was constituted by a core PKS-NRPS (*phy*S), tailoring redox enzymes and a single *O*-MeT. The presence of the *O*-MeT encoding gene was determinant in the choice of *phy*, as phyllostictines do have an *O*-methylation. Domain analysis of the core PKS-NRPS revealed a broken ER domain, whose activity is compensated by the *trans*-acting *phy*L4 enoyl reductase.

2.4.3 KO experiments

The candidate BGC was confirmed by KO experiments, *via* protoplast transformation. Establishment of a transformation protocol to obtain healthy and numerous protoplasts

90

was not straight forward. Overgrown cells could not generate abundant protoplasts, nor could clumped cells obtained from shaking condition. The ideal conditions to maintain protoplast homeostasis were found to be osmotic buffer 1.3 M NaCl and agar buffed with 1 M sorbitol at 24-30 °C. Incubating the protoplasts in ice for less than 30 minutes did not lead to major cell mortality. The best result was achieved digesting young cells (2-3 days old), grown statically in rich medium, with *Trichoderma* lysing enzymes (10 mg/ml) and driselase from *Basidiomycetes* (5 mg/ml) in 1.3 M NaCl solution for ∼3 hours. Longer digestion would also give a similar result, but it is not recommended as protoplasts tend to fuse, yielding syncytia.

KO experiments targeting the core *phy*S prevented the biosynthesis of **190** and related compounds, while disruption of the P450 cytochrome oxidase *phy*L6 led to the intermediate **199**, structurally related to **190**, lacking the *N*-methoxy group. This was ascertained by NMR: when comparing chemical shifts of **190** and **199**, they were very similar to each other, especially in the aliphatic lateral chain, with the exception of the C-16 signal, that disappeared in phyllostictine E dataset. The production of **199** was noticeably reduced compared to the related **190** in the WT, suggesting that the intermediate could be degraded by the fungus. For this reason, we could not isolate any analogue of phyllostictines B-D. We were unable to purify enough of phyllostictines C and D for full structural characterisation, but HRMS analysis suggested that they must be a hydroxylated congener of phyllostictine A and its carbonyl-homolog respectively. Since in the literature NMR data for phyllostictines C and D the methyl signal corresponding to C-18 is missing in both cases we assume that these changes probably mean that phyllostictine C is the C-18 alcohol (**200**, Scheme 2.16, Section 2.4.6), consistent with a new carbon resonance at 68.1 ppm, and Phyllostictine D is the corresponding C-18 aldehyde (**201**, Scheme 2.16, Section 2.4.6), consistent with the appearance of a carbon resonance at 210 ppm in the original literature data.

2.4.4 Phyllostictines and other tetramic acid compounds

Tetramic acid natural compounds structurally similar to phyllostictines have been previously reported, such as pyranonigrin E **198** from *Aspergillus niger* and spirostaphylotrichin A **202** from *Staphylotrichum coccosporum*. Other tetramic acid containing compounds worth of mention are fusarin C **64**, equisetin **203** and cytochalasin K **204** (Figure 2.43).
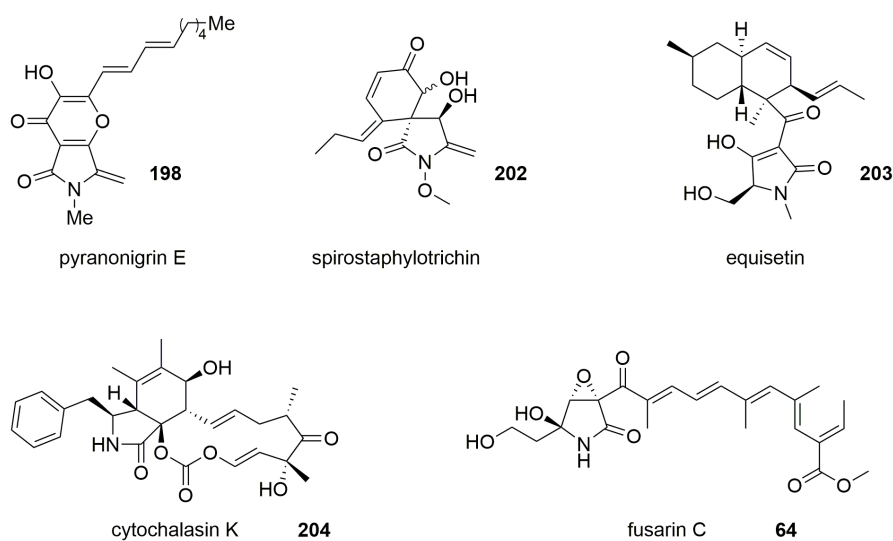
Figure 2.43: Tetramic acid natural products.

The biosynthesis of spirostaphylotrichin A **202** is relatively well understood through labeling studies performed by Peter Sandmeier and Christoph Tamm in 1989,[260] it is probably the product of a PKS-NRPS hybrid that brings together five units of acetate/malonate and aspartic acid (Scheme 2.10).
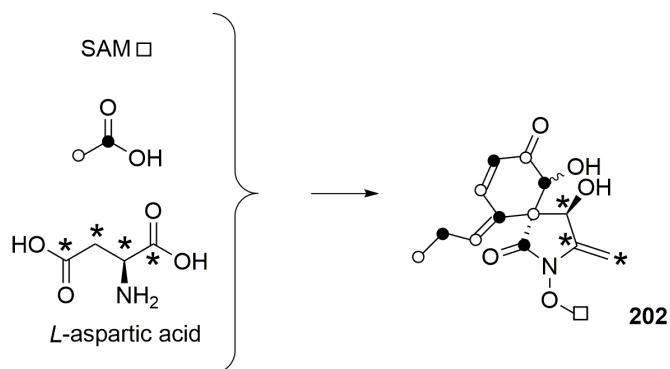
Scheme 2.10: Labeling experiments of spirostaphylotrichin A.

Although the biosynthetic building blocks of **202** are understood, the detailed characterisation of **202** biosynthesis has never been investigated. Nothing is known about the oxidation level of the PK moiety, nor the releasing mechanism from the PKS-NRPS. It is possible that the PKS system produces a highly reduced pentaketide **205** which is fused to L-aspartyl-PCP **206** to form **207** (Scheme 2.11). The chain is probably released by Dieckmann cyclisation (DKC), forming a tetramic acid intermediate **208**, which can be further processed to **202** with loss of CO_2 and, possibly, water for the formation of the

terminal methylene group (**209** to **210**). The 6-membered ring may form from nucleophilic attack on C-7 by C-2, with consequent reduction of carbonyl C-3 to hdroxyl (**210** to **211**). The conjugated double bond system in **202** is probably derived from dehydration of C-7 as the double bonds lies within one acetate unit (**211** to **212**). P450-mediated oxygenation on C-4 and SAM-dependent methyltransferase might yield the final product (**212** to **202**, Scheme 2.11).
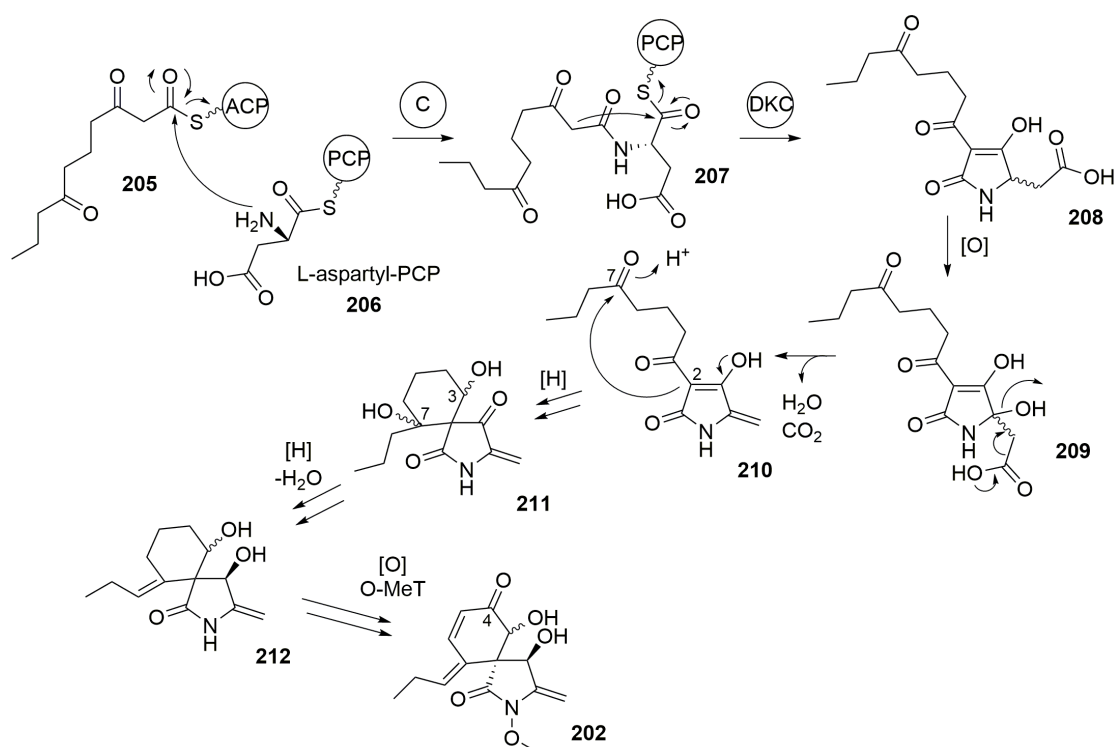
Scheme 2.11: Speculative biosynthetic patway of spirostaphylotrichin A.

The offload mechanism of the natural product from the core enzyme is a key features in PKS-NRPS systems. A reductive release (R) or a Dieckmann cyclisation (DKC) sets different chemistry in the downstream processing of the natural product (Scheme 2.12).[75] In case of pyrrolidinone systems **213**, a reductive release instead of a Dieckmann cyclisation, sets different oxidation level at the β-carbon. A functional reductive domain (R) catalyses the attack of a hydride to the thiolester to release the correspondent aldehyde (**214** to **215**), which can undergo Knoevenagel cyclisation, yielding a pyrrolidinone moiety (**215** to **216** to **213**). A small mutation in the catalytic site of the R domain leads to DKC mechanism of release that forecasts intramolecular nucleophilic attack on the thiolester and PCP-SH domain as leaving group (**214** to **217**) to yield a tetramic acid (**218** to **217**, Scheme 2.12).[261] The fate of the various tetramic acid precursors depends from their release mechanism and from tailoring enzymes, often oxidative, that expand their

scaffold complexity. For example, the epoxy group of fusarin C **64** is inserted by virtue of the olefin on the pyrrolidinone moiety of prefusarin C **78** (Scheme 2.12 **A**; Scheme 1.12 section 1.3). A pyrrolidinone precursor is also likely involved in cytochalasin biosynthesis by Diels-Alder cyclisation (**220** to **221**, Scheme 2.12 **B**).
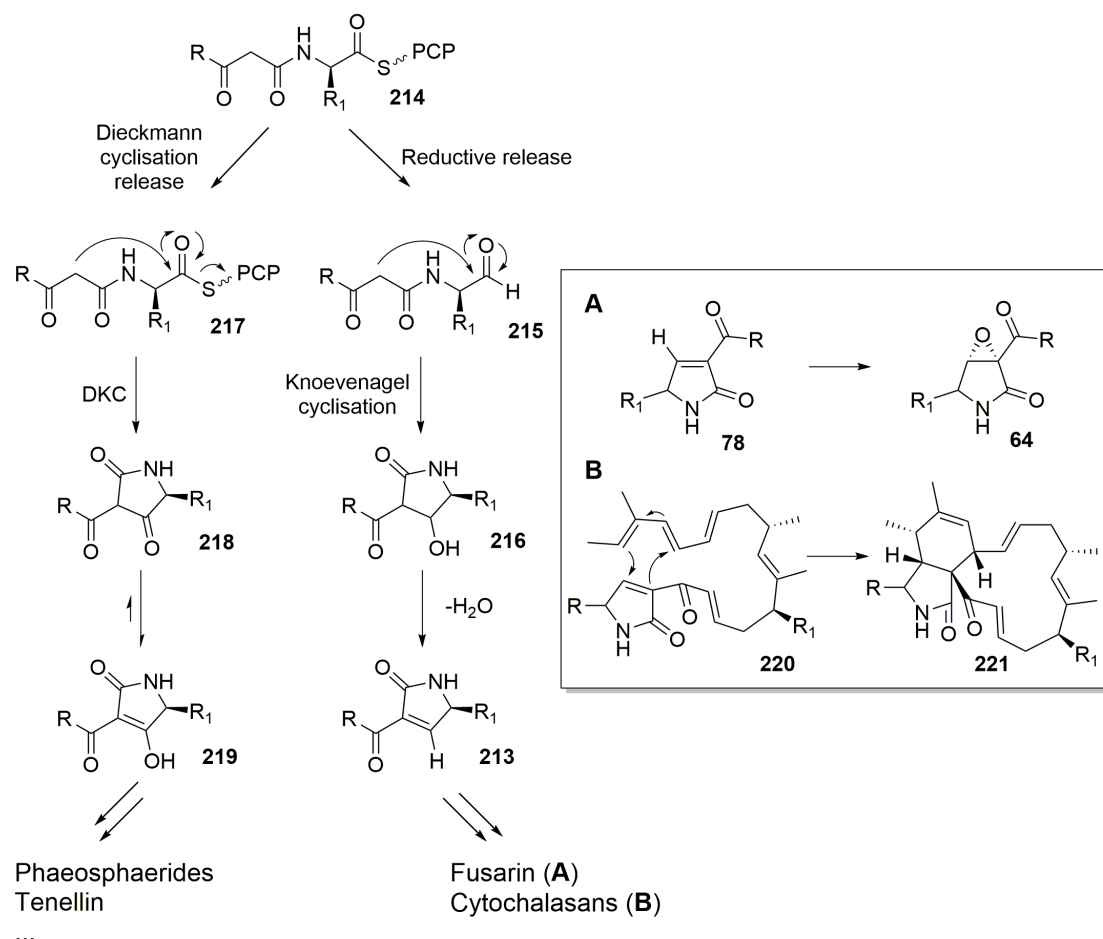
Scheme 2.12: Possible mechanisms of chain release from PKS-NRPS systems. **A** Epoxydation of pyrrolidinone (from fusarin C biosynthesis). **B** Diels-Alder cyclisation involving pyrrolidinone (cytochalasans).

Studies on lysine biosynthesis have shed light on reductive chain release mechanism for NRPS and PKS-NRPS hybrids. In particular, in the α-aminoadipate (AAA) pathway characteristic in fungi,[262, 263] the enzyme Lys2 acts on aminoadipate **222** to reduce it to aldehyde **225**. Lys2 shares many analogies to minimal NRPS systems, bearing adenylation (A), thiolation (T or PCP) and reductive release (R) domains.[264] **222** is activated by adenylation to **223** and loaded onto the PCP (**224**). The chain is then released by the R domain to yield aminoadipate semialdehyde **225** that can be further processed to

lysine (**88**, Scheme 2.13).
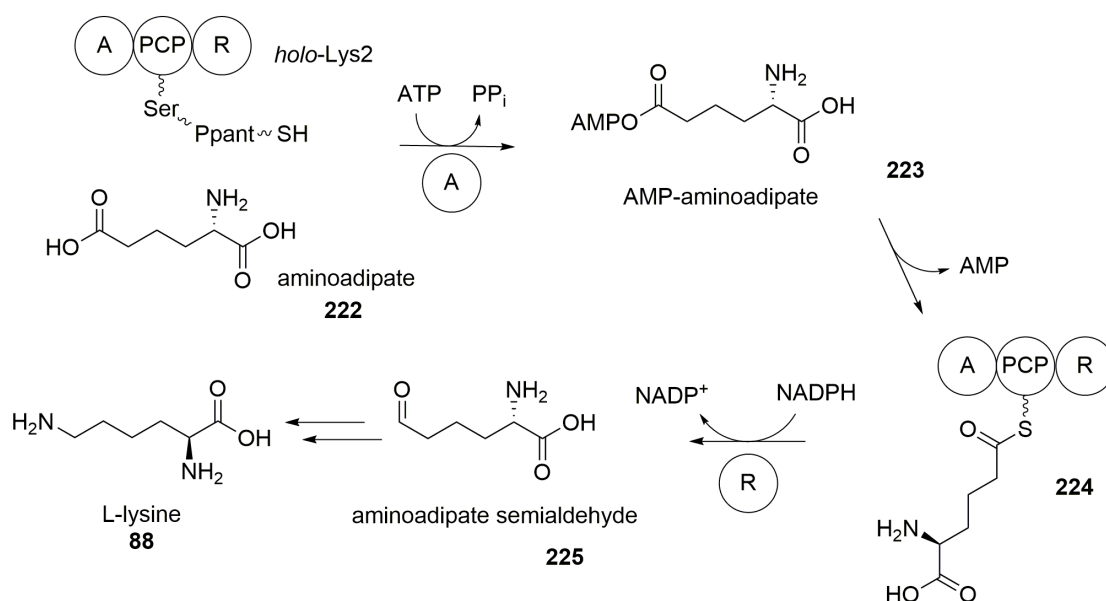
Scheme 2.13: Reduction step in α-aminoadipate pathway in fungi.

The release domain of PhyS was shown to have strong similarities to DKS release domains, such as tenellin PKS-NRPS TenS and pyranonigrin E PynA. Protein alignment showed conserved catalytic sites of the terminal chain-releasing DKC domain within TenS, PksD and PhyS, all having mutation in the NADPH binding site or in the crucial Ser-Tyr-Lys triad of the short chain redox (SDR) superfamily (Figure 2.32, Section 3.3.4).[265] This observation suggested that the mechanism of release and formation of the tetramic acid moiety is analogous between tenellin, pyranonigrin and phyllostictine.

The pyranonigrin E **198** biosynthetic gene cluster *pyn* and metabolic pathway have been characterised in *Aspergillus niger* by Watanabe and collaborators[253] (Scheme 2.14, Table 2.16). We compared *phy* and *pyn* biosynthetic clusters at protein level in order to observe any analogy. The result showed that the two clusters are rather different, being fairly different even for the core PKS-NRPS systems (Figure 2.31, Section 3.3.4). PynA has a broken C-MeT domain, resulting in the complete absence of SAM-derived methyl groups in its product, while PhyS posses a functional C-MeT that introduces a single methylation. The chain length of phyllostictines and pyranonigrins is also different, with PynA processing exclusively six malonyl-CoA units, while PhyS is able to vary the chain length of its products, suggesting a different programming of the latter. Also the tetramic acid forming amino acid differs in the two systems, being alanine in PhyS and serine in PynA. This implies different routes for the formation of the terminal methylene: it is introduced hydrolytically in the case of **198**, with water elimination (**226** to **198**, Scheme 2.14), but oxidatively in the case of phyllostictines. Moreover, the Dieckmann

cyclisation to form the tetramic acid moiety is proposed to be catalysed by a *cis*-acting hydrolase (PynI) in the case of **198** (**227** to **228**, Scheme 2.14) which is not encoded in the *phy* cluster. Watanabe and collaborators proposed the thiolesterase PynI to work in the offload of the chain as the ΔpyrI mutant could not produce any intermediate of the pyranonigrin pathway. Considering the alignment analysis performed between various releasing domain of PKS-NRPS systems, one could discuss that the PKS-NRPS PynA may be able to release the chain independently by Dieckmann cyclisation, and the real role of PynI might be proofreading of the core PynA: if the growing chain gets stuck with no chances of being further processed, it would definitely stall the PKS-NRPS, as enzyme and polyketide chain are covalently bound. PynI may release the shunt product, resetting PynA and enabling its function once again. This would explain why ΔpynI mutant is not able to produce any product: the absence of the proofreading thiolesterase may lead to the accumulation of dead PKS-NRPS unable to process correctly the substrate. Thus it appears that despite their chemical similarities **190** and **198** have probably arisen by parallel evolution.
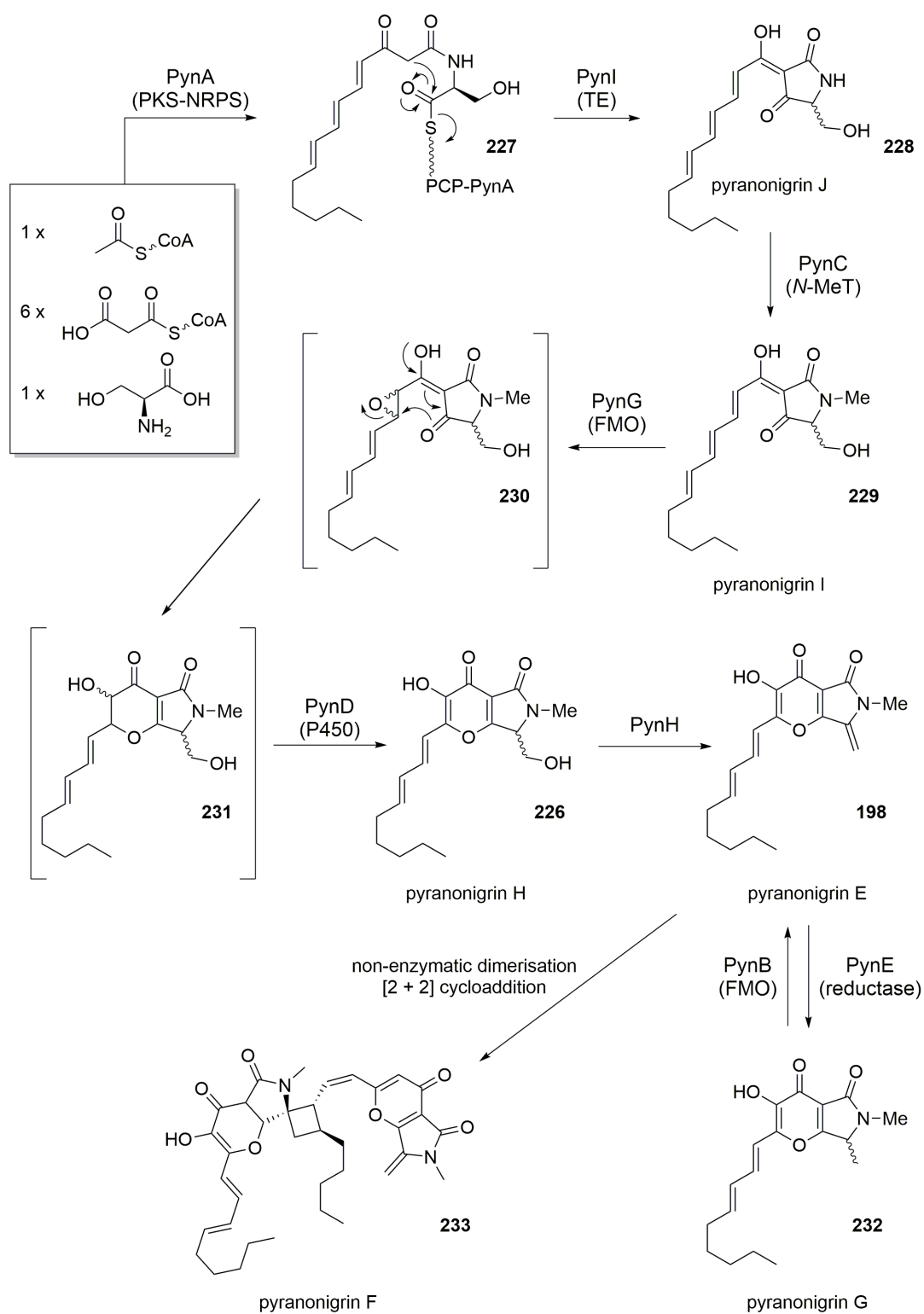
Scheme 2.14: Biosynthesis of pyranonigrin E and related compounds.

2.4.5 Homologous *phy* BGC in other fungi

PhyS showed high homology to a PKS-NRPS found in *M. mycetomatis* (namely PksD). The product of *M. mycetomatis* PksD is currently unknown, but the high identity shared with PhyS suggested that it could also form a tetramic acid-containing compound. Interestingly, the high homology is not limited to the PKS-NRPS enzymes, but to the whole BGC of *P. cirsii* and *M. mycetomatis* (Figure 2.28), suggesting that the human pathogen may produce compounds related to the class of phaeosphaerides. Since phyllostictine A itself, and the analogous phaeosphaerides showed distinct activities against human cell lines it is an intriguing possibility that similar compounds biosynthesised by *M. mycetomatis* may be involved in human pathogenicity. A major difference between the phaeosphaerides and the putative class of compounds produced by *M. mycetomatis* is the level of unsaturation of the polyketide chain. PksD has a non-functional ER domain like PhyS, but no *trans*-acting enoyl reductase such as PhyL4 in *phy* cluster. The same pattern is observable in the fusarin C BGC, which encodes the biosynthesis of a polyunsaturated compound. These are all speculations, as no extensive research has ever been carried out concerning the *M. mycetomatis* molecular arsenal. Analysis of *M. mycetomatis* secondary metabolome should be performed to observe which class of compounds are produced, and to which extent they are related to the phaeosphaerides.

The original source of phaeosphaerides, strain FA39 (related to *Phaeosphaeria avenaria*),[242] has no genome sequence data available. The closest species with an accessible genome is *Stagonospora nodorum* (telomorph *Phaeosphaeria nodorum*). *S. nodorum* is known to make PKS-NRPS compounds,[266] in particular pramanicin **234**, a modified PK-NRP pyrrolidinone antibiotic (Figure 2.44).[267]
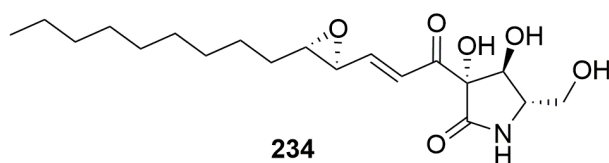
Figure 2.44: Pramanicin produced by *Stagonospora nodorum*.

Labeling studies with universally labeled amino acids and acetate showed L-serine to be incorporated as an intact entity into the pyrrolidinone moiety, and acetate to form the octaketide moiety.[268] The *anti*-1,2-diol on **234** pyrrolidinone ring suggests a reductive release from the PKS-NRPS to give the aldehyde **235** that can undergo Knoevenagel cyclisation to yield **237** through **236**. Two events of epoxidation and hydrolysis of the pyrrolidinone epoxide to form the *anti*-1,2-diol are likely to happen to yield **234** through **238** (Scheme 2.15). Considering L-serine as the amino acid selected by the PKS-NRPS hybrid and the reductive mechanism of release, we consider pramanicin not to be closely related to the phyllostictine system.
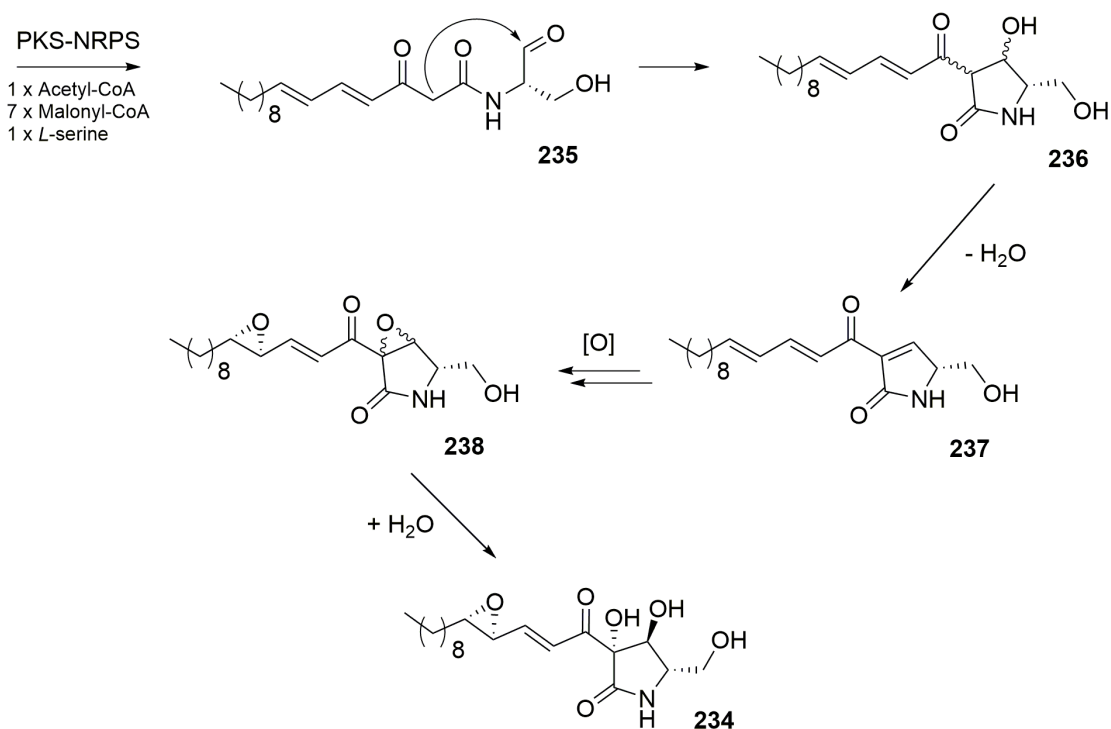
Scheme 2.15: Possible bisynthetic pathway of **234**.

The BGC *phy* was BLASTED into *S. nodorum* genome, in order to search for a possible homologous gene cluster in this organism. Interestingly, the only significantly conserved gene between the two BGC (beside the PKS-NRPS), was found to be SNOG_00312.3, a *trans*-acting enoyl reductase homologous to PhyL4, suggesting that the product of this putative BGC may be a highly reduced polyketide-NRP metabolite. Comparison at the protein level did not lead to any striking homology evidence, suggesting that *P. nodorum* may not produce the class of phaeosphaerides. Since our evidence showed that pyllostictine B is identical to phaeosphaeride A **181**, we propose that the endophyte strain FA39, and possibly *Phaeosphaeria avenaria*, possess a highly homologous, or even identical, *phy* BGC. For this reason, *Phyllosticta cirsii* could be renamed *Phaeosphaeria cirsii* or *Stagonospora cirsii*.

2.4.6 Biosynthesis proposal

Based on the labeling studies and the analysis of the *phy* biosynthetic gene cluster we could propose a likely biosynthetic pathway for the synthesis of the phyllostictines and the closely related metabolites (Scheme 2.16). In the first step a hybrid PKS-NRPS, aided by the *trans*-acting ER PhyL4, produces a monomethylated β-keto unsaturated hexaketide or pentaketide from acetate and methionine and links this to alanine to form

the enzyme-bound **239**. The Dieckmann release domain (DKC) of PhyS then catalyses cyclisation and release to form a 3-methyl tetramic acid **240**. Oxidation at the 17-methyl and elimination then forms the 3-methylene tetramic acid **241** and is possibly catalysed by the PhyL8 encoded non-heme iron-dependent oxygenase. Formation of the dihydropyran **243** could involve epoxidation (possibly PhyL9) and then ring-closure *via* **242.a**. Alternatively, Michael addition followed by later hydroxylation at C-7 would lead to **243** *via* **242.b**. The carbonyl at C-6 is reduced to give the known compounds **199** and **184**, possibly by the NAD(P)H-dependent long-chain ketoreductase PhyL5. Our knockout results show that the last steps of biosynthesis are the *N*-hydroxylation and final *O*-methylation catalysed by the P450 PhyL6 and SAM-dependent methyl transferase PhyL7 to form phyllostictine A and B. Proteins involved in pyridone *N*-oxidation, for example in the cases of tenellin **197** and desmethylbassianin **245** (Figure 2.45), have been previously reported, but this is the first report of a protein involved in *N*-hydroxylation of a tetramic acid (Scheme 2.16).
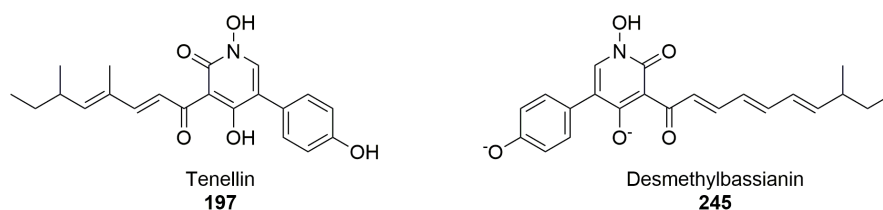
Figure 2.45: Tenellin and desmethylbassianin.

Phaeosphaeride A was reported to have potent inhibitory effect against the STAT3 (Signal Transducer and Activator of Transcription) regulator, a nuclear receptor involved in the promotion of cell growth and survival, constitutively activated in many cancers, including breast, prostate, ovarian, and skin tumors.[269–272] Since this class of compounds possesses interesting and useful bioactivities the discovery and manipulation of the biosynthetic gene cluster opens up the possibility of metabolic engineering to produce new and related compounds in either the producing organism itself or in related fungi. Our results presented here show that new family members such as phyllostictine E can be generated in this way. The revised structure should also allow more effective synthetic strategies to be devised for the construction of this class of compounds.
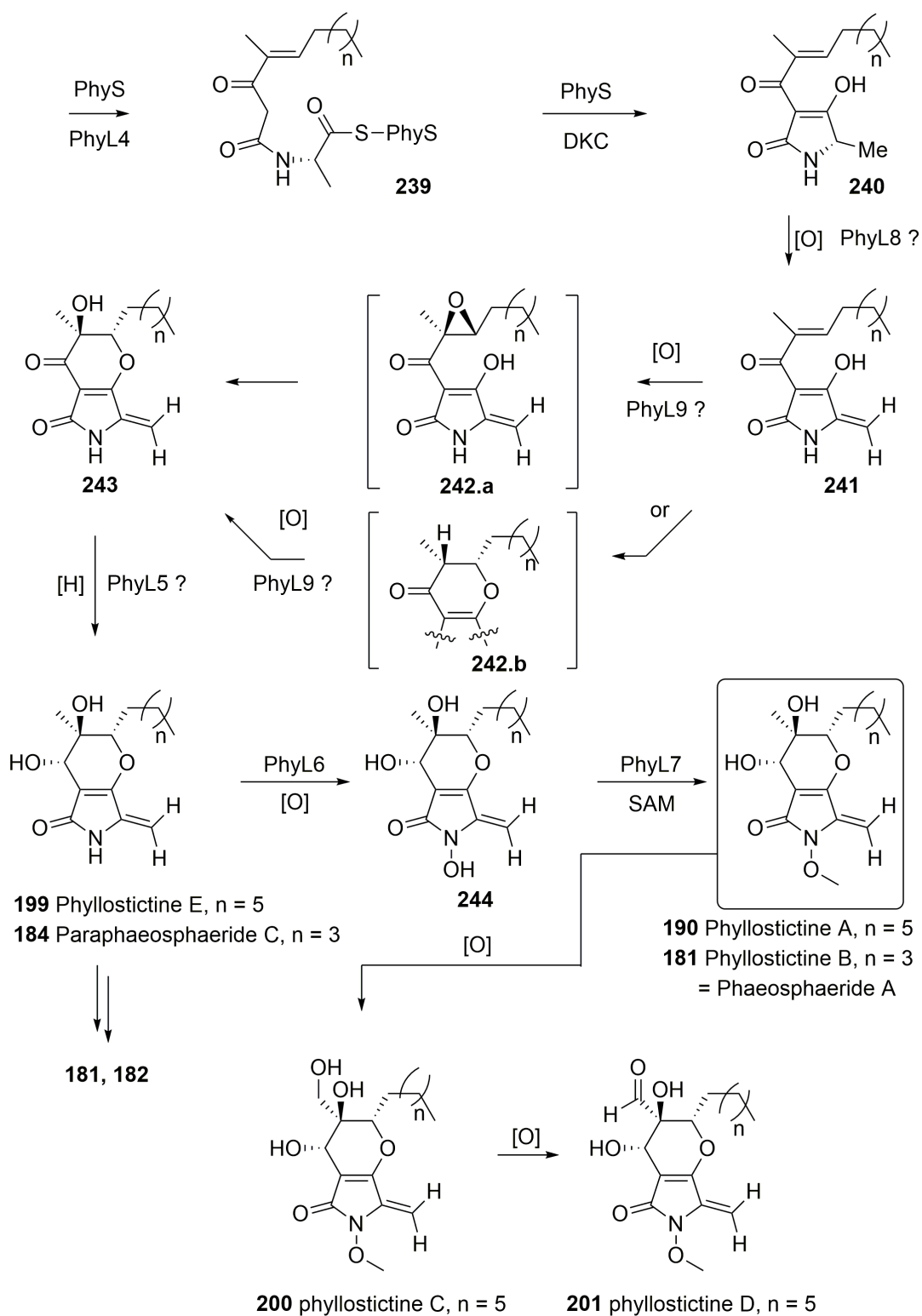
Scheme 2.16: Proposed bisynthetic pathway of **190** and related metabolites.

3 Insights into the biosynthesis of Sch-642305, a potent cytotoxic and antiviral heptaketide from *Phomopsis CMU-LMA*

3.1 Introduction

10-membered lactone polyketides are commonly found in plants, fungi, and marine and terrestrial bacteria (Figure 3.1). Jasmine ketolactone **246** was the first decalactone isolated in 1942 from essential oil of *Jasminum grandiflorum*.[273] This family of compounds presents a wide array of biological activities: diaplodialide A **247** inhibits the 11-α-hydroxylase involved in steroid biosynthesis; microcarpalide **248** interferes with microtubule assembly;[274] modiolide **249** is a fungal antimicrobial and fungicide;[275] muggelone **250** blocks the maturation of *Danio rerio* revealing the molecular mechanisms of its embryonal development.[276] Decalactones isolated from filamentous fungi are commonly referred as decanolides. This group includes: stagnolide I **251**; herbarumin III **252**; nonenolide **253**; decarestrictine C1 **254**; and phomolide A **255**.[277–281]
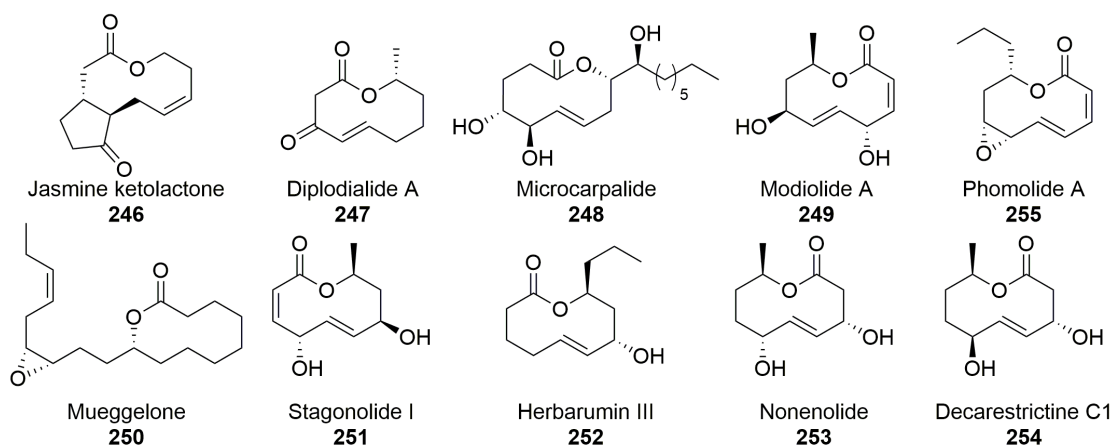
Figure 3.1: Decalatones natural products.

Sch-642305 **169** was originally isolated from *Penicillium verrucosum* ILF-16214 in 2003

by the pharmaceutical company Schering-Plough (Figure 3.2).[282] The filamentous fungus *Phomopsis CMU-LMA* was also found to be a consistent source of **169** by Ouazzani and coworkers in 2011.[192] The bicyclic 10-membered macrolide structure was determined by NMR spectroscopy and the absolute stereochemistry was assigned by X-ray crystallographic analysis of the *p*-bromobenzoate ester.[282]
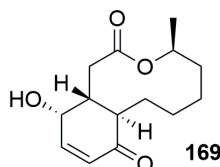
Figure 3.2: Sch-642305 isolated from filamentous fungi *Penicillium verrucosum* and *Phomopsis CMU-LMA*.

Efforts to accomplish **169** total synthesis have been made because of its potent inhibitory activity against Gram-positive DNA primase ($EC_{50} = 70\ \mu M$[282]) and its anti-viral action against HIV-1 Tat protein ($IC_{50} = 1\ \mu M$[283]), proposing its usage as an antibiotic against multi drug resistance *Staphylococcus aureus* and an active substance in AIDS therapy.

Since its discovery, **169** captured the attention of various synthetic chemists, resulting in eight major publications between 2005 and 2009.[284] Total synthesis of **169** has been successfully achieved by Mentha and Shinde using a ring closing metathesis.[285, 286] The most elegant strategy was described by Wilson and Trauner, with only 7 steps and a 12% yield, which includes Mukaiyama-Michael addition followed by allylation and, again, ring-closing metathesis to form the decalactone.[284] The only total synthesis that attempts to mimic the natural biosynthesis has been proposed by Snider and Zhou: they reported a 17-step strategy *via* the transanular Michael reaction of the hypothetical mutolide-derived ketolactone **256** to yield **169** (Scheme 3.1). In particular, treatment of the ketolactone **256** with 1.2 equivalents of NaH in THF (0 °C, 30 minutes) afforded the transannular Michael adduct as a single isomer **257** in 80% yield. Microwave irradiation of **257** in 1.5% TFA in CDCl$_3$ for 3 hours at 120 °C, gave a separable mixture of **257-258** in 1:3 ratio. Finally, hydrolysis of the TBDPS ether of **258** with TBAF/HOAc in THF afforded (+)-Sch-642305 **169** in 86% yield.[193]
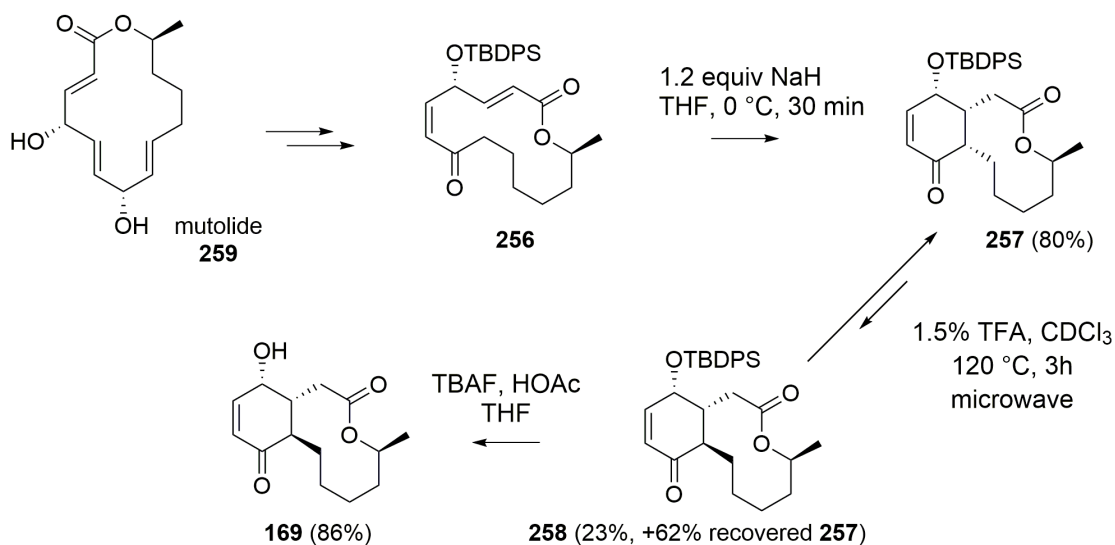
Scheme 3.1: Synthesis of **169** by a biomimetic transannular Michael reaction by Snider and Zhou.

Although Snider could reproduce the chemical reactions from mutolide **259** to **169**, neither the hypothetical intermediate **256** nor **259** have ever been purified from any **169**-producing strain. On the tracks of **169** biosynthesis, Ouazzani proposed a different mechanism (Scheme 3.2), involving benquoine **260** as the real precursor, as it was isolated from the **169**-producer organism *Phomopsis CMU-LMA*.[192] Two hypothetical biosynthetic pathways stemming from benquoine were suggested by Ouazzani: the **A** route, which forecasts carbonyl protonation followed by intramolecular rearrangements, while the **B** path involves a succession of oxido-reductions and enzyme-catalyzed Michael cyclisation (**265** to **169**, Scheme 3.2).[192,194] Route **A** appears unlikely because of the C-4/C-5 olefin switch geometry: *E* becoming *Z* with no apparent reason (**261** to **262**). Moreover the cationic chemistry, more typical of terpene biosynthesis, is uncommon in fungal polyketide biosynthesis. However, route **B** appears more typical of fungal pathways.
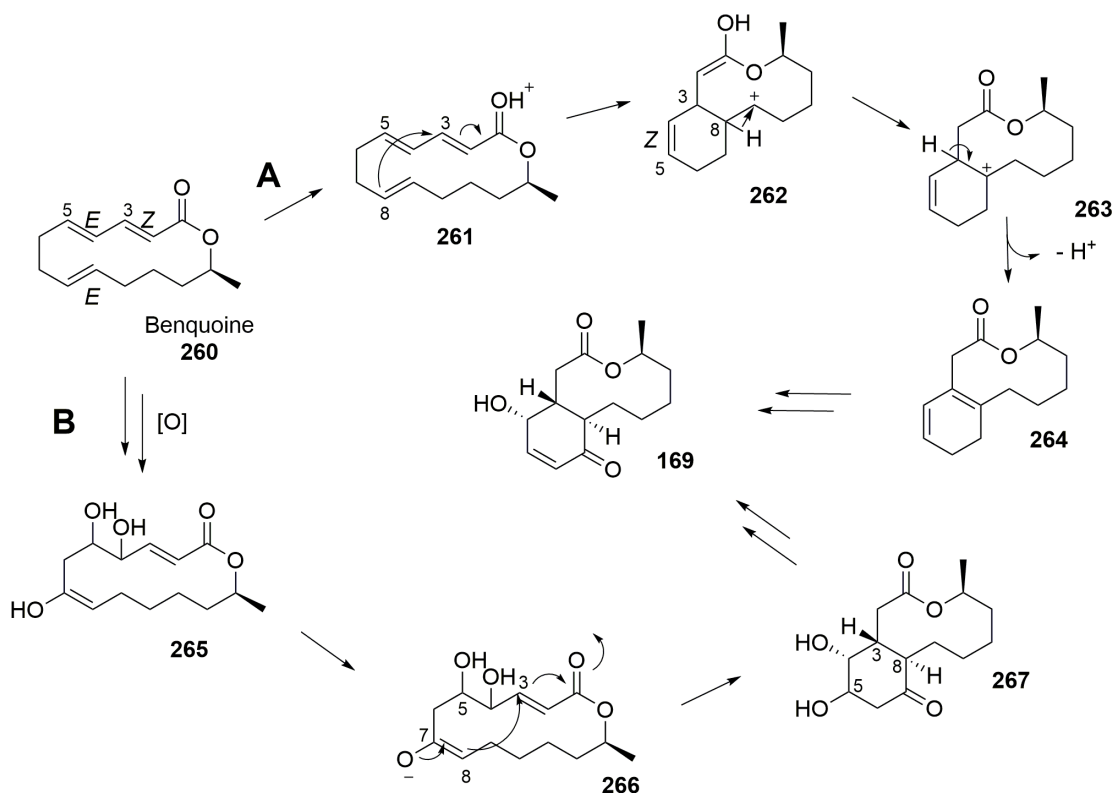
Scheme 3.2: Mechanisms proposed by Ouazzani for **169** biosynthesis stemming from benquoine.[192]

Phomopsis CMU-LMA is an endophytic ascomycete of the *Diaporthaceae* family. It was isolated from seemingly healthy *Alpinia malaccensis* leaves in the forest of the National Park Doi Suthep-Doi Pui in northern Thailand.[287–289] *Phomopsis CMU-LMA* produces the macrolide **169**, along with a variety of compounds, mostly polyketides, whose structures have been characterized by Ouazzani *et al* (Figure 3.3).[192] *Phomopsis CMU-LMA* is a strong source of secondary metabolites and it was recently found that it produces 50 mg/L of **169**, 20 times the amount compared to the original source *P. verrucosum*.[192]
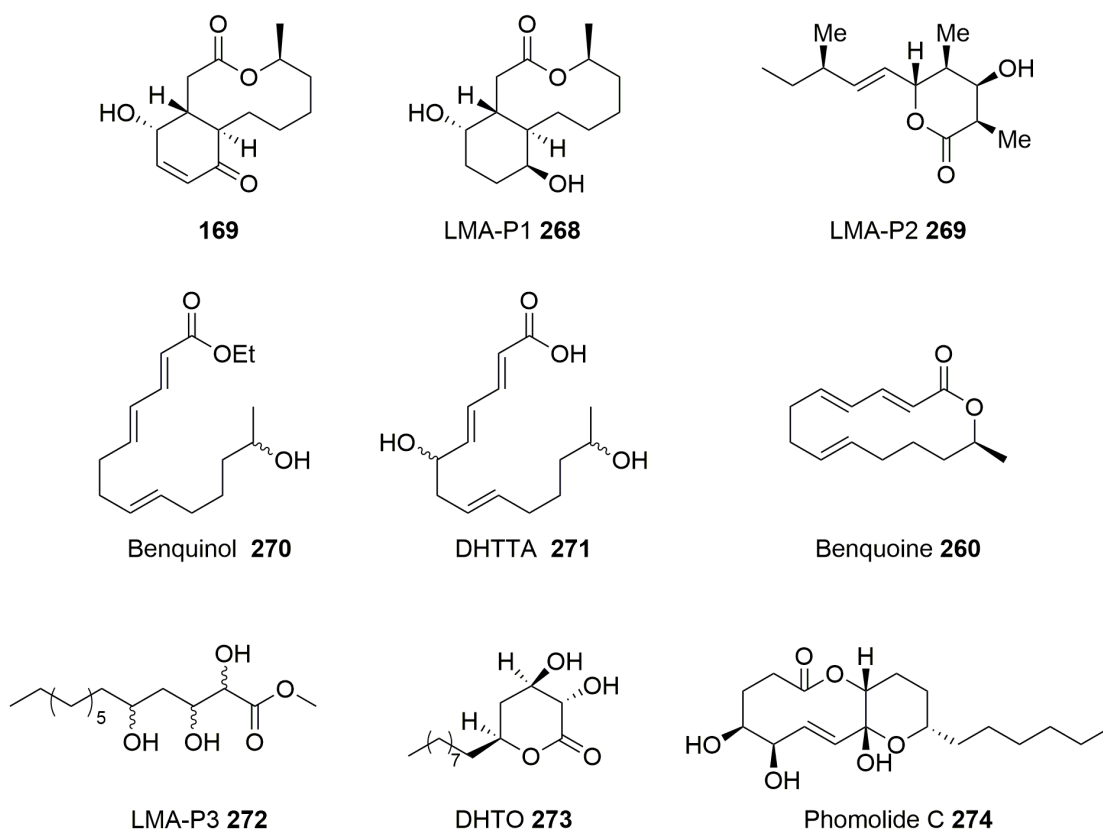
Figure 3.3: Isolated compounds from source *Phomopsis CMU-LMA*.[192]

A metabolite related to **169** is Brefeldin A **275**, a protein-transport inhibitor isolated from several species of filamentous fungi (Figure 3.4).[290–294] **275** is used to study protein transport among eukaryotes as it prevents protein transport from the endoplasmic reticulum to the Golgi apparatus by interfering the association of COP-I proteins,[295] but has also been found to have antiviral, antifungal, and antitumor properties.[296]
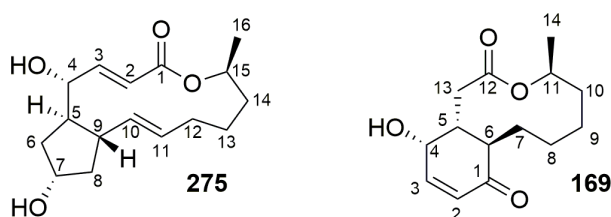
Figure 3.4: Brefeldin A and Sch-642305

The polyketide origin of this 16- membered macrolactone was previously established through feeding studies with [$^{18}O_2$, 2-$^3H_3$]-acetate and $^{18}O_2$[297–299] (Scheme 3.3 **A**). Feed-

ing the **275**-derivative [4-^{2}H]-brefeldin C **276** to the **275**-producer *Eupenicillium brefeldianum* resulted in efficient conversion of [4-^{2}H]-**276** to [4-^{2}H]-**275**, determining the C-7 hydroxylation to be a late stage event that does not contribute to the cyclopentane ring closure (Scheme 3.3 **B**).[300] The formation of the cyclopentane ring was proposed to be P450-mediated,[300] but it is still an open question.
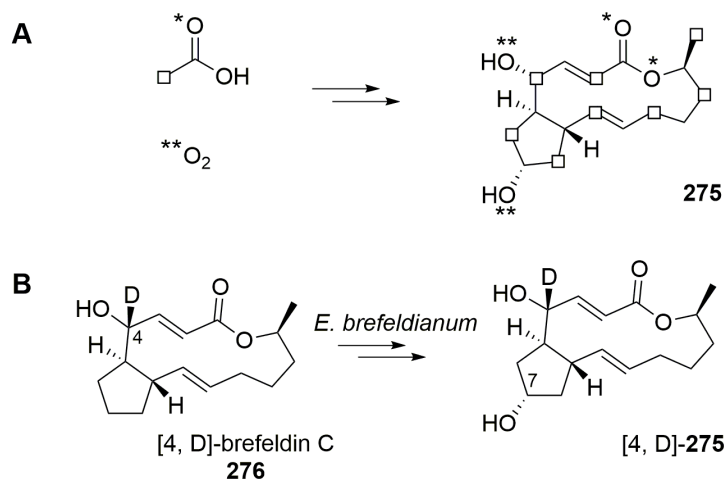
Scheme 3.3: Incorporation of precursors into **275**. **A** incorporation of labeled [^{18}O$_2$, 2-^{3}H$_3$]-acetate and ^{18}O$_2$. **B** Conversion of [4-^{2}H]-brefeldin C to [4, D]-brefeldin A by *E. brefeldianum*.

Brefeldin A **275** is the octaketide product of an hrPKS (namely Bref-PKS) that works with a partner thiohydrolase (Bref-TH). It is a common feature of hrPKS to lack the thiolesterase domain (TE), relying instead on *trans*-interaction with discrete thiohydrolases or acyltransferase-like enzymes for product release.[61] Such partnering enzymes play an important role in terminating the chain elongation and consequently in determining the length of the final product.

Tang and collaborators demonstrated that Bref-PKS is a nonaketide synthase in the absence of Bref-TH.[301] *In vitro* and yeast heterologous expression evidences showed that Bref-TH and Bref-PKS work together yielding the four linear octaketides **277** - **280**, with **277** being the major product. *In vitro* experiments found Bref-PKS alone to be a nonaketide synthase with major production of **281** and trace amounts of **282** (Scheme 3.4). Importantly, the polyketide starter-unit C-1 position remains oxygenated (Scheme 3.3) and the same moiety is preserved during **169** biosynthesis.

The reconstitution experiments showed that the dominant product of Bref-PKS in the presence of Bref-TH is the acyclic polyketide **277**, which has the same length as **275** and β-reduction pattern for downstream conversion. It was found that Bref-TH has an important role in controlling the release of the octaketide product from Bref-PKS before an extra cycle of chain extension. It is also interesting to notice that the minor compounds

279, **280** and **282** have skipped one β-reduction at the starting unit, retaining a ketone instead of an alcohol. This showed the KR domain to be prone to omit ketoreduction of the PKS-bound diketide intermediate. Further investigation showed the NADPH concentration to determine the KR efficiency at the first ketoreduction step at least *in vitro*: NADPH < 2 mM (still physiological concentration) led compound **279** to be the major product, while NADPH > 2 mM gave **277** as dominant product. Hence, the KR domain is highly sensitive to availability of reducing cofactors, adding one more degree of complexity in generating scaffold variability.[301]
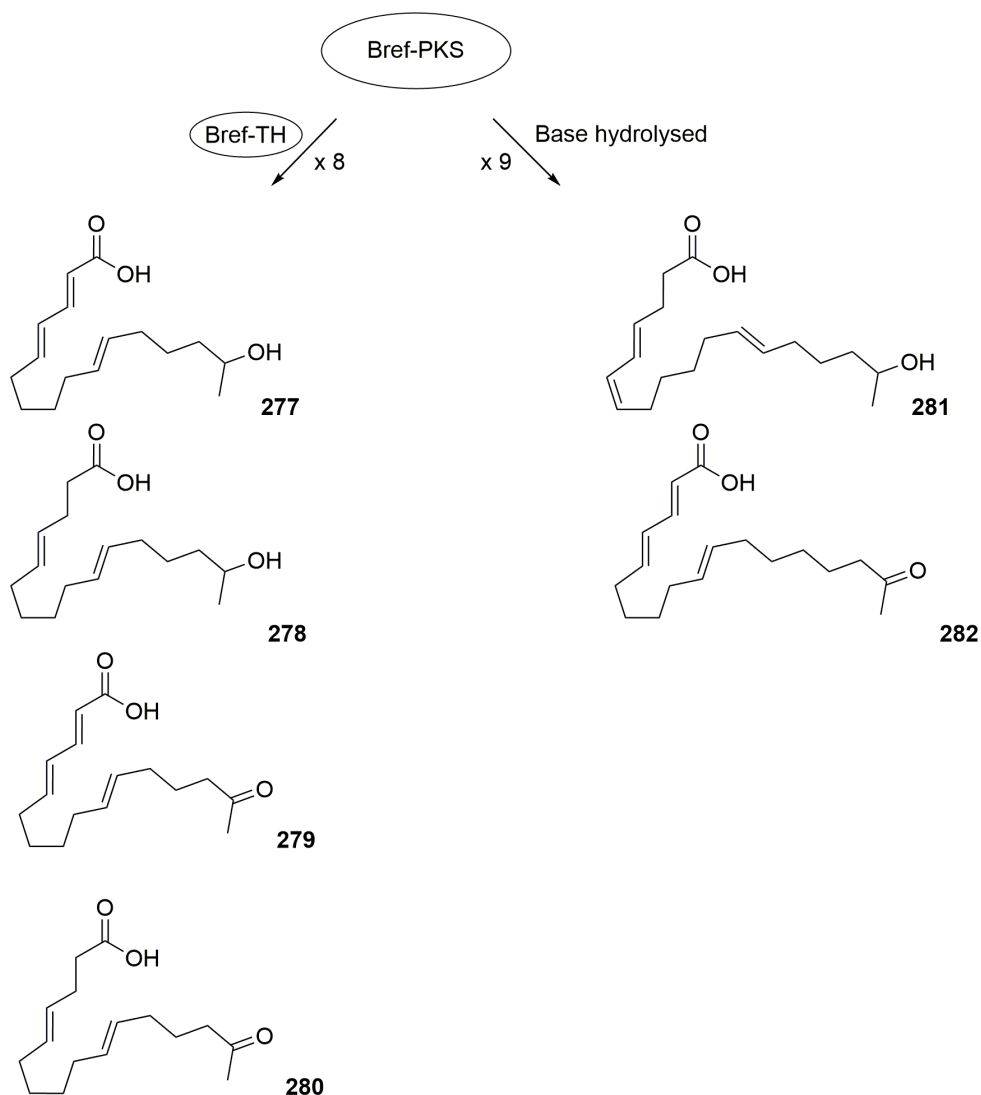
Scheme 3.4: Chain length control Bref-PKS product by the *trans*-acting thiolesterase Bref-TH *in vitro*.

The macrolactonisation mechanism of **275** is still a mystery. In non-reducing PKS (e.g. RAL biosynthesis) often the chain is released by lactonisation by the action of a *cis*-thiolesterase. Tang *et al* proposed that the Bref-TH could act similarly to a *cis*-TE performing both offloading and lactonisation, with the catalytic *His276* of Bref-TH that could deprotonate the OH-15 and make it a suitable nucleophile for lactonization. Clearly, the OH-15 must be in place to obtain the cyclisation, as observed for compounds **277**, **278** and **281**. However, the observation that all of the products obtained *in vitro* and in yeast were linear indicates the necessity of having the cyclopentane ring form first, in order to bring the nucleophile closer to direct lactonisation.

The mechanism behind the formation of the cyclopentane ring of **275** is still not elucidated. Yamamoto *et al* proved the OH-7 to be a non-necessary decoration for the 5-membered ring formation, and Mabuni *et al* proposed a mechanism involving a C-4/C-5 epoxide intemediate (**283** to **284**, Scheme 3.5).[298] However, this epoxide-opening mechanism would require an alkene at C-9/C-10, which needs post elongation tailoring to be put in place, as it is an unusual position in a polyketide chain (Scheme 3.5).
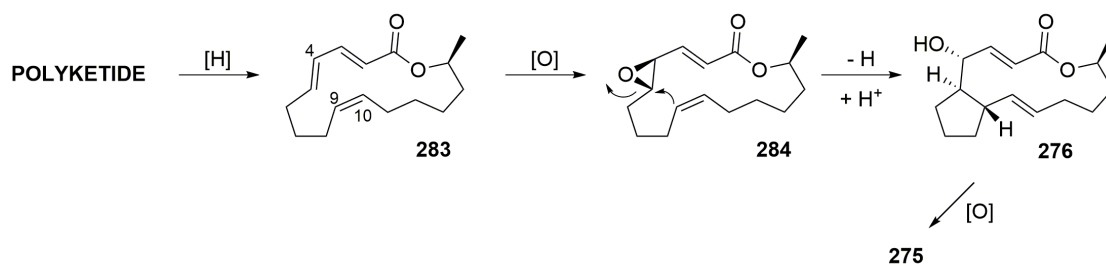
Scheme 3.5: Proposed mechanism of the 5-membered ring formation by Mabuni *et al via* nucleophilic attack of a C-4/C-5 epoxide intermediate.[298]

Tang *et al* observed intermediate **277** as an early product of Bref-PKS, thus they proposed a mechanism *via* C-9 radical intermediate catalyzed by a P450 enzyme (**277** to **285** *via* **286** and **287**, Scheme 3.6). Interestingly Tang suggested that the P450-mediated ring formation may also happen at a PKS-bound intermediate stage **288**, although there are no known examples of tailoring events simultaneous to chain elongation by iterative fungal PKS.
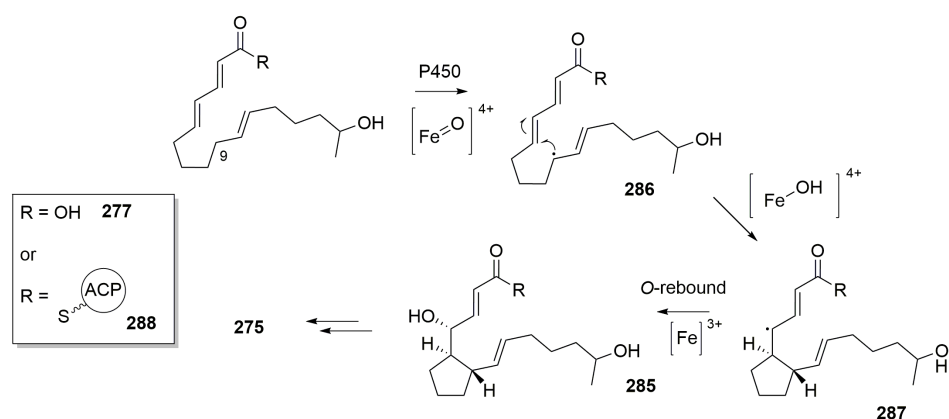
Scheme 3.6: Proposed mechanism of the 5-membered ring by Tang *et al*.[301]

3.2 Aims

Although **169** has a rather simple structure, its biosynthetic steps have not been elucidated until today, therefore we resolve to study the biological mechanisms that happen *in vivo* during its biosynthesis, with particular attention to the chemistry behind the 6-member ring formation. We propose that **169** is the product of a highly-reducing PKS decorated by post-elongation oxidative events.

The organism of choice for the study is *Phomopsis CMU-LMA*, as it grows fast, produces **169** reliably and its genome is sequenced (although at the start of this project the data had not been assembled). The high number of polyketides produced by this filamentous fungus is noteworthy. For example, the linear **272** possesses a polyketide chain with an odd number of carbons, suggesting the use of an unusual starting unit. Compound LMA-P1 **268** and **169** are clearly products of the same pathway, one being the precursor of the other; also benquoine **260** displays a 14-member macrolactone similar to **268** and **169**, suggesting that it may be a early precursor, such as the product of the core hr-PKS prior to downstream processing. The hr-PKS could release the chain through macrolactonisation, or by partnering with a *trans* acyltransferase-like protein or thiohydrolase analogously to the macrolactone brefeldin A **275**.[301] Benquinol **270** correspond for oxidation pattern to **260**, and it could be the real product released by the PKS in analogy to **277**, or a degradation product of **260**. This would explain the ethyl ester, which is probably due to **260** hydrolysis followed by solvent acylation during the extraction procedure, but it is not to exclude that the PKS releases a linear chain that is cyclised enzymatically, with some shunt ethylation due to endogenous acyl-transferases. DHTTA **271** is also an open form of a downstream oxidised **260**. Indeed, the OH-6 cannot arise from a ketoreductive action of the PKS. If this is true, it may be that a hypothetical 6-hydroxy-benquoine would be part of **169** biosynthesis.

No genomic data of *Phomopsis CMU-LMA* are available in the literature, and no biosynthetic cluster has been reported in any other **169** producer strain, therefore the composition of the **169** BGC is yet to be determined. The genome of *Phomopsis CMU-LMA* was sequenced by Ouazzani in 2005 at the *Institut de Chimie des Substances Naturelles* (ICSN, Gif-sur-Yvette, France), but the assembly has never been performed (unpublished data). Raw data can be assembled and annotated by antiSMASH, in order to select the best BGC candidates. Clusters of choice should include genes encoding a highly reducing PKS with an inactive *C*-MeT domain, as **169** scaffold does not contain any SAM-derived methyl-groups, and iron-dependent oxidative enzymes or flavoproteins. Directed disruption of the core and/or tailoring enzymes encoding genes will confirm the role of the putative **169** BGC. In particular, disruption of the core hrPKS will prevent the synthesis of **169** and related metabolites, and KO of tailoring enzymes will lead to the accumulation of intermediates to reveal the biochemical steps of **169** biosynthesis.

3.3 Results

The fungus *Phomopsis CMU-LMA* was provided by Dr. Jamal Ouazzani, who recently isolated and characterised various compounds from this strain.[192,302,303]

3.3.1 *Phomopsis CMU-LMA* culturing and Sch-642305 harvesting

In our hands *Phomopsis CMU-LMA* grew rapidly on many generic fungal media such as Czapek Dox Agar (CDA) and in Potato Dextrose Broth (PDB) developing with different phenotypes: colonies varied their colour from a pale greenish brown to a deep black. *Phomopsis CMU-LMA* grew as highly branched mycelia and formed spores after 10 days of growth on PD Agar medium at 28 °C, that can be stored for unlimited time in 25% glycerol at -80 °C.
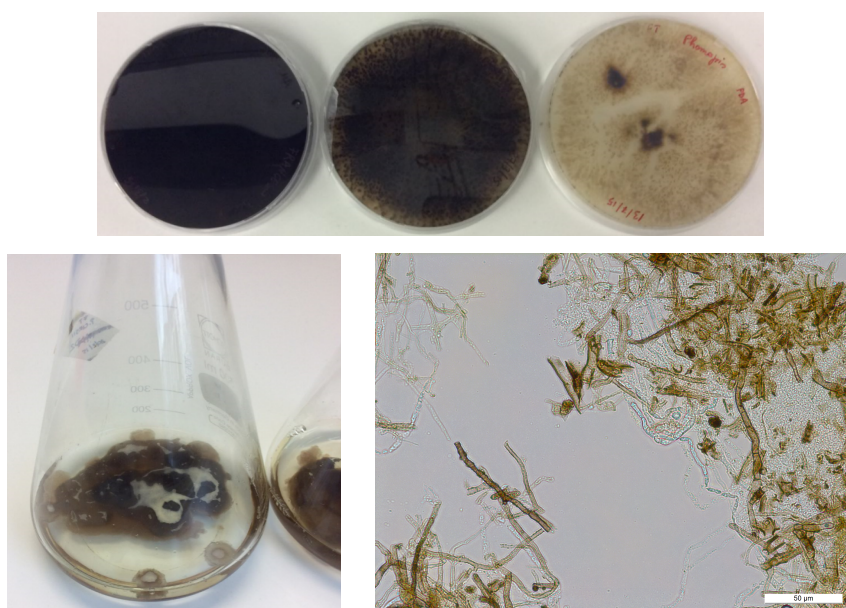
Figure 3.5: *Phomopsis CMU-LMA* in solid and liquid cultures, and under optic microscope.

A crude harvesting routine involved blending mycelia together with supernatant and then filtering the mycelia. Double extraction with ethyl acetate was performed in a *1 : 1.5 = supernatant : solvent* ratio. The solvent was dried using $MgSO_4$ and then removed by rotavapor. The crude extract was stored at -20 °C or dissolved in acetonitrile at a concentration of 10 mg/ml and analysed by LCMS. A peak with mass corresponding to **169** was observed in extracts obtained from fermentation in any liquid or solid medium after 5 days of incubation at 28 °C, accumulating in the media. Other peaks with masses corresponding to LMA-P3 **272**, phomolide C **274** and DHTO **273** were found (Figures 3.6 and 3.8). Extraction from mycelia showed trace amounts of **169** with a signal intensity 4 fold weaker compared to supernatant extracts (Figure 3.7).
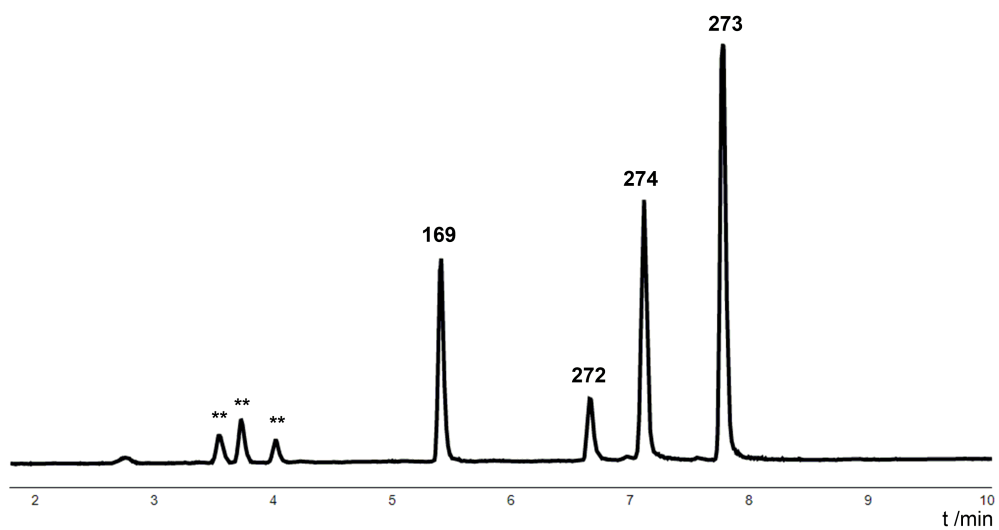
Figure 3.6: crude ELSD chromatogram of WT supernatant (PDB, 5 days, 28 °C, 110 rpm). ** marks unrelated peaks.
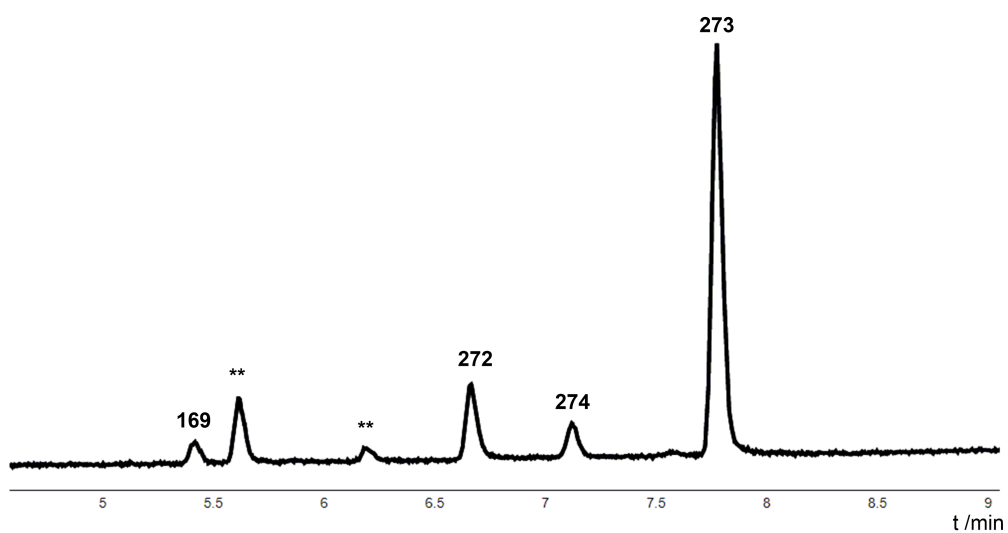
Figure 3.7: crude ELSD chromatogram of WT mycelia (PDB, 5 days, 28 °C, 110 rpm). The scale of *y*-axis is identical to Figure 3.6. ** marks unrelated peaks.
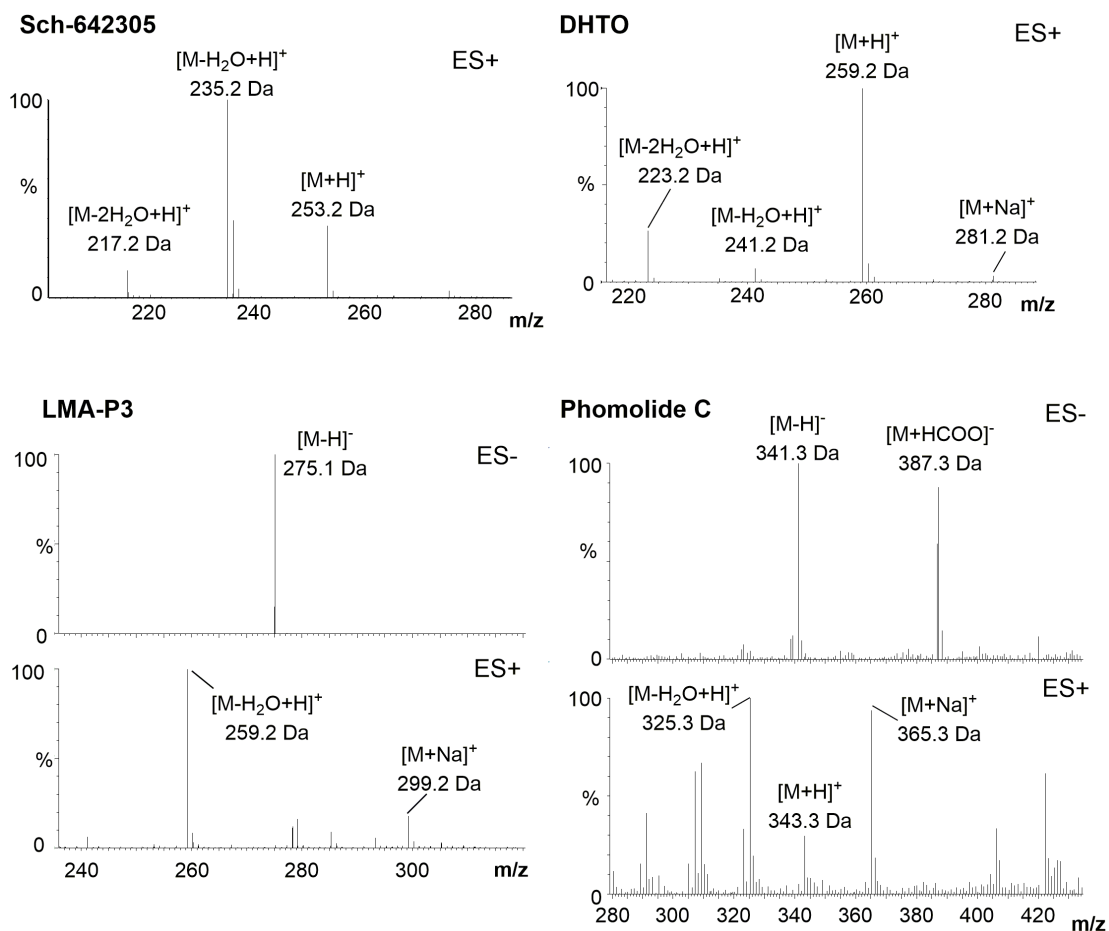
Figure 3.8: Mass spectra (ES+ and ES-) of peaks with mass corresponding to Sch-642305 **169**, LMA-P3 **272**, DHTO **273** and phomolide C **274**.

Interestingly, after two years of continuous culturing *Phomopsis CMU-LMA* started producing different aromatic polyketides: dothiorelone A **289** and cytosporones B **290** and C **291** (Figure 3.9 and 3.10). The identification of those compounds was performed by comparison of NMR spectra found in the literature (data not shown). In particular, Reaxys[241] was a valuable source of literature records: natural product searches gave over 300 metabolites isolated from *Phomopsis spp.*, and we could focus on the ones with masses corresponding to **289** (338.2 Da), **290** (322.2 Da) and **291** (294.1 Da). Production and titres of compounds **289**, **290** and **291** was unpredictable also maintaining the same condition of fermentation, varying consistently without an apparent reason. The same was observed concerning phomolide C **274**.

114
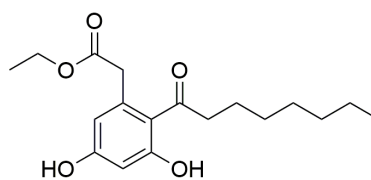
Figure 3.9: ELSD chromatogram of a crude extract obtained from WT *Phomopsis* (PDB, 7 days, 28 °C, 110 rpm). ** marks unrelated peaks.
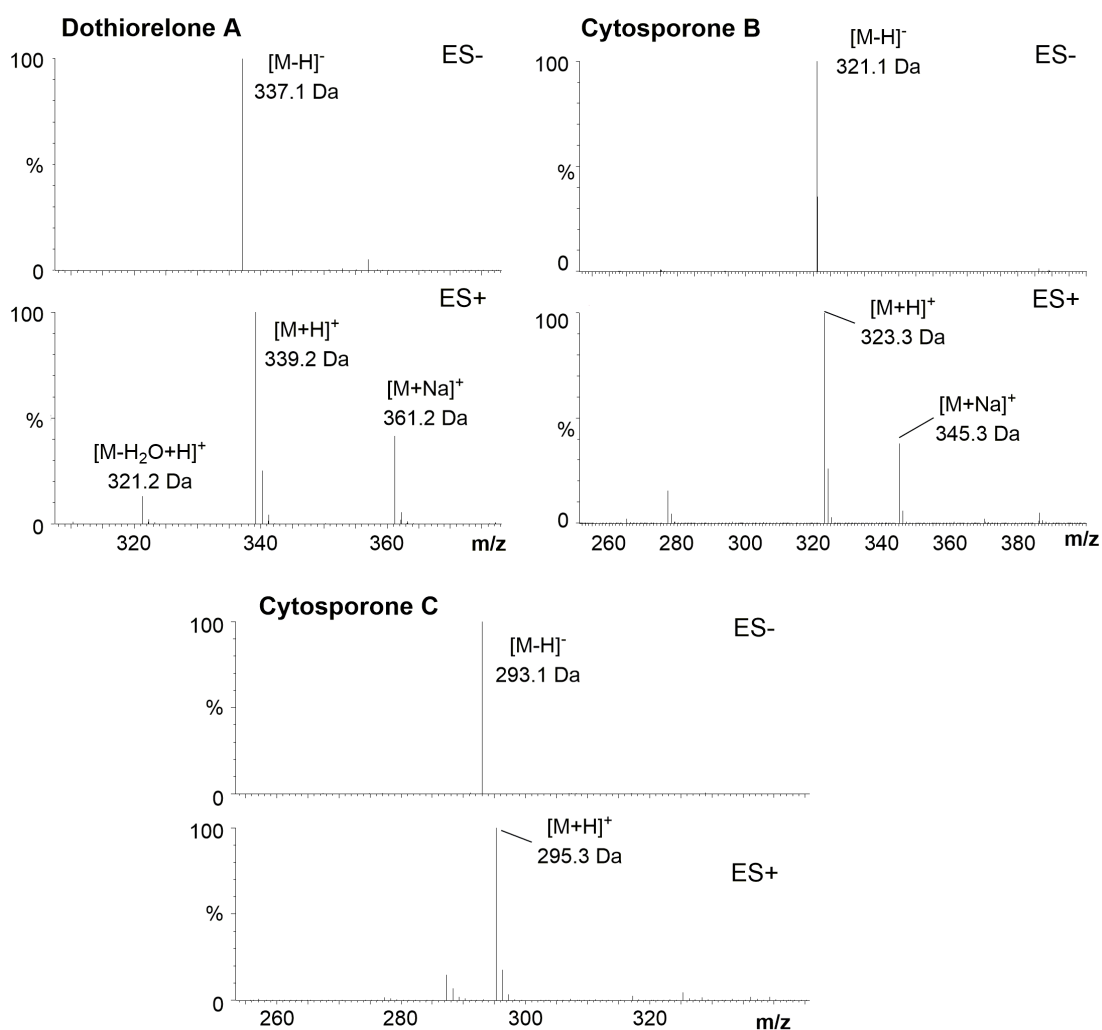
Figure 3.10: Mass spectra (ES+ and ES-) of peaks with mass corresponding to dothiorelone A **289**, cytosporone B and C **290**, **291**.

3.3.2 Sch-642305 structural data

NMR data of pure **169** were compared to literature values, and this confirmed the structure of Sch-642305 (Table 3.1, Figures 3.11 and 3.12.)

CDCl$_3$, 293 K

| | Measured data | | Literature data[303] | |
| | δ_H/ppm | δ_C/ppm | δ_H/ppm | δ_C/ppm |
Position	125 MHz	500 MHz	125 MHz	500 MHz
1	-	200.3	-	200.3
2	6.02 d (1H, 9.9 Hz)	131.1	6.02 d (1H, 9.9 Hz)	131.1
3	6.99 dd (1H, 4.3 Hz, 5.5 Hz)	146.6	6.99 dd (1H, 4.3 Hz, 5.5 Hz)	146.6
4	4.25 dd (1H, 3.4 Hz, 5.6 Hz)	67.1	4.25 dd (1H, 3.4 Hz, 5.6 Hz)	67.1
5	2.80 m (1H)	37.1	2.80 m (1H)	37.1
6	2.66 m (1H)	46.3	2.66 m (1H)	46.3
7	2.11 m (2H)	22.9	2.11 m (2H)	22.9
8	1.29, 1.58 m (2H)	22.3	1.29, 1.58 m (2H)	22.3
9	1.13, 1.78 m (2H)	22.9	1.13, 1.78 m (2H)	22.9
10	1.29, 2.01 m (2H)	30.3	1.29, 2.01 m (2H)	30.3
11	5.06 m (1H)	73.9	5.06 m (1H)	73.9
12	-	172.2	-	172.2
13	2.62 dd (2H, 2.4, 16.8 Hz)	38.8	2.62 dd (2H, 2.4, 16.8 Hz)	38.9
14	1.27 d (3H, 6.6 Hz)	18.6	1.27 d (3H, 6.6 Hz)	18.6

Table 3.1: NMR chemical shifts of compound **169** in CDCl$_3$.
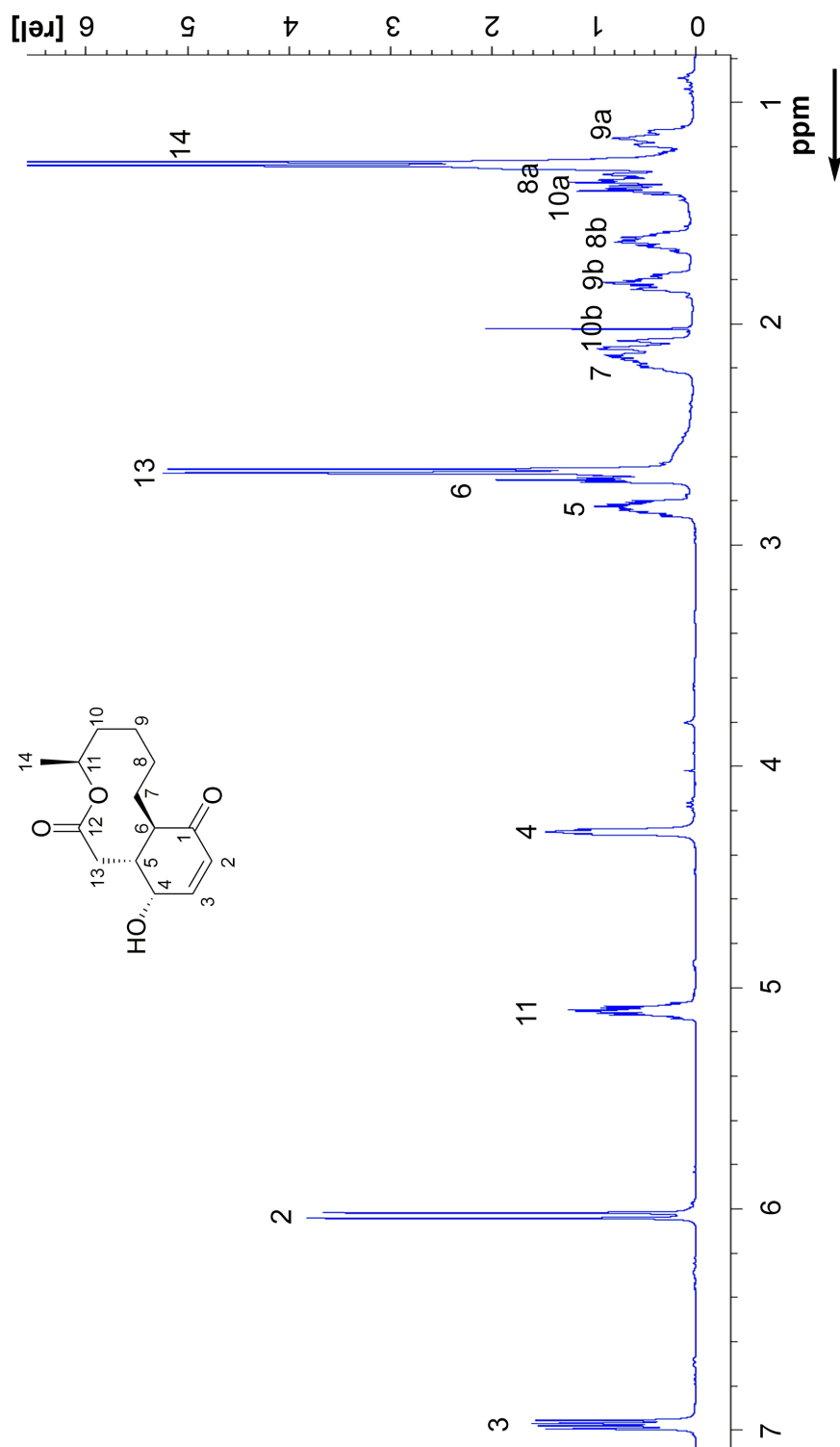
Figure 3.11: ^{1}H-NMR of **169** in CDCl$_3$.
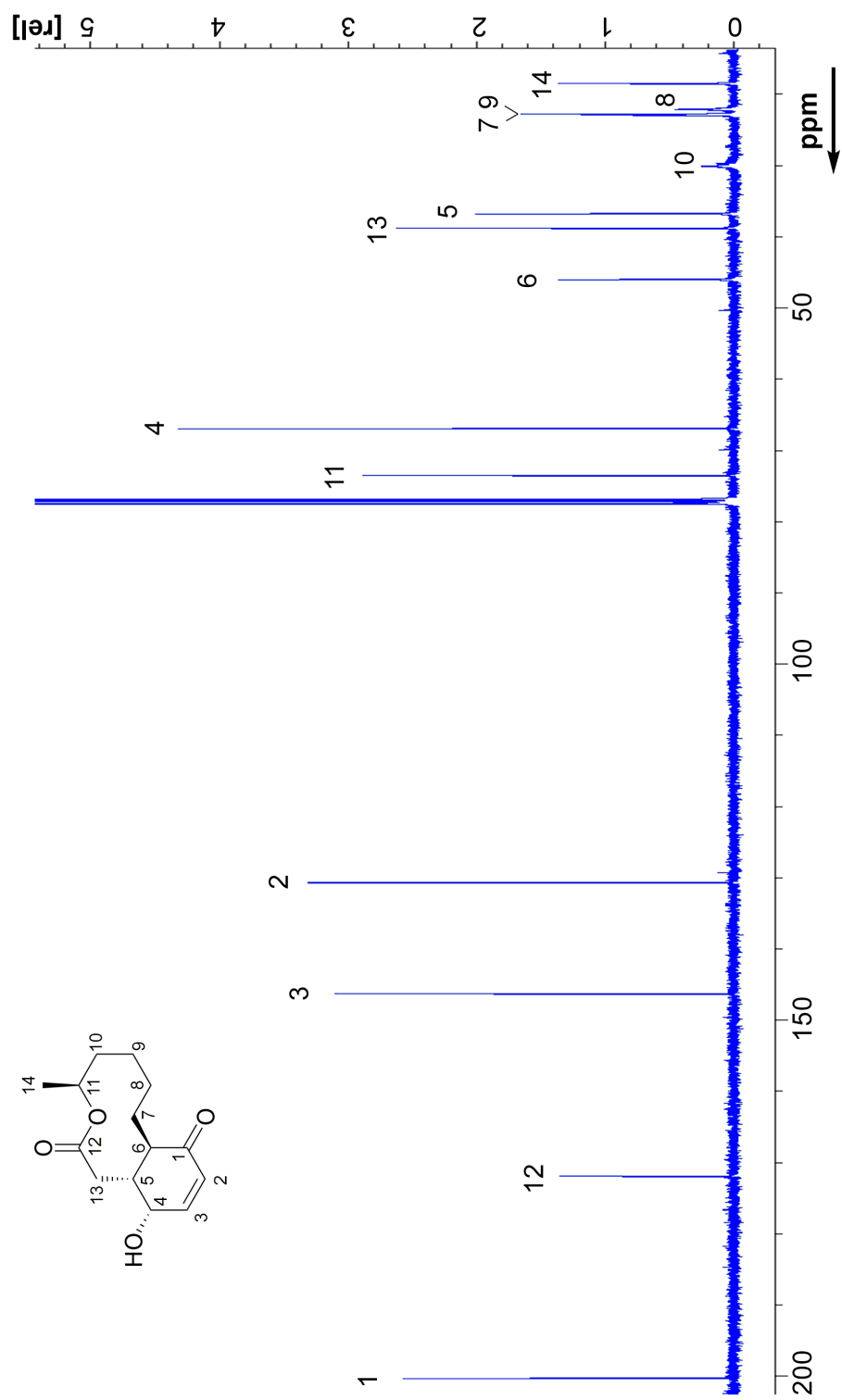
Figure 3.12: ^{13}C-NMR of **169** in CDCl$_3$.

3.3.3 [1-^{13}C]-acetate feeding

Cultures growing in producing conditions (PDB pH 5.2, 28 °C, 110 rpm, 8 days) were supplemented with [1-^{13}C]-acetate to a final concentration of 15 mM. Labeled acetate was fed at days 4, 5, 6 and 7, and the cultures extracted on day 8. A total of 1 mg labeled-**169** was purified by preparative LCMS and submitted to ^{13}C-NMR. The resonance of seven carbons were enhanced, corresponding to C-1, C-3, C-5, C-7, C-9, C-11 and C-12. Signals corresponding to the remaining carbons were barely visible, given the scarce amount of the sample. This result showed **169** to be a heptaketide, most likely derived from a highly reducing type I PKS.
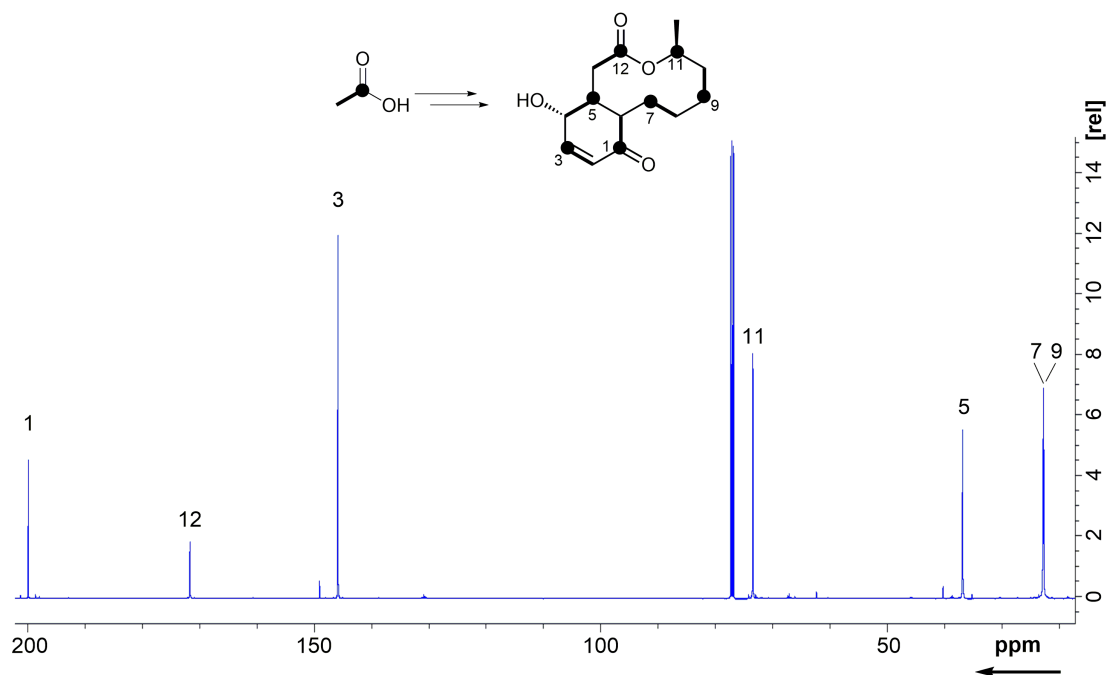
Figure 3.13: ^{13}C-NMR of labeled **169** (1 mg) in CDCl$_3$.

3.3.4 Genomics

To elucidate the biosynthesis of **169** we searched for the corresponding BGC. Genomic DNA of *Phomopsis CMU-LMA* was sequenced in 2005 by Ouazzani, but the raw data were never assembled until 2015 by Daniel Wibberg, who performed a paired-end assembly at the Center for Biotechnology (CeBiTec) in Bielefeld. The output is of poor quality by today's standard, with numerous short scaffolds. The assembly generated 2042 scaffolds with an average length of 31.5 KBp and a scaffold N$_{50}$ of 69.4 KBp (Table 3.2). Considering fungal BGC to possibly have ~40-50 KBp, it may be difficult to obtain the complete BGC for the **169** pathway.

Aligned Reads (All/Paired End)	54,377,401/23,842,462
Assembled Bases	4,201,735,013
PE-Size(s)	883 $\pm$ 239; 884 $\pm$ 239
Scaffolds (All/True)	2,042/2,042
Contigs (Scaffolded/Large($>$500 Bp)/All)	5,742/7,578/28,816
Bases in Scaffolds	64,382,824
Coverage	65.26 x
GC content (%)	52.92
Avg. Scaffold (Bp)	31,529
N_{50} Scaffold	69,395
Largest Scaffold	420,401
Avg. Scaffold Contig	10,656
Avg. Contig	8,321
N_{50} Contig	19,566
Largest Contig	119,136

Table 3.2: Assembling details of *Phomopsis* genome.

Automatic cluster prediction was performed using AntiSMASH[149], the dedicated browser for secondary metabolism queries. Other bionformatical tools were employed to validate intron prediction, sequence homology and protein alignment, such as Softberry FGE-NESH, NCBI BLAST interface, Conserved Domains Database (CDD), Artemis Comparison Tool and Augustus.[232–237] AntiSMASH predicted over 150 gene clusters dedicated to secondary metabolism (Table 3.3). Theoretically, this high number of BGC supposes *Phomopsis CMU-LMA* to produce up to 150 potential compounds, but given the poor quality of the genome assembly we have to be cautious and critical with this result.

BGC type	n°
PKS	60
Terpene	26
Fatty acid	5
NRPS	17
PKS-Terpene	3
Linaridin	1
Lantipeptide	1
Siderophore	1
type III PKS	1
PKS-NRPS	19
PKS-NRPS-Terpene	1
Other	21
TOTAL	**156**

Table 3.3: AntiSMASH prediction of secondary metabolites BGC.

The identification of the **169**-related gene cluster (namely *sch* BGC) was done by rational analysis of **169** structure. **169** belongs to the polyketide family; in addition to the 10-membered macrolactone ester, it has a sole apparent acetate-derived C-1 oxidation at C-7 which presumably arise from a highly reducing PKS that might leave in place the acetate C-1 carbonyl after the first two cycles of chain elongation and reductive events. This keto group could also derive from a monooxygenase or an oxidase such as P450 cytochrome or flavoprotein. The absence of α-methylation suggests a PKS with a non-functional *C*-MeT domain. The origin of the *Z* C-2/C-3 double bond is unknown: it is not the result of the hrPKS DH domain because it lays within an acetate unit, so it could arise from a post elongation dehydratation event, or could be formed during the formation of the 6-membered ring, which may be the result of an enzyme-directed Michael cyclisation.

Other clues point at tailoring events: OH-4 is surely inserted by an oxygenase, most likely a P450 cytochrome or a non-heme iron oxygenase as it is attached at an acetate C-2-derived carbon. Another interesting feature is the terminal C-13 hydroxy group involved in the macrolactonisation of the 10-membered ring: this alcohol could arise from β-keto reduction or from a post elongation decoration driven by a tailoring oxygenase. It is not clear whether the macrolactonisation occurs after the elongation, but some literature speculation suggest the 10-membered ring to be formed after chain release.[301]

AntiSMASH predicted 60 PKS encoding genes, 36 of which encode hrPKS (found by manual BLAST in NCBI and CDD). Manual annotation performed by Softberry FGE-NESH and NCBI BLAST of the genes present within $\sim$50 KBp around the core hrPKS decreased the number of candidate BGC to four, as the remaining did not include P450 cytochromes or a sufficient number of redox enzymes. Finally, domain analysis of the four best candidate hrPKS revealed only one hrPKS with an inactive *C*-MeT domain, making it the best BGC candidate for **169** biosynthesis (namely *sch* BGC, Table 3.4). Moreover, protein BLAST in NCBI of the hrPKS (*sch*PKS) had as first match the PKS involved in brefeldin A **275** biosynthesis. The structural analogies between **169** and **275** (Figure 3.14) suggested that the putative *sch* BGC might be the real **169**-gene cluster. The genes on the right side of the *sch*PKS (*sch*R1 - *sch*R7) encoded for proteins with enzymatic activities, including P450s, non-heme dependent oxidoreductase and hydrolases. Genes on the left side of the core *sch*PKS were found to be DNA-binding proteins such as helicases, DNA binding hypothetical protein, and nesprin-2 like proteins, so they were not included into the borders of *sch* BGC as they probably do not take part in **169** biosynthesis.

Annotation	Function	Organism	Cofactor
*sch*L4	Helicase	*P. solitum*	ATP
*sch*L3	Nesprin-2	*P. expansus*	ATP
*sch*L2	Hypothetical protein	*P. expansus*	-
*sch*L1	Hypothetical DNA binding protein	*P. solitum*	Zn^+
*sch*PKS	hr-PKS	*P. brefeldianum*	-
*sch*R1	α/β-hydrolase	*P. italicum*	-
*sch*R2	Cytochrome p450	*P. italicum*	haem-thiolate
*sch*R3	Cytochrome p450	*P. italicum*	haem-thiolate
*sch*R4	Serine hydrolase	*M. phaseolina*	Ca^{2+}
*sch*R5	Oxidoreductase	*O. sinensis*	NAD
*sch*R6	Transporter	*T. islandicus*	-
*sch*R7	Oxidase	*P. italicum*	FAD

Table 3.4: Candidate *sch* BGC. The core *sch*PKS is annotated in blue; redox encoding genes are marked in red; transporters in gray; and uninvolved genes in white.

The putative *sch*PKS displayed high homology with *bref*-PKS, a related enzyme involved in the biosynthesis of brefeldin A **275** in *Penicillium brefeldianum* (Figure 3.14).[301] Identity between the PKS of the two organisms was calculated using Clustal Omega with default settings at amio acid level, resulting in 56% global identity.
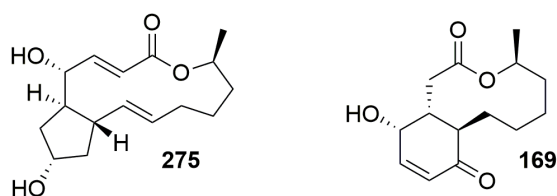
Figure 3.14: Brefeldin A and Sch-642305.

Homology analysis was carried out between the putative **169** BGC and brefeldin A *bref* cluster using the Artemis Comparison Tool (ACT). The *Phomopsis CMU-LMA* cluster was also compared to a 50 KBp region of *P. verucosum* genome (namely *PVSCH* cluster) found by blasting the *sch*PKS amino acid sequence into *P. verrucosum* gDNA (Figure 3.15, Table 3.5). The brefeldin A cluster *bref* is available in the literature.[301] The genome of *P. verrucosum* is accessible in NCBI (BioProject PRJNA276626), submitted by Max Rubner-Institut in Karlsruhe, Germany. We defined and manually annotated cluster *PVSCH* by Softberry FGENESH 2.6 prediction with default settings using *Penicillium chrysogenum* as reference organism.

The hrPKS in the three organisms showed high global homology and also at domain level (Table 3.6). This result was not surprising, as we specifically selected these BGC by

virtue of the homology of their respective PKS. Interestingly, the *trans-α/β*-hydrolase (*sch*R1), was conserved in *bref* and *PVSCH* clusters with identity >50% (Table 3.6). This hydrolase is involved in the release of the growing chain in brefeldin A biosynthesis, as proven by Tang and collaborators,[301] and it is reasonable to assume the same role in the other two systems. Interestingly, two P450s of *Phomopsis CMU-LMA* (*sch*R2 and *sch*R3) showed homology in the other two BGC. Notably, both *sch*R2 and *sch*R3 displayed high homology with three P450 in brefeldin A cluster and with a P450 cytochrome in *P. verucosum* and therefore also with each other. This may indicate a case of gene duplication, and in case of *P. brefeldianum* even gene triplication. The *P. verucosum* BGC would suggest that one P450 is enough for **169** biosynthesis, but the different organisms may have evolved parallel strategies. Moreover, the results showed a conserved oxidase between *P. verucosum* and *Phomopsis CMU-LMA* (*sch*R7), which is not present in *P. brefeldianum*.
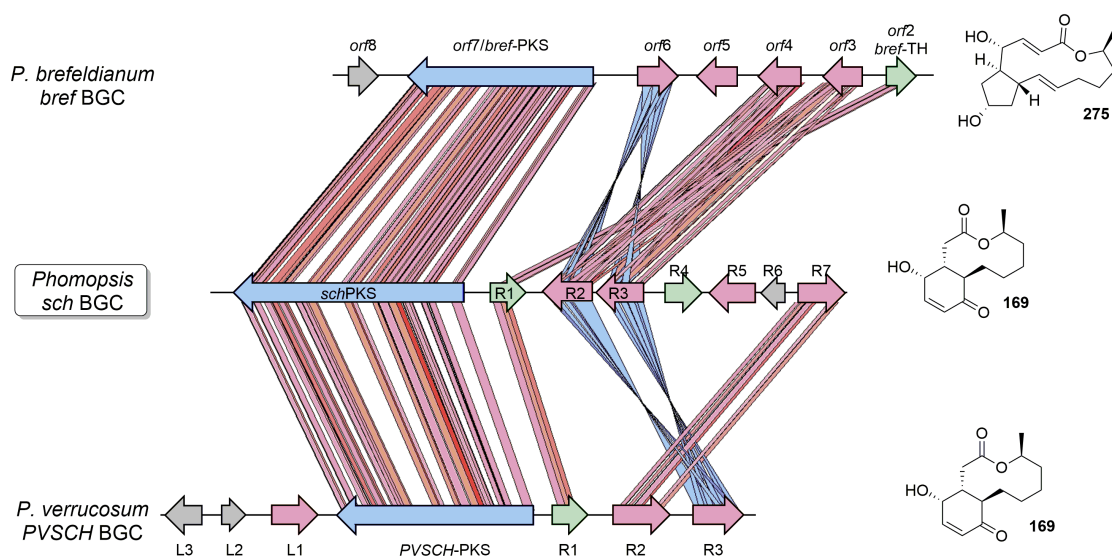
Figure 3.15: BGC homology between *P. brefeldianum*, *Phomopsis*, and *P. verrucosum*. Artemis double ACT was used for the pairwise comparison at translated-nucleotide level, setting a threshold at 20. Red genes encode for oxidoreductive enzymes, green for hydrolase activity, blue for the core PKS. Red lines imply genes on the same DNA strand, blue lines for genes on opposite strands.

Brefeldin A *bref* BGC

Name	Annotation	Cofactor
*orf*8	Large tegument protein	
*orf*7/*bref*-PKS	hr-PKS	
*orf*6	Cytochrome P450	iron, heme
*orf*5	Cytochrome P450	iron, heme
*orf*4	Cytochrome P450	iron, heme
*orf*3	Cytochrome P450	iron, heme
*orf*2/*bref*-TH	α-β hydrolase	

PVSCH cluster

*PVSCH*L3	Phosphatase-like	metal ion
*PVSCH*L2	Hypothetical protein	
*PVSCH*L1	Oxidoreductase	NAD(P)$^+$
*PVSCH*PKS	PKS	
*PVSCH*R1	α-β hydrolase	
*PVSCH*R2	Oxidase	FAD
*PVSCH*R3	Cytochrome P450	iron, heme

Table 3.5: Annotation of the brefeldin A *bref* cluster from *P. brefeldianum* and PVSCH cluster from *P. verrucosum*.

Domain analysis of the core SchPKS was performed using InterPro[304, 305] and confirmed the C-MeT domain to be absent. SchPKS had the complete domain set for total β-reduction, counting DH, ER and KR domains besides the minimal KS-AT-ACP triad. The same analysis was performed on the homologous Bref-PKS and PVSCH-PKS and the homology of the single domains of the three highly reducing systems was calculated by identity matrix using Clustal Omega. In the analysis SchR1, Bref-TH and PVSCH-L1 were also included: the *trans*-acting thiohydrolases necessary for chain release. High homology was observed among the single domains, especially between SchPKS and Bref-PKS (Table 3.6).

SchPKS

KS AT DH ER KR ACP
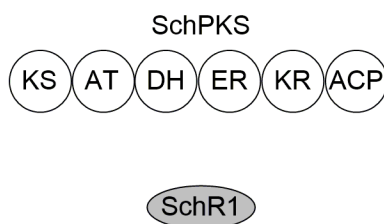
SchR1

| | Identity matrix | |
SchPKS domain	*vs* Bref-PKS	*vs* PVSCH-PKS
KS	70%	66%
AT	55%	56%
DH	52%	43%
ER	63%	61%
KR	61%	58%
ACP	64%	66%
Global identity	56%	51%

	vs Bref-TH	*vs* PVSCH-L1
SchR1	66%	55%

Table 3.6: Comparison of SchPKS with Bref-PKS and PVSCH-PKS domains. Identity was calculated also for the *trans*-thiohydrolase SchR1 against its correspondents of the other BGC. Domain boundaries were estimated by InterPro with default settings and the identity matrix calculated with Clustal Omega with default settings.

Taken together, this evidence is not enough to prove the linkage between the putative *sch* BGC and **169**, but they are sufficient to push forward with the generation of targeted KO against the *sch* BGC. Interference in **169** biosynthesis will prove the role of the BGC and the accumulation of intermediates will reveal the chemical events behind its formation. With this purpose in mind, we established a transformation protocol for *Phomopsis CMU-LMA*.

3.3.5 *Phomopsis CMU-LMA* transformation

In order to find the best selection marker, different antibiotics were tested at increasing concentrations.
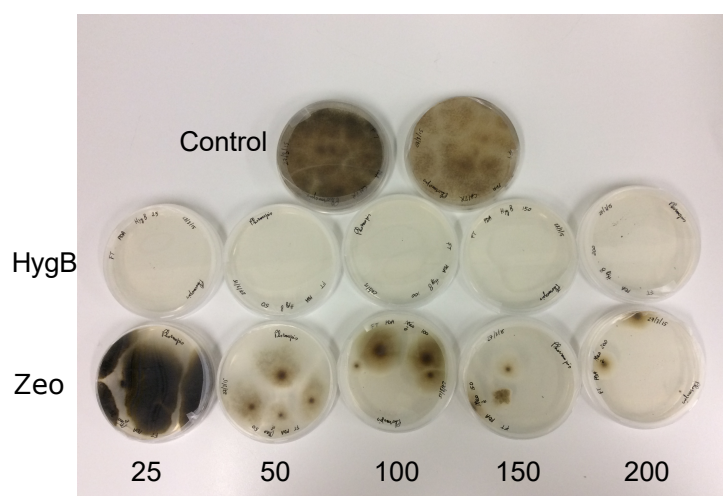
Figure 3.16: Antibiotic testing: hygromycin B (HygB) and zeocin (Zeo) at increasing concentrations (μg $\cdot$ ml^{-1}).

The most effective antibiotic established the corresponding resistance gene to use in transformation experiments. *Phomopsis CMU-LMA* grows even at the highest concentration of zeocin, while hygromycin B (HygB) is effective at low concentrations, with scarce growth at 25 μg/ml and no colonies observed at 50 μg/ml. This antibiotic prevents mRNA translation by binding the 30S ribosomal subunit and inhibiting the translocation step of elongation.[306–308] HygB resistance gene *hph* consists of a kinase that inactivates the antibiotic through phosphorylation. The gene *hph* was isolated in 1983 from *E. coli*[254, 255] and it was previously engineered under the control of the constitutive fungal promoter P_{gdpA} in the vector pTH-GS-eGFP. This plasmid was used to define a transformation protocol, as it contains the enhanced Green Fluorescent Protein reporter gene *eGFP* under the control of the amylose-inducible promoter *amyB* (P_{amyB}), and to test whether the P_{gdpA}-*hph* construct confers resistance in *Phomopsis CMU-LMA*. Onset of a fluorescent phenotype in the transformant cells grown in starch-rich medium will also confirm P_{amyB} to be functional in the fungus.

We developed a transformation protocol *via* protoplast formation: the cell wall of young mycelia was digested by a mixture of lysing enzymes extracted from *Trichoderma harzianum* 10 mg/ml in conjunction with driselase from *Basidiomycetes spp.* 5 mg/ml in osmotic buffer NaCl 0.8 M. Protoplasts were counted by hemocytometer (average 2 x 10^7 cells $\cdot$ ml^{-1}) and incubated in an osmotic buffer with 10 μg of plasmid DNA (Figure 3.17).
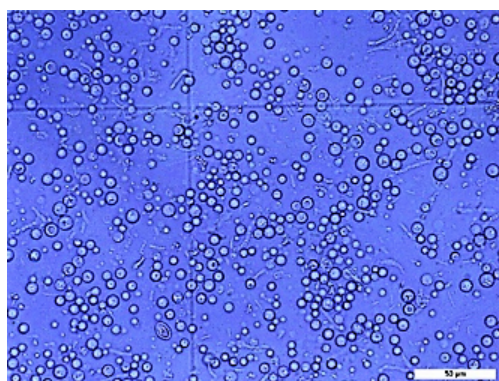
Figure 3.17: Round protoplasts of *Phomopsis CMU-LMA* in osmotic buffer.

Cells were plated on agar with hygromycin B (100 μg/ml) and growing mycelia were picked for secondary HygB plates (also 100 μg/ml) and microscopy to detect any fluorescence activity (Figure 3.18). Fluorescence microscopy, resistance to antibiotic and PCR evidences (not shown) confirmed that *Phomopsis CMU-LMA* was possible to transform with relatively little effort. For future KO experiments bipartite transformation will be adopted.
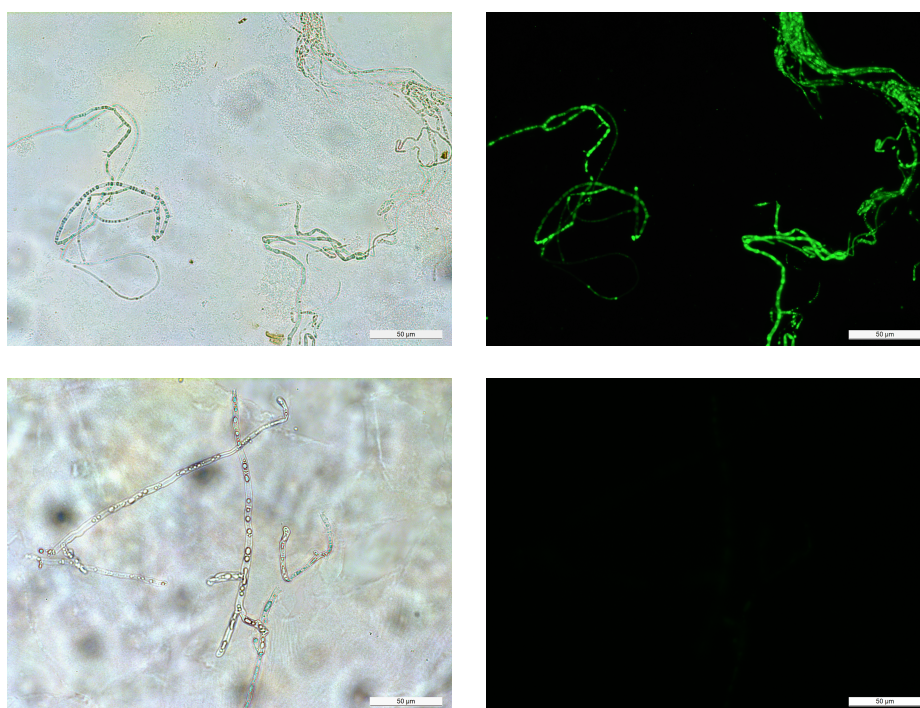
Figure 3.18: pGS-TH-eGFP transformant mycelia under white and UV light (right and left top). WT cells display no fluorescence (right and left bottom).

3.3.6 Building the knock out cassette by yeast recombination

We designed a KO-cassette composed of three DNA fragments cloned into the shuttle pE-YA vector (Scheme 2.6 section 2.3.8). In particular, the hygromycin B resistance cassette $HygR$ (P_{gdpA} + hph, ~4 KBp) was flanked by sequences homologous to the target gene (Target Left T_L and Target Right T_R, ~1 KBp each). T_L, $HygR$ and T_R were linked together and cloned into pE-YA by yeast recombination. T_L and T_R were obtained by Taq PCR from gDNA, and $HygR$ by high fidelity Q5 PCR to preserve the resistance effectiveness. Primers were designed with 30 nt tails homology to guide the site-directed recombination in yeast (Scheme 3.7).
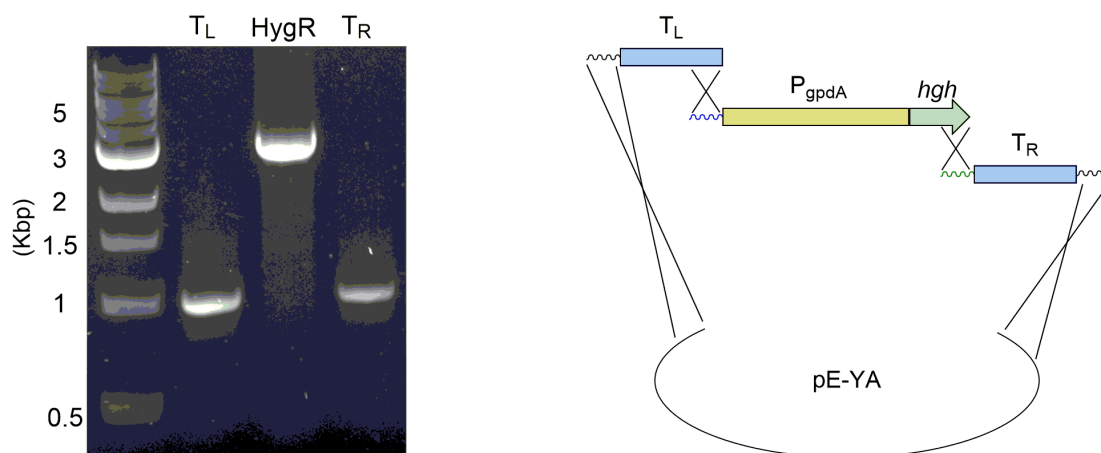
Scheme 3.7: KO cassette construction into pE-YA vector by yeast recombination.

A total of 5 constructs were built, to target genes schPKS (hrPKS), schR2 (P450), schR3 (P450), schR5 (oxidoreductase) and schR7 (FAD oxidase). The plasmids were confirmed by enzymatic digestion (not shown).

3.3.7 Fungal Transformation

We performed bipartite knock out aiming at sch cluster genes. Protoplasts were generated and transformed with two non-functional overlapping fragments (α and β) of the KO cassette and plated on hygromycin B (50 μg · ml^{-1}). Colonies resistant to antibiotic were picked and transferred on secondary hygromycin B plates (100 μg · ml^{-1}), Figure 3.19).
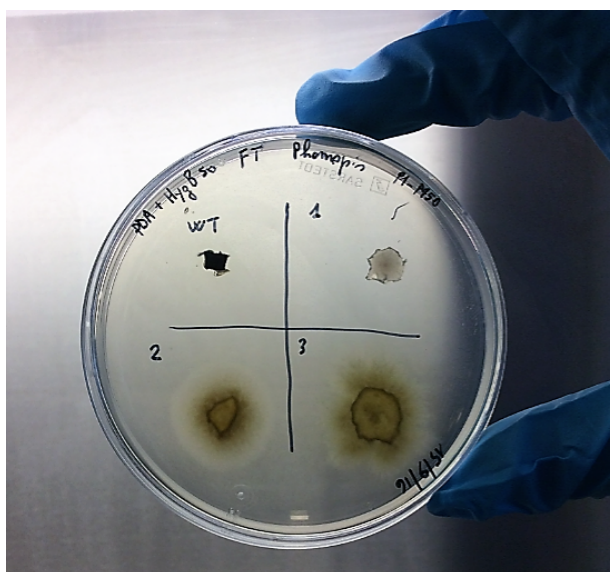
Figure 3.19: *Phomopsis CMU-LMA* transformants on secondary HygB plates. The WT control cannot grow (top left corner). Transformant mycelia survive the antibiotic.

Two PCR on genomic DNA (gDNA) were set up to confirm the success of the transformation: *whole* PCR (wPCR) and *external* PCR (extPCR). wPCR reveals whether the target gene was successfully disrupted: it covers the target gene from start to end (except for *sch*PKS), and it normally gives a product exclusively in the WT. Special mention for the *sch*PKS gene: instead of covering the whole gene, wPCR spans ~1 KBp of the internal region, which must be removed in case of successful disruption. wPCR products are also observable in the transformant at ~4.5 KBp, which correspond to the whole KO cassette. extPCR reveals whether the recombination happened at the correct site: it spans from the 5'end of the *hph* resistance to the genomic DNA adjacent the target gene. extPCR can give a product only in case of a successful disruption in the correct locus, and cannot give a product in the WT, since it lacks the *hph* gene (Scheme 2.8 section 2.3.8, Figure 3.20).
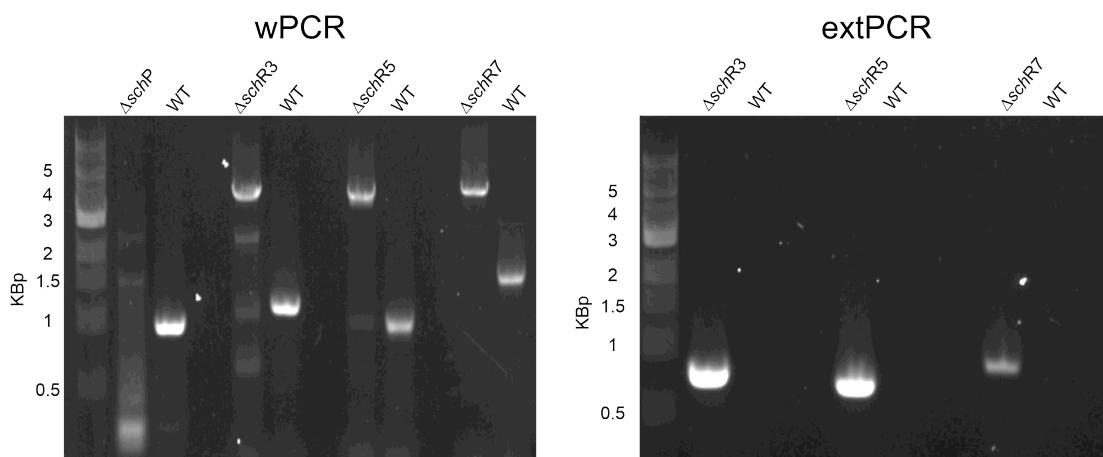
Figure 3.20: wPCR and extPCR products. See scheme 2.8 for primer positions.

Phomopsis CMU-LMA was prone to transformation: it generated numerous protoplasts and integrated DNA with high efficiency, resulting in a large number of colonies in every transformation round. Nevertheless, the rate of successful targeted integration was quite low: only 8% of screened transformant colonies displayed a site-specific integration of the KO cassette (Table 3.7).

Target gene	Screened colonies	True KO mutant	% Success
*sch*PKS	10	2	20.0
*sch*R2	30	2	6.7
*sch*R3	48	3	6.3
*sch*R5	31	2	6.5
*sch*R7	18	1	5.6
TOT	137	10	7.3

Table 3.7: *Phomopsis CMU-LMA* transformation rate.

3.3.8 Δ*sch*PKS mutant fermentation

The two true Δ*sch*PKS transformants were fermented in producing condition (PDB pH 5.2; 28 °C; 100 rpm) for 5 days. Interestingly the Δ*sch*PKS strain lost the characteristic dark pigmentation of the WT.

After fermentation, the culture supernatant was extracted with 1.5 volume equivalent ethyl acetate and the crude examined by LCMS. Analytical LCMS of supernatant extracts revealed compounds LMA-P3 **272** and DHTO **273** to be produced and **169** to disappear (Figure 3.21). This showed that the correct BGC was identified and that **272** and **273** are unrelated biosynthetically.
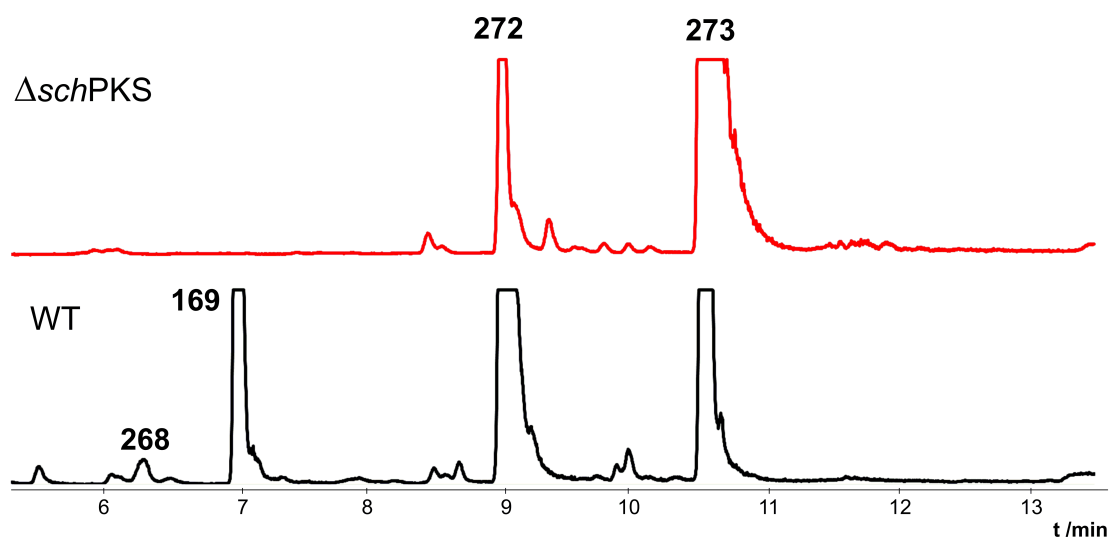
Figure 3.21: ELSD chromatograms of WT *Phomopsis* extract (black) and ΔschPKS crude (red).

3.3.9 ΔschR7 mutant fermentation

The single true mutant ΔR7 grown in producing conditions (PDB pH 5.2, 5 days, 28 °C, 120 rpm) yielded a new metabolite that we called R7A **292** (Table 3.8, Figures 3.22 and 3.23). Compound **169** disappeared in the mutant extraction, proving that its biosynthesis was interrupted.

Compound	RT /min	UV$_{max}$ /nm	Mass /Da	Amount /mg
R7A **292**	4.8	218	252.1	2.7

Table 3.8: New peak **292**.
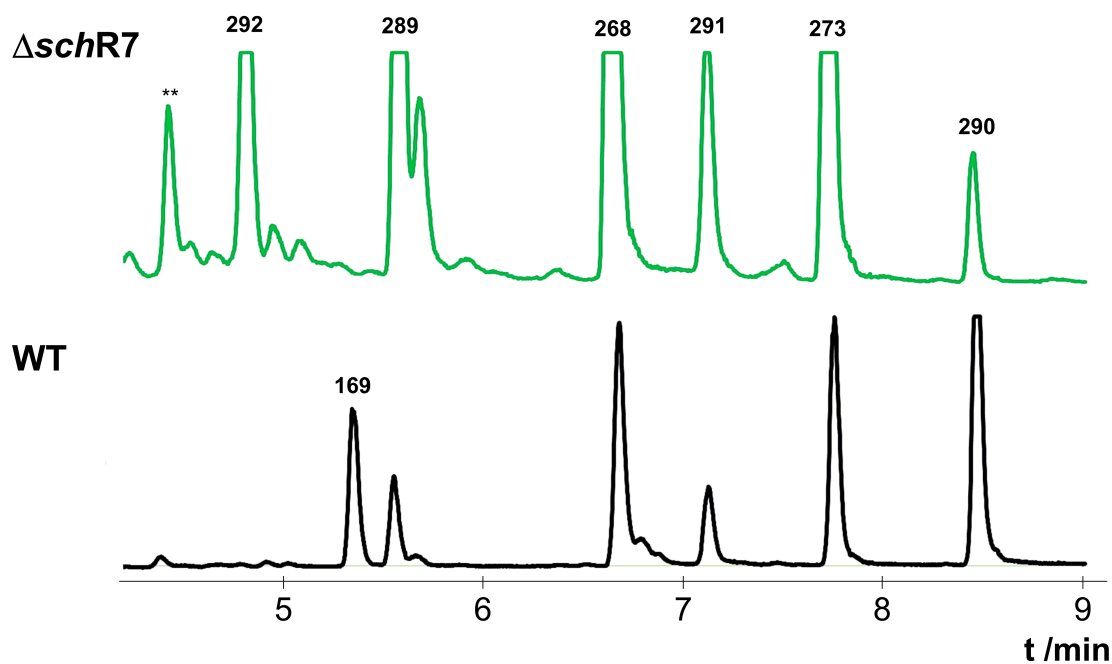
Figure 3.22: ELSD chromatograms of crude extracts from WT *Phomopsis* (black) and ΔR7 (green).
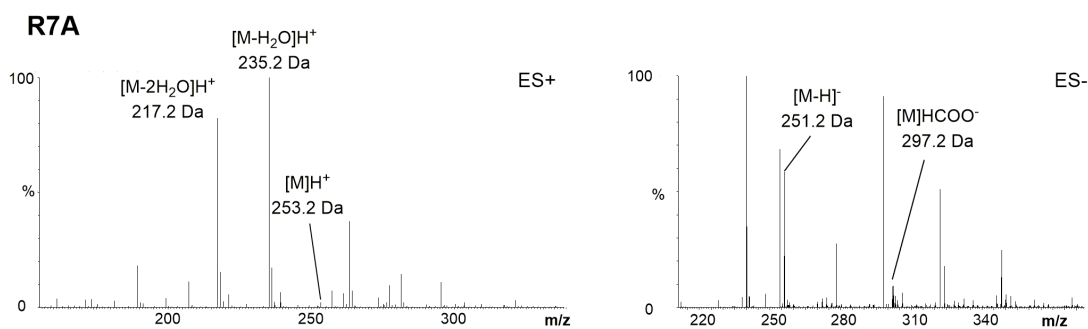
Figure 3.23: Mass spectra of peak **292**.

HRMS data for peak R7A **292** ($[M]H^+$ calculated $C_{14}H_{20}O_4$ 253.3180, found 253.3178) **292** suggested compound **292** to be related to **169** because of the same number of carbons and mass in the range of a possible intermediate. HRMS data are hereby reported:

R7A (**292**): UV (diode array HPLC, H_2O/CH_3CN) λ_{max} = 218 nm LRMS m/z 253.2 $[M]H^+$; $[M - H_2O]H^+$ 235.2; $[M - 2H_2O]H^+$ 217.2; $[M - H]^-$ 251.2 HRMS m/z 253.3178 $[M]H^+$ (calcd for $C_{14}H_{21}O_4^+$, 253.3180); 5 DBEs.

Compound R7A was purified (2.7 mg) and submitted to NMR using $CDCl_3$ or acetone-d_6 as solvents (Table 3.9, Figures 3.24 - 3.27). ^{1}H-NMR confirmed **292** to be a macrolactone, displaying the terminal C-14 methyl as a doublet at 1.26 ppm (Figure 3.24). The COSY spectrum showed atom connectivity that reflected those of mutolide **259**[309] and nigrosporolide **293**,[310] but the literature data of neither of these compounds had a perfect match with the chemical shifts of **292** (Figure 3.27, Tables 3.9 and 3.10), opening the question of the stereochemistry of compound R7A.

292
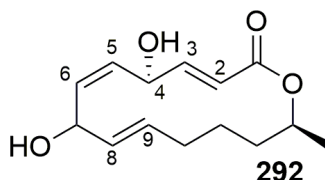
		CDCl₃		Acetone-d₆	
		δ_C/ppm	δ_H/ppm	δ_C/ppm	δ_H/ppm
	Position	100 MHz	400 MHz	100 MHz	400 MHz
CO₂	1	166.1		166.7	
CH=CH	2	118.2	5.91 dd (1 H, 15.6, 1.8)	117.7	5.81 dd (1H, 15.6, 1.8)
CH=CH	3	148.2	7.03 dd (1H, 15.6, 4.5)	151.3	7.13 dd (15.6, 4.5)
CH-OH	4	68.5	5.27 m (1H)	68.5	5.26 m (1H)
CH=CH	5	132.0	5.54 m (1H)	132.9	5.45 m (1H)
CH=CH	6	131.2	5.46 m (1H)	132.8	5.36 m (1H)
CH-OH	7	70.6	4.82 t (1H)	70.5	4.82 td (1H, 16, 7.9, 3.4)
CH=CH	8	134.4	5.87 m (1H)	133.0	5.79 m (1H)
CH=CH	9	130.2	5.51 m (1H)	132.7	5.4 m (1H)
CH₂	10a	32.8	2.08 m (1H)	33.6	
	10b		1.97 m (1H)		1.9 m (1H)
CH₂	11a	24.4	1.82 m (1H)	25.5	1.79 m (1H)
	11b		1.12 m (1H)		1.04 m (1H)
CH₂	12a	34.5	1.75 m (1H)	35.2	1.7 m (1H)
	12b		1.5 m (1H)		1.5 m (1H)
CH-OH	13	72.05	4.76 m (1H)	72.3	4.6 m (1H)
CH₃	14	20.2	1.26 d (3H, 6.2)	20.6	1.2 d (3H, 6.2)
OH					3.95 d (1H, 3.6)

Table 3.9: Chemical shifts of ∼3 mg of **292** in $CDCl_3$ and acetone-d_6. Coupling constant are measured in Hz.

Figure 3.24: ^{1}H-NMR of compound **292** in CDCl$_3$.
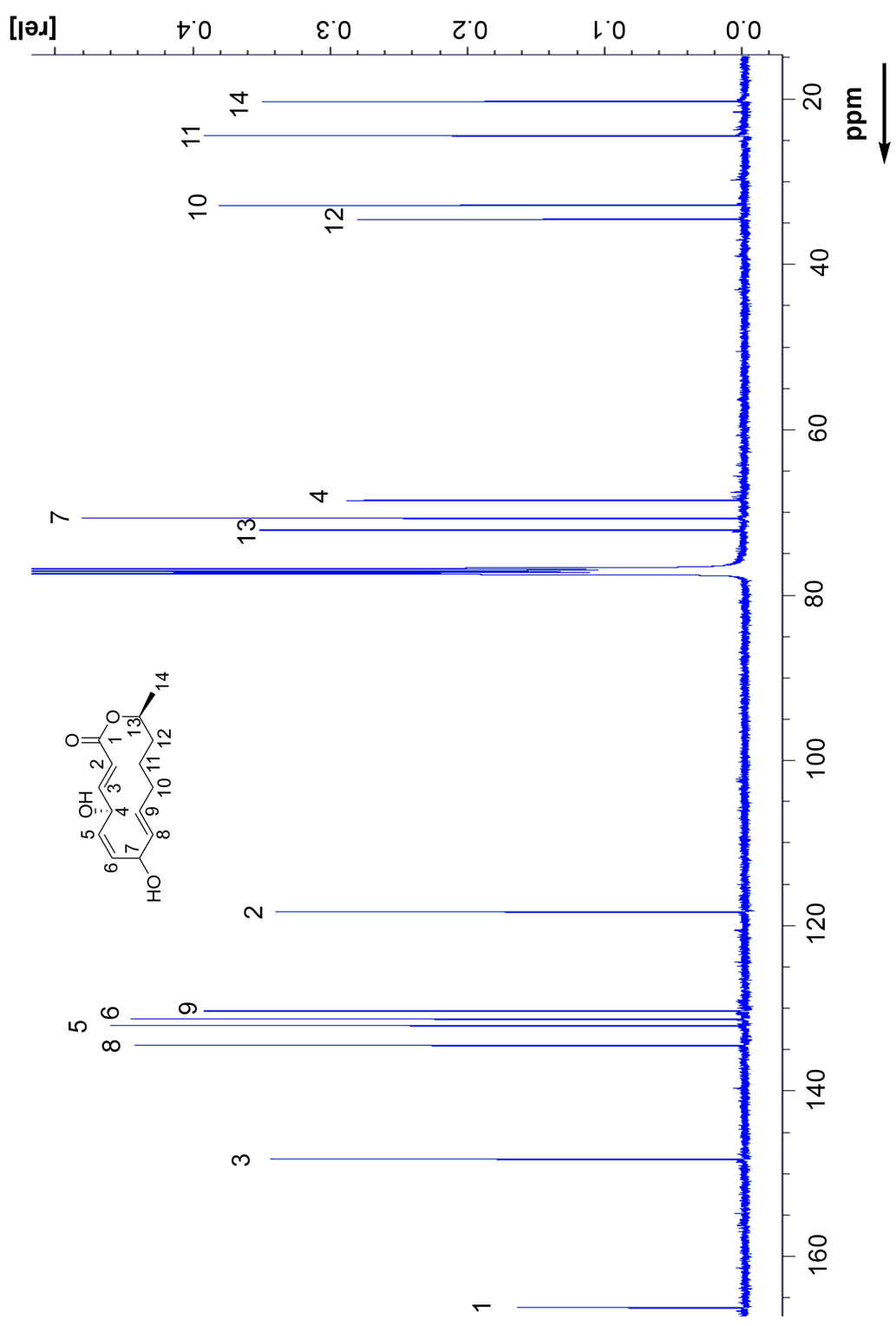
Figure 3.25: ^{13}C-NMR of compound **292** in CDCl$_3$.
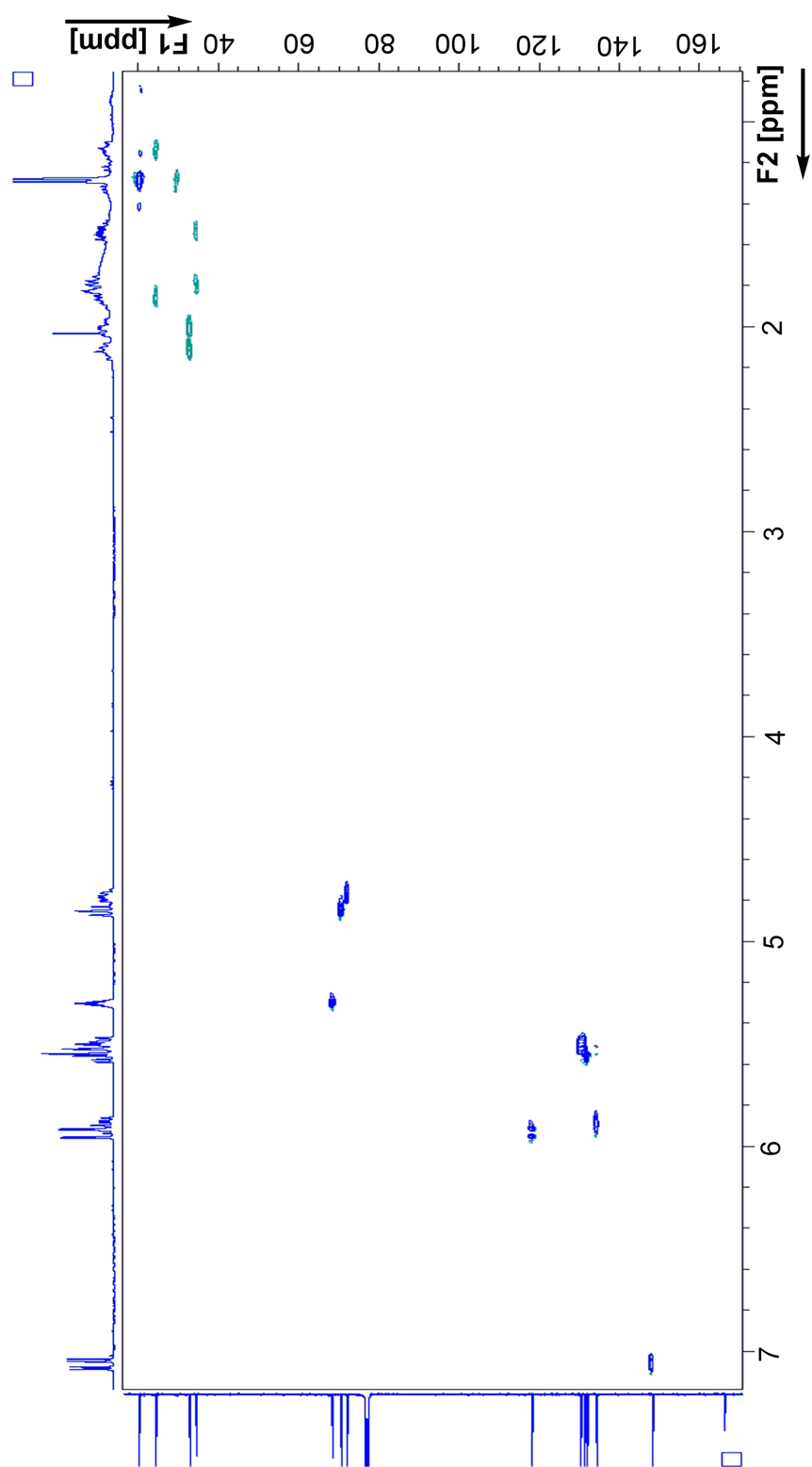
Figure 3.26: HSQC NMR of compound **292** in CDCl$_3$.
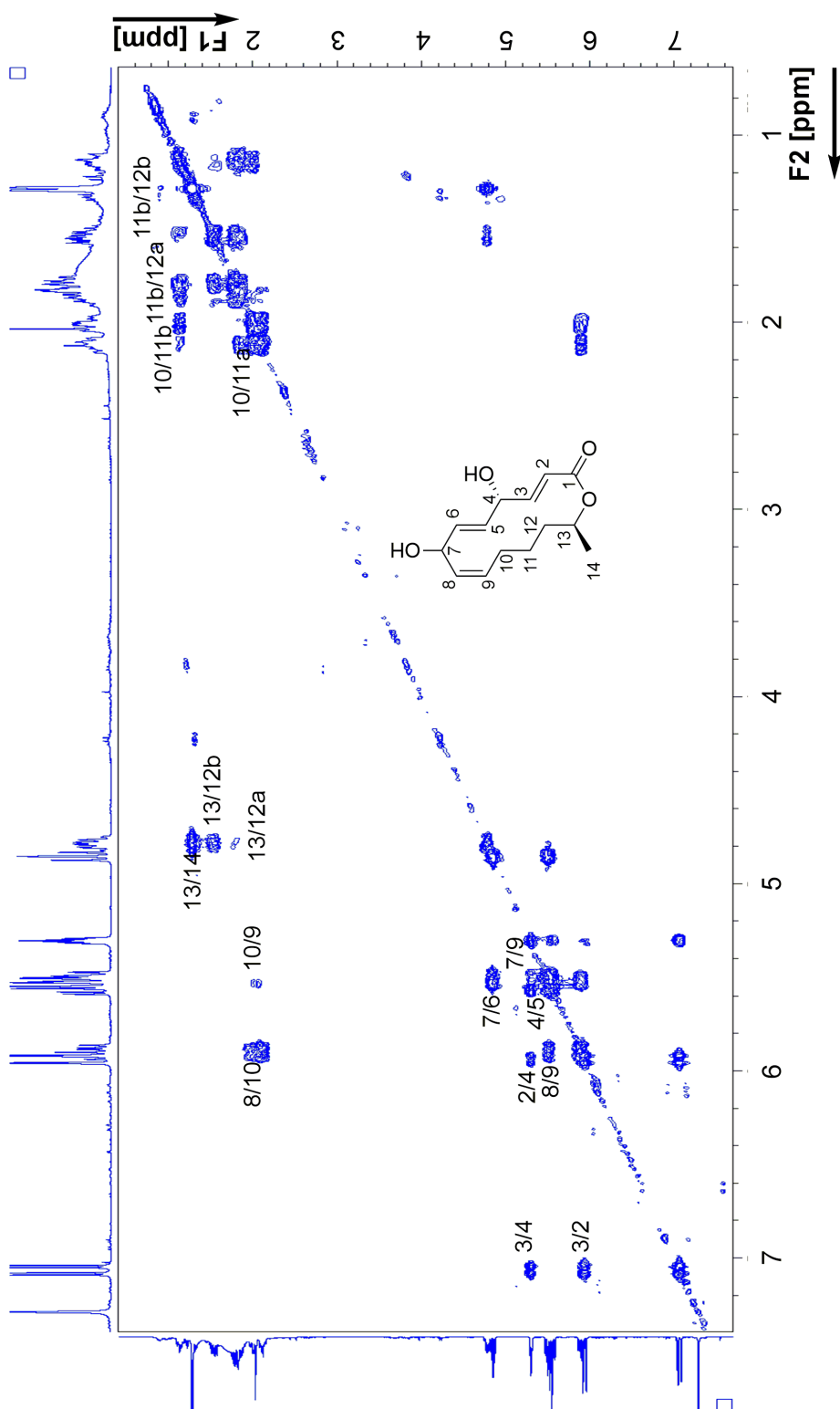
Figure 3.27: COSY NMR of compound **292** in CDCl₃.

Nigrosporolide **293**

Mutolide **259**

	Nigrosporolide 293		**Mutolide 259**	
	δ_C/ppm	δ_H/ppm	δ_C/ppm	δ_H/ppm
Assignment	CDCl$_3$	CDCl$_3$	Acetone-d$_6$	Acetone-d$_6$
1	166.4		168.6	
2	121.3	6.08 dd (1H, 15.6, 1)	118.8	5.81 dd (1H, 16, 1)
3	147.3	6.93 dd (1H, 15.6, 7.1)	152.9	6.7 dd (1H, 16, 7)
4	69.1	5.25 m (1H)	70.8	4.89 m (1H)
5	133	5.59 dd (1H, 11.4, 4.3)	131.1	5.41 m (1H)
6	131.5	5.45 dd (1H, 11.4, 2.2)	135	5.77 ddd (1H, 16, 7, 1)
7	63.7	5.12 bm (1H)	73	4.62 s, br (1H)
8	129.3	5.32 bdd (1H, 10.6, ca 2)	135.6	5.57 m (1H)
9	133.2	5.37 dd (1H, 10.6, 3.5)	132.6	5.41 m (1H)
10	29.5	2.48 m (1H)	31.7	1.94 m (2H)
		1.98 m (1H)		
11	25.6	1.72 m (1H)	25	1.4 m (2H)
		1.15 m (1H)		
12	34.5	1.87 m (1H)	35.5	1.52 m (2H)
		1.45 m (1H)		
13	73	4.99 m (1H)	72.8	4.98 m (2H)
14	20.4	1.27 d (3H, 6.3)	18.9	1.18 d (3H, 6.5)
OH-4				4.5 s, br (1H)
OH-7				3.96 s, br (1H)

Table 3.10: Literature NMR of nigrosporolide (*2E,5Z,8Z* isomer) and mutolide (*2E,5E,8E* isomer). Coupling constants are measured in Hz.

$|\Delta\delta_C|$ was calculated for nigrosporolide and mutolide respectively as the difference between the carbon chemical shift of **292** *vs* **293** and *vs* **259** carbon shifts. **292** NMR spectra were recorded in CDCl$_3$ and acetone-d$_6$ to have a better comparison with the literature data. $|\Delta\delta_C| \geq 2$ ppm were considered significantly different. Interestingly, comparison between **292** and **293** showed C-2, C-7, C-8, C-9 and C-10 as nuclei with significant differences in chemical shift, suggesting that possibly the C-8/C-9 olefin configuration might be inverted. Similarly, comparison between **292** and **259** showed C-4, C-6, C-7 and C-8 as centers with major chemical shift difference, suggesting the *E* 5-6 double bond to be *Z* in **292**. This observation was in accord with the fact that $|\Delta\delta_C|$ between **292** and **293** at C-5 and C-6 positions are < 1 ppm (Table 3.11).

Table 3.11: Difference of carbon chemical shift ($|\Delta\delta_C|$) of **292** *vs* **293** and *vs* **259**. Significant threshold was set at $|\Delta\delta_C| \geq 2$ ppm (marked with • on structures). **292** δ_C were measured in different solvents to match the literature records of **293** and *vs* **259**.
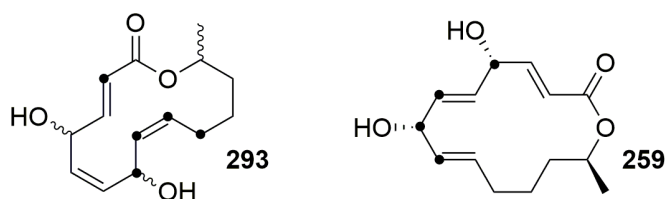
Position	$|\Delta\delta_C|$ *vs* **293** / CDCl$_3$ (ppm)	$|\Delta\delta_C|$ *vs* **259** /acetone-d$_6$ (ppm)
1	0.3	1.9
2	3.1	1.1
3	0.9	1.6
4	0.6	2.3
5	1.0	1.8
6	0.3	2.2
7	6.9	2.5
8	5.1	2.6
9	3.0	0.1
10	3.3	1.9
11	1.2	0.5
12	0.0	0.3
13	0.9	0.5
14	0.2	1.7

[1]H-decoupling experiments were performed to simplify the **292** ^{1}H-NMR spectrum, suggesting a *2E,5Z,8Z* stereoisomery of the latter basing on coupling constants. In particular, removal of coupling to H-4 converted the H-2 and H-3 doublet of doublets into doublets with the same coupling constant (15.6 Hz), indicative of a *E*-olefin; the *E* stereochemistry of the C-2/C-3 olefin was already clear from standard ^{1}H-NMR (Figure 3.28 **A**). Much more interesting was the region between 5.45 and 5.60 ppm, where H-5, H-6 and H-9 signals were superimposed (Figure 3.28 **B**). H-4 decoupling considerably simplified this region, allowing a more accurate peak assignment. The coupling constant between H-5 and H-6 was found to be 11.3 Hz, corresponding to a *cis*-double bond. This was in accord with the observation that the $J_{5\text{-}6}$ of nigrosporolide **293** (isomer *2E,5Z,8Z*) had a very similar value of 11.4 Hz (Table 3.10).
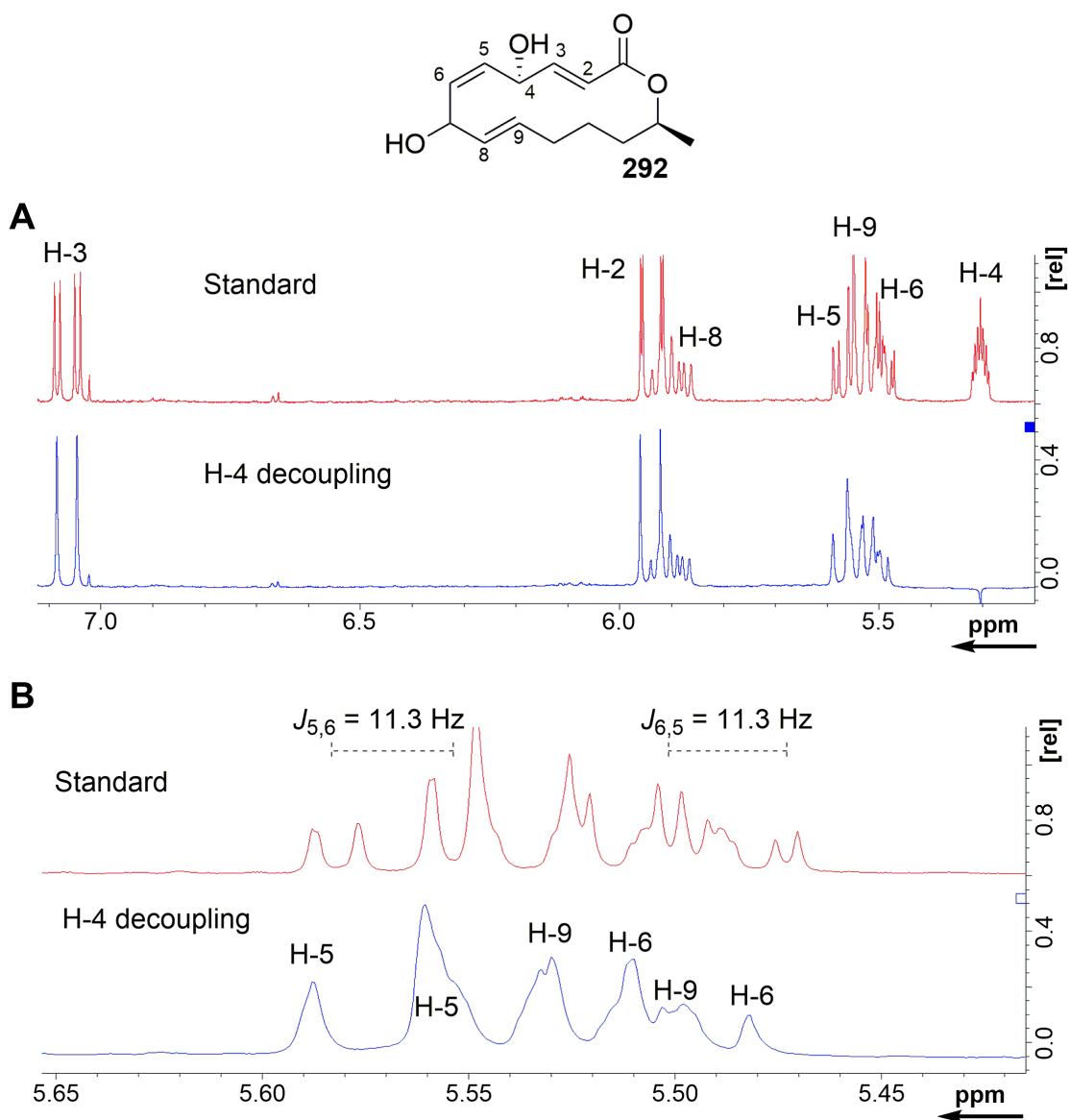
Figure 3.28: Decoupling of H-4 in **292** ^{1}H-NMR.

H-7, H-10 decoupling was performed in two separate experiments and both simplified signals relative to H-8 and H-9. $^3J_{8,9}$ was measured 15.1 Hz, reflecting a *trans*-double bond between these two centers (Figure 3.29). Interestingly, decoupling H-7 did not alter signal H-8, meaning that these two protons do not couple because of their geometry: the dihedral angle between each other must be around 90°. The *E* double bond between nuclei 8 and 9 was confirmed by comparing the coupling constant with **293**, which has $^3J_{8,9} = 10.6$ Hz, corresponding to a *Z* configuration (Table 3.10).
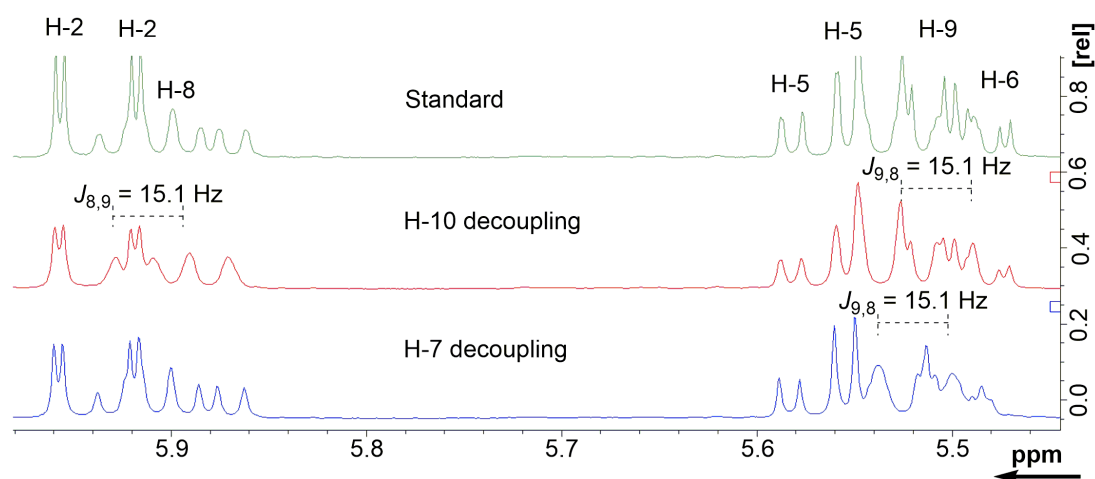
Figure 3.29: Decoupling of H-7 and H-10 in **292** ^{1}H-NMR.

Given the literature comparison between **292**, **293** and **259**, and the decoupling experiments, we could assess compound **292** to have a $2E$, $5Z$, $8E$ stereochemistry, making it a distinct new compound generated by KO.

3.3.10 ΔschR3 mutant fermentation

The three mutants ΔschR3 (P450) were fermented in the dark in M1D medium for 10 days at 28 °C in static conditions. The culture was homogenised, the cells filtered and the supernatant was extracted with ethyl acetate. Low resolution LCMS analysis showed the disappearance of **169** and a new peak, that we labeled R3A **294**, with an apparent mass of 236.2 Da (Table 3.12, Figure 3.30, 3.31). The mycelia was extracted separately, but no new compounds were detected (data not shown).

Compound	RT /min	UV$_{max}$ /nm	Mass /Da	Amount /mg
R3A **294**	17.1	210	236	0.8

Table 3.12: New peak **294**.

High resolution MS analysis confirmed the mass of compound R3A ([M]H$^+$ calculated $C_{14}H_{20}O_3$ 237.3185, found 237.3181) suggested that compound **294** might be related to **169** because of the same chain length. The retention time at 17 minutes indicated a rather aliphatic compound. This observation was in line with the fact that mutant ΔschR3 lacked a P450 oxidase, consistent with interrupting **169** pathway at a less oxidised stage. HRMS data are hereby reported:

142

R3A (**294**): UV (diode array HPLC, H_2O/CH_3CN) $\lambda_{\max} = 210$ nm m/z 237.3181 $[M]H^+$ (calcd for $C_{14}H_{21}O_3{}^+$, 237.3185); 5 DBEs.
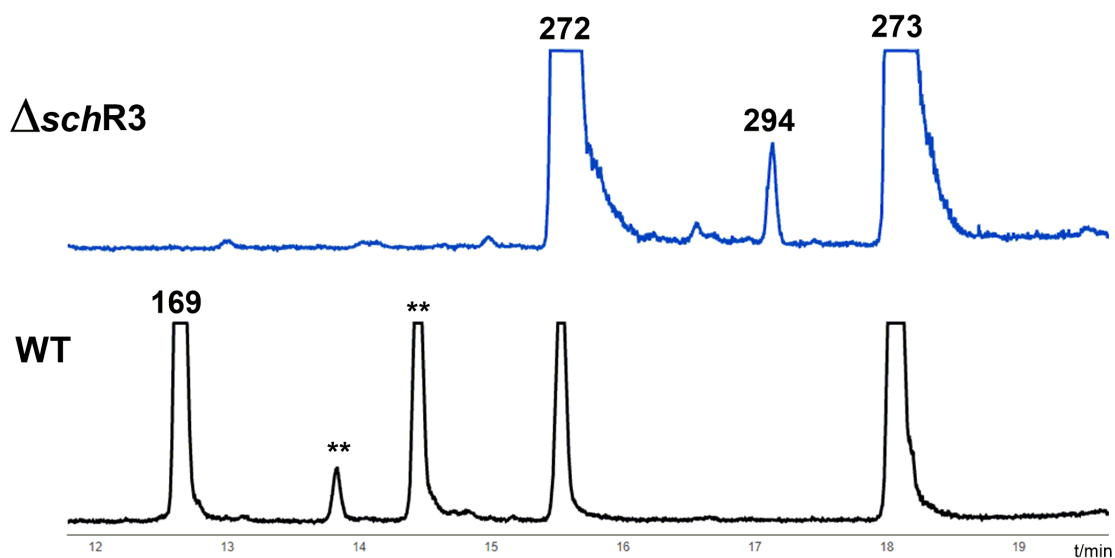
Figure 3.30: ELSD chromatograms of WT *Phomopsis* extract (black) and ΔR3 crude (blue). Unrelated peaks are marked with ** .
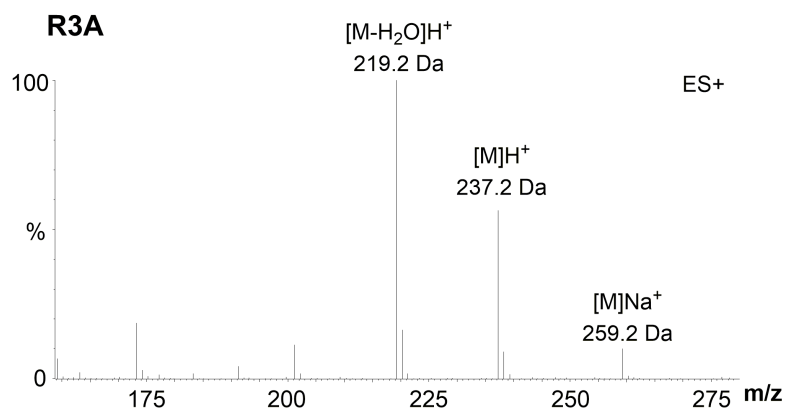
Figure 3.31: ES+ Mass spectrum of compound **294**.

Full NMR data acquisition revealed compound R3A to be a 14-member macrolactone related to benquoin and to **292**, with a single hydroxylation at C-4 (Table 3.13, Figures 3.32, 3.33).
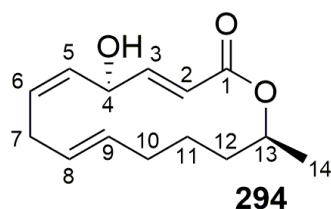
	Position	δ_C/ppm (100 MHz)	δ_H/ppm (400 MHz)
		294, CDCl$_3$	
CO$_2$	1	166.3	
CH=CH	2	117.5	5.94 dd (1 H, 15.7 Hz, 1.9 Hz)
CH=CH	3	149.6	7.18 dd (1H, 15.7 Hz, 3.9 Hz)
CH-OH	4	68.3	5.26 m (1H)
CH=CH	5	131.5	5.51 m (1H)
CH=CH	6	128.6	5.54 m (1H)
CH$_2$	7	32.1	2.82 m (1H)
CH=CH	8	132.1	5.55 m (1H)
CH=CH	9	126.6	5.43 m (1H)
CH$_2$	10a	32.9	2.08 m (1H)
	10b		1.91 m (1H)
CH$_2$	11a	24.4	1.78 m (1H)
	11b		1.19 m (1H)
CH$_2$	12a	34.3	1.74 m (1H)
	12b		1.51 m (1H)
CH-OH	13	71.8	4.82 m (1H)
CH$_3$	14	20.1	1.27 d (3H, 6.4 Hz)

Table 3.13: Chemical shifts of ~1 mg **294** in CDCl$_3$.

144
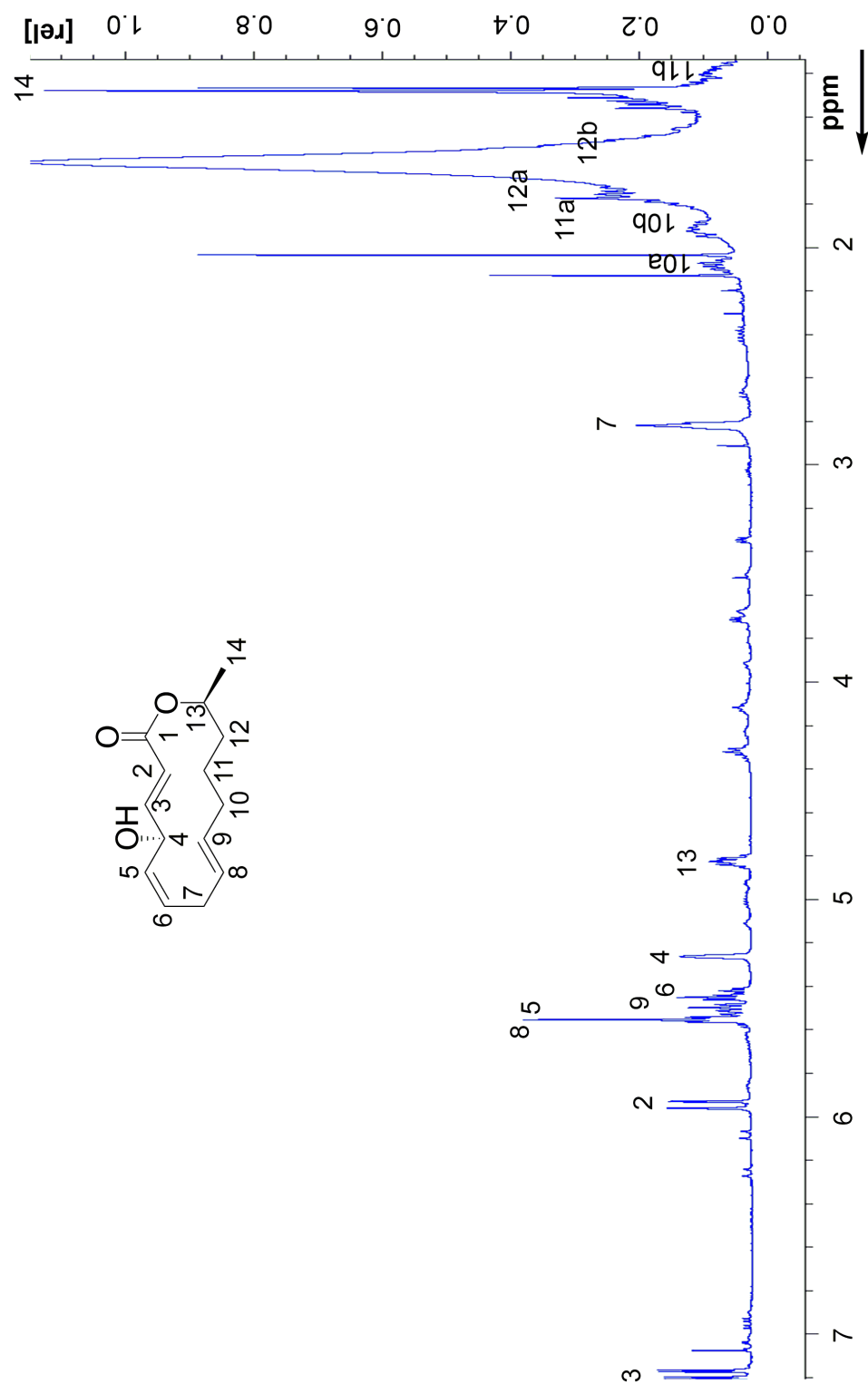
Figure 3.32: ^{1}H-NMR of **294** in CDCl$_3$.
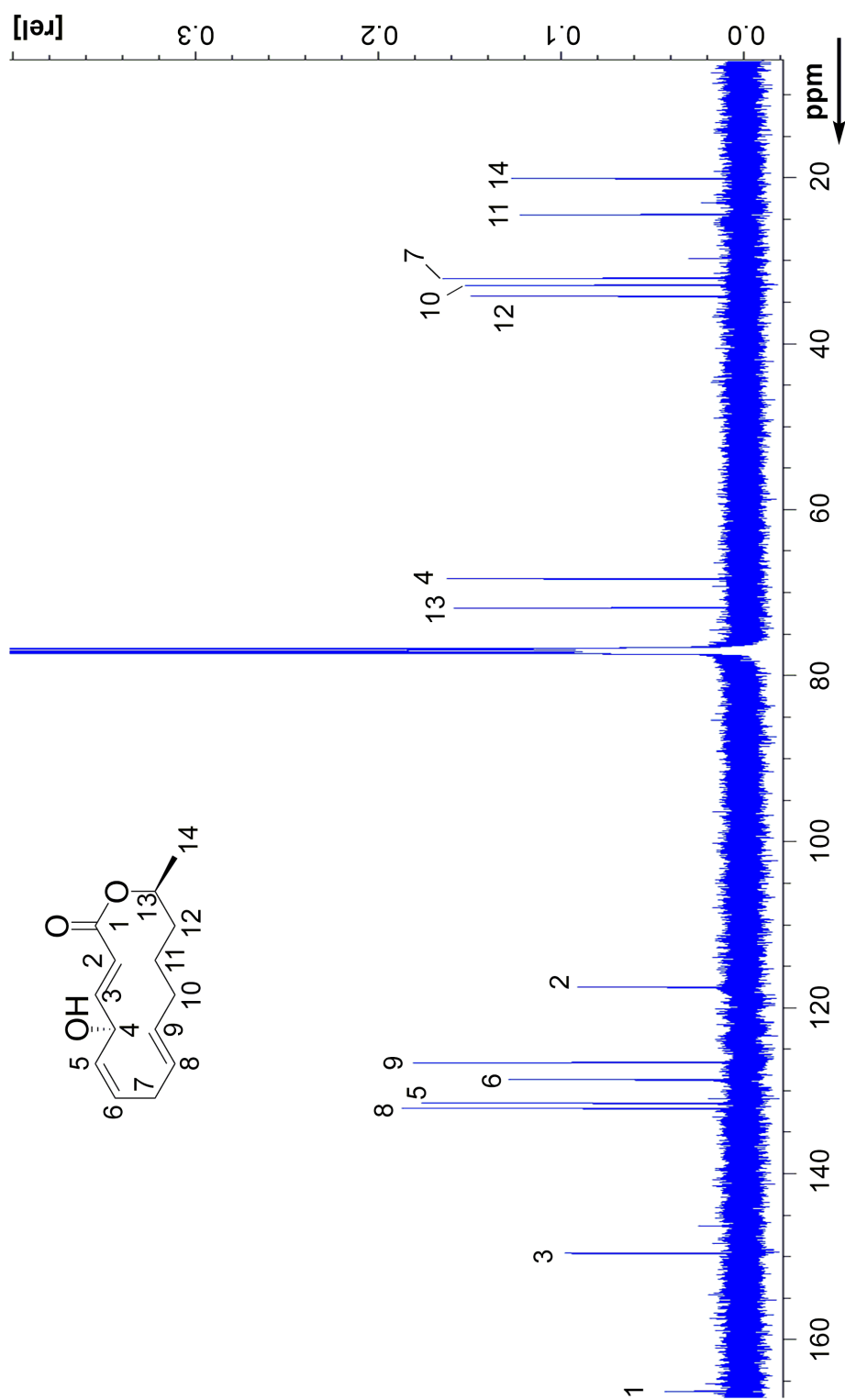
Figure 3.33: ^{13}C-NMR of **294** in CDCl$_3$.

NMR spectra of **292** and **294** were extremely similar, with the exception of signals relative to nuclei 7. The disappearance of chemical shift 70.6 ppm, relative to C-7 of **292**, in **294** ^{13}C-NMR showed that the hydroxylation on this center was absent in the ΔschR3 isolate (Figure 3.34). Appearance of a multiplet at 2.82 ppm in **294** ^{1}H-NMR, was coherent with the doubly allylic position of center 7 (Figure 3.35).
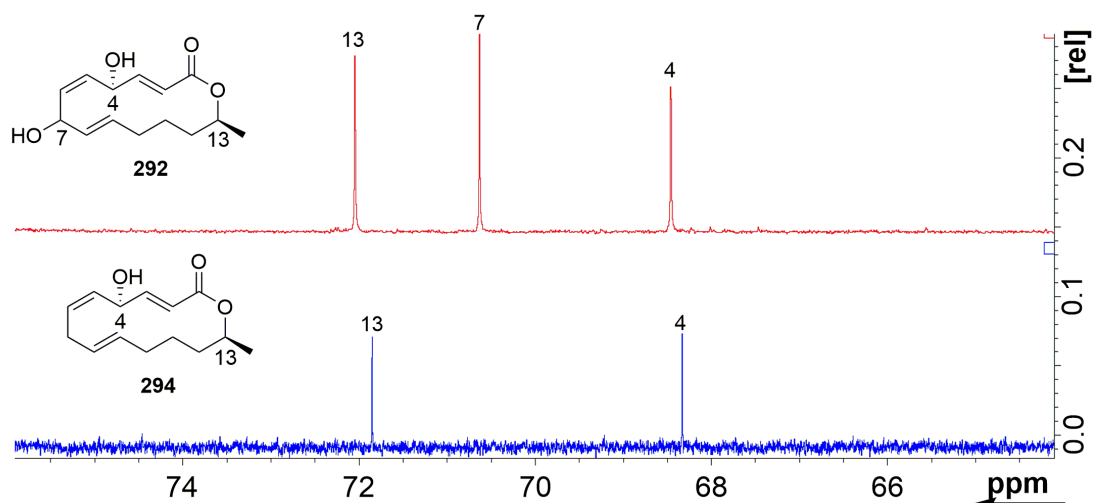
Figure 3.34: Disappearance of C-7 signal in **294** ^{13}C-NMR within the C-OH range.
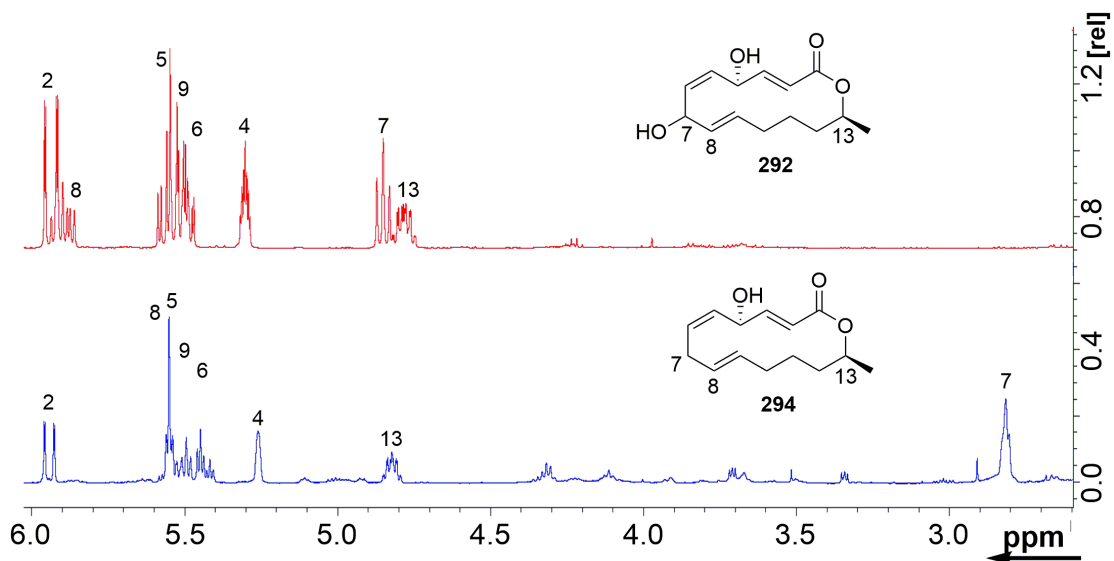
Figure 3.35: Shift of H-7 from 4.82 ppm (**292**) to 2.82 ppm (**294**).

3.3.11 Δ*sch*R5 mutant fermentation

Two true Δ*sch*R5 KO mutants were obtained by bipartite transformation and confirmed by PCR evidence. Neither of the two Δ*sch*R5 mutants could yield **169** in producing conditions, meaning that SchR5 is an essential biosynthetic gene (Figures 3.36, 3.37). Interestingly, the peak corresponding to the intermediate LMA-P1 **268** also disappeared in the Δ*sch*R5 mutants (Figure 3.37), strengthening the hypothesis that **169** and **268** are related. Different media, culture volumes and fermentation times were tried with no significant outcomes. Medium and mycelia were extracted together in order to collect the major number of compounds as possible. The production of two unprecedented compounds was observed, but preliminary analysis proved they did not belong to the **169** pathway. In particular, a 4 day-old culture grown in producing conditions (PDB, 28 °C; medium + mycelia sample 090118) yielded the new peak R5A, **295** (Table 3.14, Figures 3.36 and 3.38). Fermentation in M1D medium for a longer time (10-15 days) showed variability in the production of secondary metabolites, with one new peak R5B **296** in sample #1 (Table 3.14, Figures 3.37 and 3.38). Peaks related to compounds **272**, **273**, **289**, **290** and **291** (Figures 3.36 and 3.37), which are not in the WT, or seem enhanced in the mutant, were not considered of interest regarding **169** biosynthesis because of their nature. A more detailed discussion of the reasons for rejecting these compounds as **169** intermediates is presented in Section 3.4.1.

Compound	RT /min	UV$_{max}$ /nm	Mass /Da
R5A **295**	8.9	220, 296	290
R5B **296**	6.9	283, 356	336

Table 3.14: New peaks from Δ*sch*R5 mutant fermentation.

Compound R5A **295** had a mass of 290 Da (LCMS, Figure 3.38), which is 38 mass unit greater than the **169** mass of 252 Da. Compound R5B **296** instead presented a mass of 336 Da confirmed by ion pattern fragmentation in ES+ and ES- in the low resolution LCMS (Figure 3.38). Compound **296** is 84 mass units heavier than **169** and has a UV absorption rather different compared to compounds **169**, R7A **292** and R3A **294** (Figure 3.39). Because of these major differences in molecular weight and UV absorption, we did not push forward the analysis of **296**.

HRMS proved compound **295** to be unrelated to **169**, with a calculated molecular formula of $C_{15}H_{30}O_5$ (calculated [M]H$^+$ 291.2171, found 291.2167) and a total double bond equivalents of 1. Therefore, compound **295** is most likely an acyclic chain similar to fatty acids devoid of double bond system. Its complete HRMS data are hereby reported:

R5A (**295**): UV (diode array HPLC, H_2O/CH_3CN) λ_{max} = 220, 296 nm; LCMS m/z 291.3 [M]H$^+$, 313.3 [M]Na$^+$, 273.3 [M - H_2O]H$^+$; HRESIMS m/z 291.2167 [M]H$^+$ (calcd for $C_{15}H_{31}O_5{}^+$, 291.2171); 1 DBE.
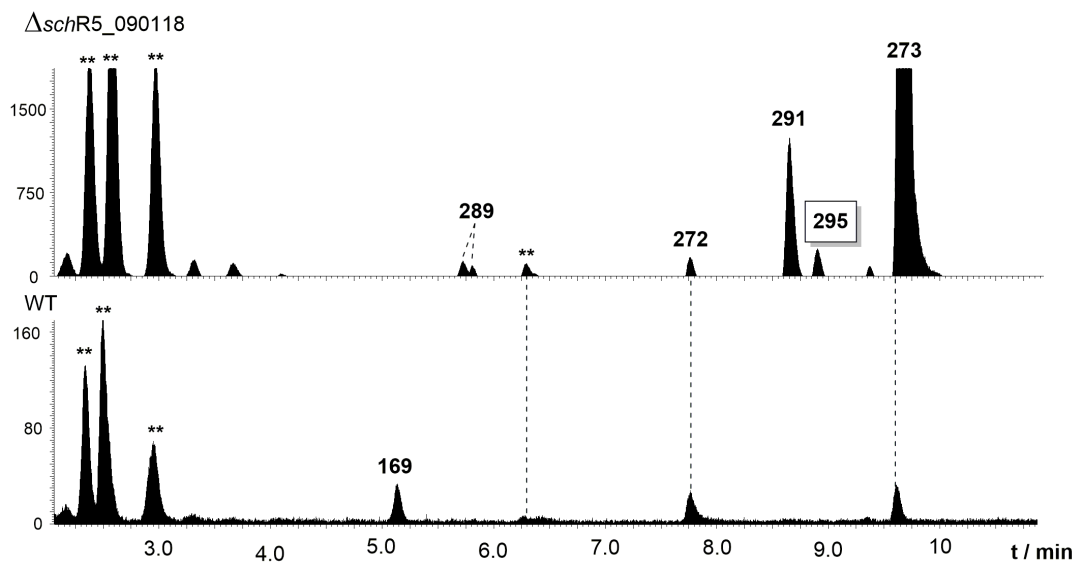
Figure 3.36: ELSD chromatograms of Δ*sch*R5 mutant (sample 090118 medium + mycelia) *versus* a wild type control grown in the same conditions. y axes are not linked. ** marks unrelated peaks.
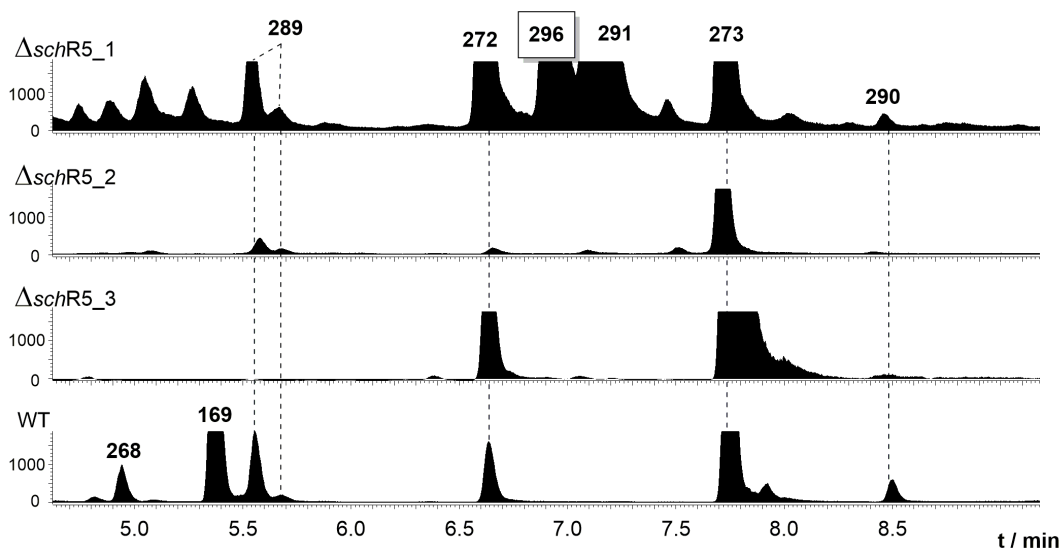
Figure 3.37: ELSD chromatograms of Δ*sch*R5 (samples #1, #2 and #3 medium + mycelia) fermented in M1D *versus* a wild type. y axes are linked.
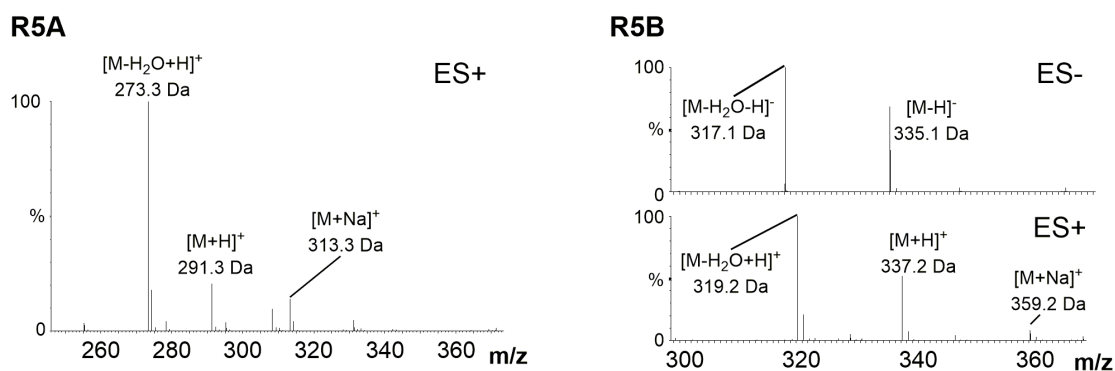
Figure 3.38: Mass spectra of R5A **295** and R5B **296**.
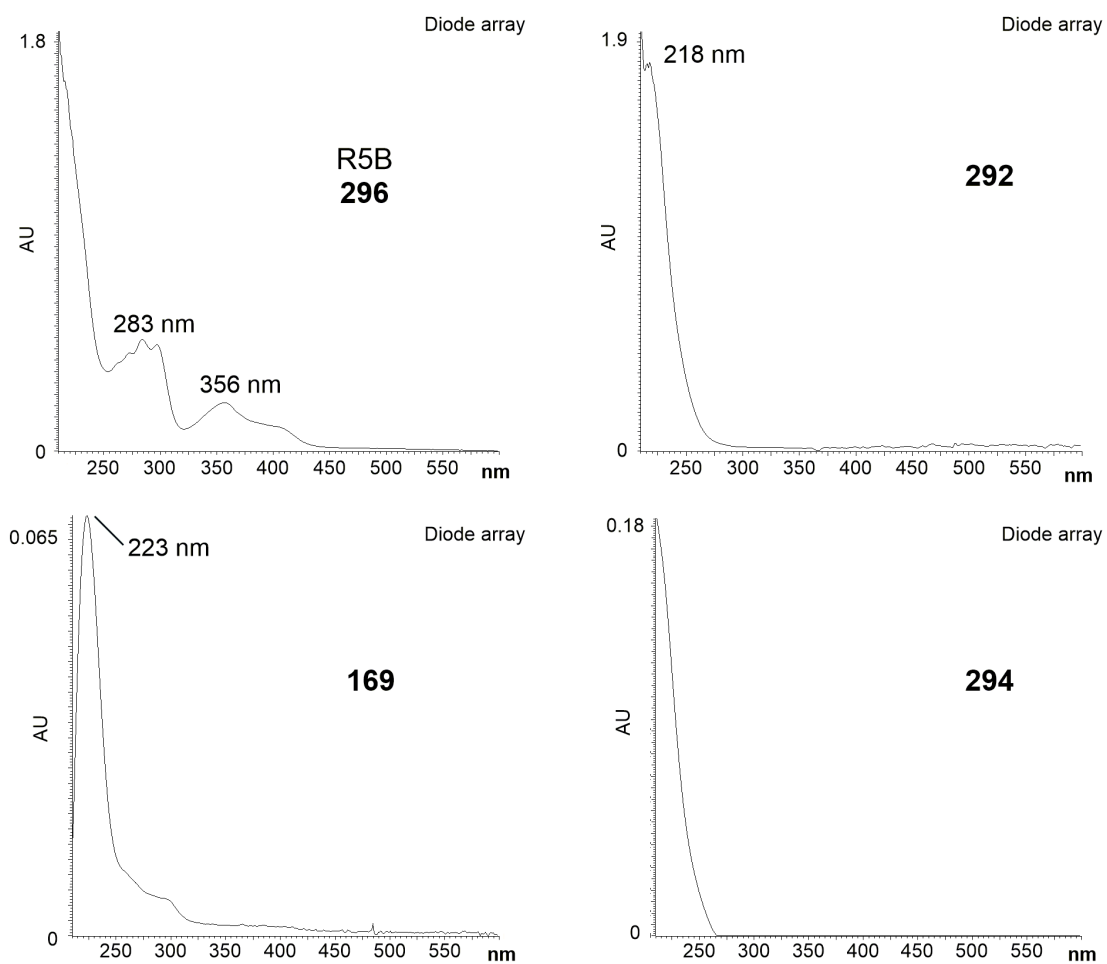
Figure 3.39: UV absorption of R5B **296** *vs* **169**, **169** R7A **292** and R3A **294**.

3.3.12 Δ*sch*R2 mutant fermentation

Two true Δ*sch*R2 (P450) mutants were obtained by bipartite transformation and confirmed by PCR. The two Δ*sch*R2 mutants did not lead to either **169** or **268** production (Figures 3.40 and 3.41), confirming SchR2 to be a P450 with a different function than SchR3. Analogously to mutant Δ*sch*R5, different media and time of growth were tested. Extractions were performed on mixed supernatant and blended mycelia in order to collect the most out of the cultures, but no new peaks were detected beside compounds **274**, **289**, **290** and **291**, which we observed during the years to be produced unpredictably in the WT and mutants (Figures 3.9, 3.22; Sections 3.3.1 3.3.9). Some peaks seemed to be enhanced in Δ*sch*R2, in particular dothiorelone A **289** and LMA-P3 **273** (Figures 3.40 and 3.41), suggesting that it might be related to the **169** pathway. However, we discarded such hypothesis because of the nature of **289**. A more detailed discussion of the reasons for rejecting compounds **273**, **274**, **289**, **290** and **291** is presented in Section 3.4.1.

During culturing of mutant Δ*sch*R2, production of compound LMA-P2 **269** was observed in the WT used as control. **269** was characterised by Ouazzani and collaborators (Figure 3.3, Section 3.1) and mass fragmentation and UV absorbance matched the literature (Figure 3.42).[192] We never observed its production in the past years until this moment.
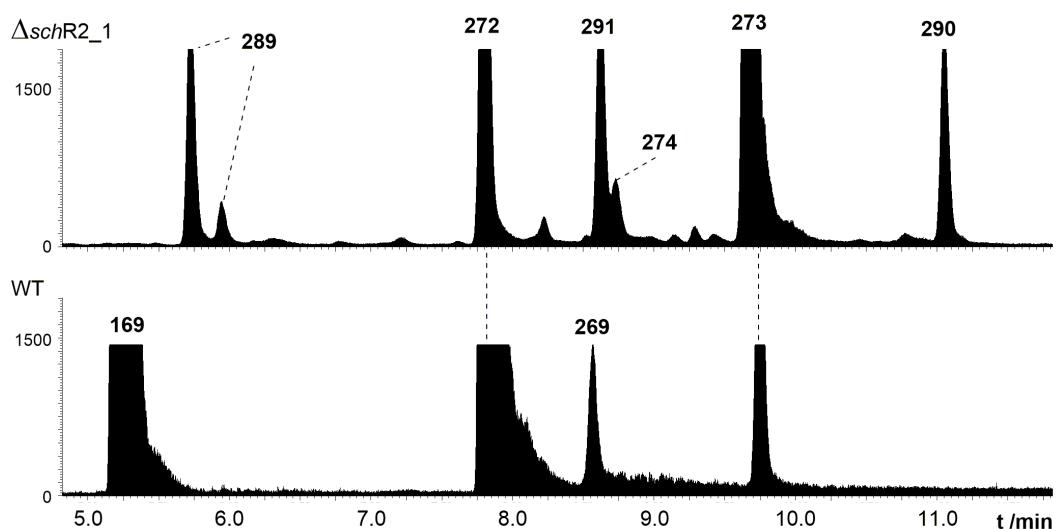
Figure 3.40: ELSD chromatograms of Δ*sch*R2 raw extract (medium + mycelia sample #1) *versus* a WT grown in the same conditions (PDB, 5 days, 28 °C). y axes are linked.
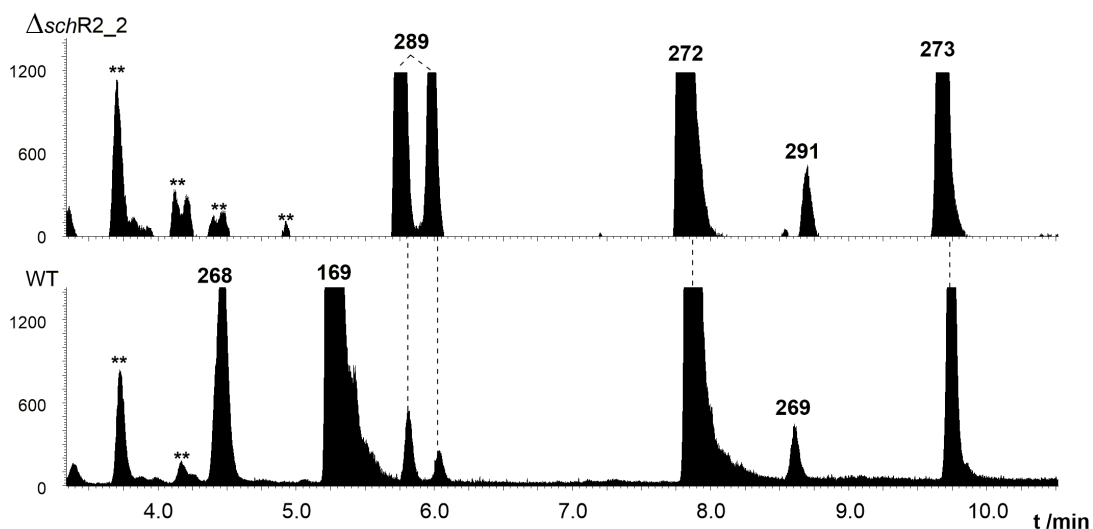
Figure 3.41: ELSD chromatograms of ΔschR2 raw extract (medium + mycelia, sample #2) *versus* a WT grown in matching conditions (M1D, 10 days, 28 °C). y axes are linked.
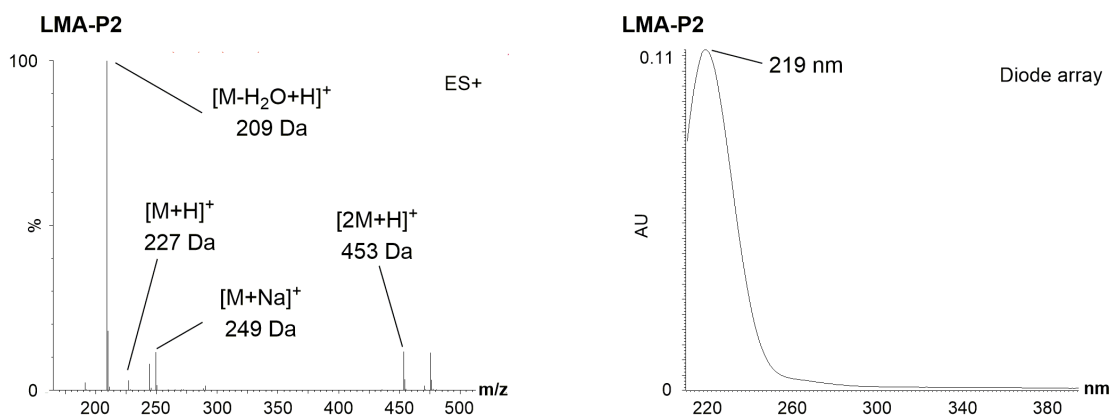
Figure 3.42: Low resolution MS fragmentation and diode array of compound **269**.

3.4 Discussion and Conclusions

3.4.1 Sch-642305 unrelated compounds

Besides Sch-642305 **169**, *Phomopsis CMU-LMA* produces many polyketides which we observed and/or are reported in the literature. It is useful to discuss which of these might be on the **169** biosynthetic pathway and which are not.

The fungus consistently synthesised compounds LMA-P3 **272** and DHTO **273** in all conditions of culturing (Figure 3.43).
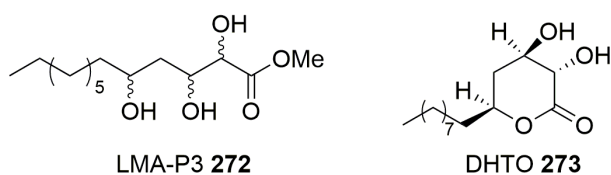
Figure 3.43: 169-unrelated compounds produced by *Phomopsis CMU-LMA*.

Compounds **272** and **273** are probably polyketides and their oxidation pattern suggests that they might be related to each other, although something unusual must happen during **272** formation. Both compounds have 14 carbons, but in the case of **272** an *O*-methylation occurs, leaving a total of 13 carbons in the main chain. An odd number of carbons in fungal polyketides is quite rare, as it suggests an unusual starting unit or a loss of a carbon. KO experiments proved **272** and **273** to be unrelated to **169** biosynthesis, as the production of both compounds was retained in Δ*sch*PKS and all other mutants, while **169** disappeared (Figure 3.21, Section 3.3.8). This was an important insight, especially because **273** and **169** posses the same number of carbons in the main chain.

A metabolic switch was observed after years of cultivation keeping the condition unaltered. Compounds LMA-P2 **269**, dothiorelone A **289**, phomolide C **274**, cytosporones B **290** and C **291** (Figure 3.44) were produced unpredictably and in different quantities depending on the fermentation (Figures 3.9, 3.22, 3.40, 3.41, Sections 3.3.1, 3.3.9, 3.3.12). This happened in conjunction with the creation of various mutants, making their analysis rather complicated.
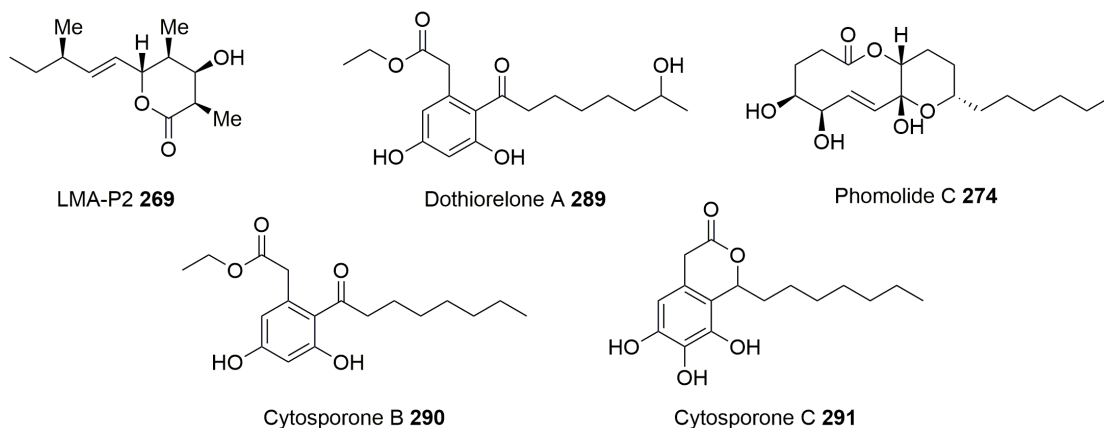
Figure 3.44: 169-unrelated compounds produced by *Phomopsis CMU-LMA*.

The observation that Dothiorelone A **289** had higher titres in mutant Δ*sch*R2 suggested

153

that it might have been a precursor in **169** biosynthesis. However **289** is an aromatic compound that is most likely derived by the action of a non reducing PKS probably related to dihydroxyphenyl acetic acid lactone (DAL)-type compounds.[311] Moreover **289** is an octketide, while **169** is a heptaketide. Studies on the brefeldin A system showed that *bref*-PKS is able to vary the chain length of its product,[301] suggesting that also *sch*PKS might generate an octaketide. However, the chain length control of *bref*-PKS was observed exclusively *in vitro* without the partnering thiolhydrolase, a scenario which might not be significant in a biological system. Moreover, as already mentioned, compounds **289**, **290** and **291** had an unpredictable production and varied their titres from one fermentation to the other. Compounds **289**, **290** and **291** seemed to be related to each other and they might belong to the same pathway. The similarity between **289** and **290** is obvious, one being the hydroxylated form of the other. Both may be ethylated shunt products on the pathway of the DAL **291**: a free carboxylic acid might cyclise to form the bicyclic system and an extra hydroxylation on the aromatic ring would yield **291**. The biosynthesis of DAL-related compounds involves the synthesis of highly reduced starter units by either hr-PKS or FAS, as the case of norsolinic acid **19**[312,313], and direct transfer to the starter domain (SAT) of a non-reducing PKS for further chain extension, aromatic cyclisation and release (Scheme 3.8). In this case, a highly reduced tetraketide would be extended by four acetate units by a non-reducing PKS (**297** to **298**), following the formation of the aromatic system (**299** to **300**), lactonisation and P450-mediated hydroxylation (**300** to **289**, **290** and **291**, Scheme 3.8). We searched the *Phomopsis CMU-LMA* genome for possible DAL biosynthetic clusters expected to contain either a FAS and a nr-PKS, or a highly reducing and non reducing PKS pair, but this type of BGC could not be found. However, given the low quality of the genome sequence these results should be viewed as tentative.
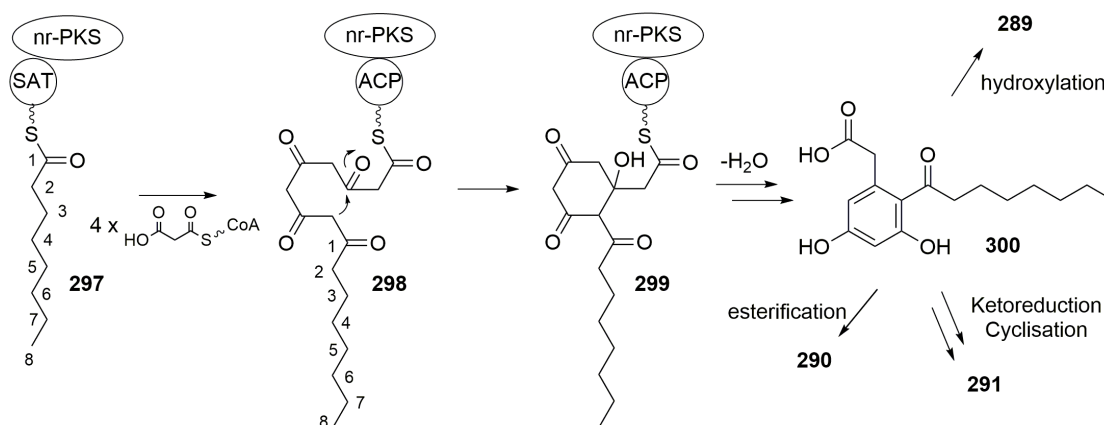
Scheme 3.8: Speculative biosynthesis of compounds **289**, **290** and **291**.

Compound LMA-P2 **269** was observed after three years of culturing in PDB in the WT and in mutant Δ*sch*R2 (Figures 3.40 and 3.41, Section 3.3.12). **269** is most likely not an

intermediate of **169** because of its methylation pattern and chain length: *sch*PKS has an inactive *C*-MeT domain, but **269** is probably a pentaketide derived from a hr-PKS with an intact *C*-MeT domain.

Phomolide C **274** is a decalactone fused to a tetrahydropyran. It has some structural analogies to **169** although it is probably not related to the latter because **274** is a nonaketide, while **169** is a heptaketide and they have different oxidation patterns. It is not to be excluded that a different programming of the *sch*PKS might extend the chain for two units more, in analogy to the brefeldin A system. However, the different oxidation pattern suggests that the two compounds derive from different systems. Production of **274** was not observed in any Δ*sch*PKS mutants, so we cannot exclude with certainty that they are unrelated. However, even if the hypothesis that they might derive from the same hr-PKS system is plausible and fascinating, they are the result of different processing and are not one the precursor of the other.

Assuming the speculation to be sensible, we can focus on the compounds that are most likely to be part of the **169** pathway.

3.4.2 Sch-642305

Sch-642305 **169** was produced in every condition of culturing, independently from the media, shaking or temperature. Recorded NMR data of purified **169** perfectly matched literature values, and feeding [1-^{13}C]-acetate confirmed **169** to be a highly reduced heptaketide, with a consistent alternating pattern of incorporation of the label (Section 3.3.3). Based on the feeding experiments we could also exclude atom rearrangement during the formation of the 6-membered ring. The oxidation pattern of **169** suggested post-elongation oxidative events, especially regarding the hydroxy group at position 4, which is probably derived from atmospheric oxygen. Also the C-2/C-3 *Z* double bond is not created by the DH domain of the PKS, as it lays within the atoms of an acetate unit, suggesting a different mechanism. Moreover, DH domains typically form *E* double bonds,[314,315] whereas **169** presents *Z* stereochemistry, therefore this double bond could be formed oxidatively. The absence of SAM-derived methyl groups on the polyketide chain suggested an inactive MeT domain of the core PKS.

3.4.3 Bioinformatics

The poor quality of the genome assembly was attributed to the relatively old raw data utilised or to a low quality DNA preparation prior to sequencing. The genome had an average scaffold length of ∼30 KBp, quite a limiting size, considering that fungal BGC normally span ∼50 KBp. Nevertheless, the automatic AntiSMASH annotation was performed and led to the best BGC candidate *sch* cluster (31.2 KBp long, on a scaffold of 76.2 KBp), which was later proved by KO experiments (Sections 3.3.8 - 3.3.12). AntiSMASH annotation gave as output a total of 156 clusters associated with secondary

metabolites, meaning that *Phomopsis CMU-LMA* could potentially produce over 150 natural products including polyketides, terpenes and NRP (Table 3.3, Section 3.3.4). However, considering the low assembly quality of the genome, we regarded this number to be an overestimation of the real total: a BGC split into two separate scaffolds may be considered as two separate hits by the software, unbalancing the count. Even if AntiSMASH might have overestimated the real number of BGC, Reaxys search showed over 330 compounds isolated from *Phomopsis spp.*, counting these organisms as prolific producers of natural products. Of the 150 predicted clusters 60 were related to type I PKS metabolism. Domain analysis of these 60 PKS revealed 36 highly reducing PKS, with a complete set of KR, DH and ER domains and, among these, only 4 PKS with a non-functional MeT domain. In particular, one of the four hrPKS had as closest hit Bref-PKS, the polyketide synthase responsible for the biosynthesis of brefeldin A **275** in *P. brefeldianum.* The high homology (56% identity, Table 3.6, Section 3.3.4) was not limited to the two hr-PKS, but it extended significantly between the whole *sch* and *Bref* cluster (Figure 3.15, Section 3.3.4). Given the structural analogies between **169** and **275** we speculated that their biosynthesis might be related, therefore also the set of enzymes that take part in their assembly. This defined cluster *sch* as the best candidate for **169** biosynthesis, and we focused on it for further investigation.

The SchPKS sequence was used to find a *sch*-homologous cluster (namley PVSCH BGC) in *P. verrucosum*, another **169** producer (Table 3.5, Section 3.3.4).[282] The comparison showed high homology between the core PKS, P450 oxidases and the *trans*-acting thiolesterase, necessary for chain release. In particular, the two P450 enzymes SchR2 and SchR3 showed high homology between one P450 in *P. verrucosum* and three P450 oxidases in *P. brefeldianum*, suggesting that one P450 is enough to obtain the final product **169**. Later KO experiments refuted this hypothesis, as both $\Delta schR2$ and $\Delta schR3$ mutants could not produce **169**, meaning that both SchR2 and SchR3 are essential for **169** biogenesis. What happens in *P. verrucosum* might be different, as this result suggested that the two organisms may have evolved parallel strategies to synthesise **169**. It is possible that in *P. verrucosum* the P450 PVSCHR3 may be iterative, as in the case of in chaetoglobosin A **301** biosynthesis, elucidated by Watanabe *et al.*[316] In particular, P450 CHGG_01243 (highlighted in Scheme 3.9) performs a double hydroxylation on the two adjacent C-19 and C-20 (**303** to **304** or **302** to **305**). Epoxidation and FAD-dependent oxidation of the C-20 alcohol to ketone finally yields **301** (Scheme 3.9). Nonetheless, in *P. verrucosum* Sch-642305 (**169**) biosynthesis, the P450 PVSCHR3 should insert two hydroxylations on two non-adjacent carbons, differing from the chaetoglobosin A case. It is possible that PVSCHR3 P450 might accept benquoine **260** and R3A **294** stepwise to perform the double hydroxylation. More investigation of the oxidative tailoring events in **169** biosynthesis might be performed in *P. verrucosum*.
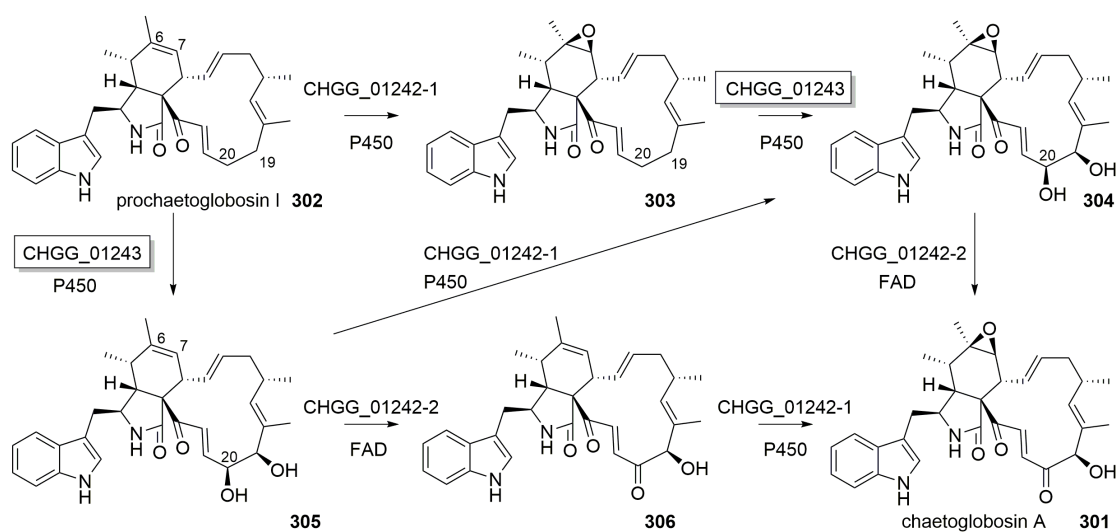
Scheme 3.9: Post-assembly oxidative tailoring steps in **301** biosynthesis.[316]

3.4.4 *Phomopsis CMU-LMA* mutants

To validate the putative *sch* BGC, knock out experiments by bipartite transformation were designed. *Phomopsis CMU-LMA* transformation has never been described before, but luckily, a general *A. oryzae* transformation protocol was successful and we could generate healthy and numerous protoplasts. Hygromycin B resistance was effective in the organism, setting the basis for the KO experiments.

The target genes were chosen according to their proposed function. The thiolesterase *sch*R1 was not taken in consideration as it is probably involved in the chain release analogously to the brefeldin A case.[301] More interesting were the core *sch*PKS, the cytochromes *sch*R2 and *sch*R3, the NAD dependent *sch*R5 oxidoreductase and the FAD dependent *sch*R7 oxidase. Attempts to create a construct aimed against the serine hydrolase *sch*R4 were performed but were not successful.

The Δ*sch*PKS mutant conclusively demonstrated the role of *sch* BGC in the biosynthesis of **169**, with total disappearance of **169** and the minor compound **268**. Mutants Δ*sch*R3 and Δ*sch*R7 produced the two compounds R3A **294** and R7A **292** respectively when fermented in large scale (>1 L), with disappearance of **169**. There is no way of knowing whether **294** and **292** are real intermediates or shunt products. True mutants Δ*sch*R2 and Δ*sch*R5 had **169** pathway interrupted, but did not produce any intermediate even in high volume fermentation nor in different conditions, suggesting that their substrates are rapidly shunted by the fungus.

Mutant Δ*sch*R3 gave a surprising result, as it showed that the SchR3 P450 is responsible for hydroxylation at C-7. This means that the hydroxy group on C-4 is probably introduced by SchR2 P450. Our results suggest that compound **292** is the substrate

for the FAD dependent oxidase SchR7. Researches in Protein Data Bank (PDB)[317,318] relative to SchR7 protein gave as first hits the flavoprotein 6-hydroxy-D-nicotine oxidase of the Gram-positive actinomycete *Paenarthrobacter nicotinovorans* (23% identity), and the ascomycete *Neurospora crassa* with even more identity to R7A (28%), a known enzyme that converts hydroxy groups to ketones. Although the rather low identity with the nicotine oxidase, we propose that SchR7 might oxidise the OH-7 of R7A to ketone as we know that such functional group is present in the final product (**292** to **307**, Scheme 3.10).
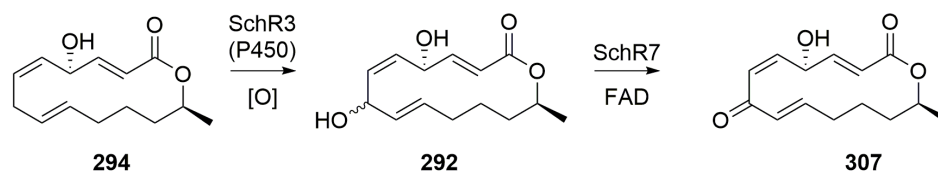
Scheme 3.10: SchR3 hydroxylation on C-7 and hypothetical oxidation to ketone by SchR7.

It is important to notice that none of the mutants could produce compound **169** nor **268**, showing that all targeted genes (*sch*PKS, *sch*R2, *sch*R3, *sch*R5 and *sch*R6) are essential in **169** biosynthesis and that **169** and **268** are products of the same pathway. This is particularly interesting regarding *sch*R2 and *sch*R3 P450-encoding genes, as bioinformatical analysis suggested that they could have been the same gene duplicated in the course of evolution (Figure 3.15, Section 3.3.4). When compared to *Bref* and *PVSCH* clusters, *sch*R2 and *sch*R3 resulted homologous to the single *PVSCH*-L3 P450 and to the three P450s in *P. brefeldianum orf*3, *orf*4 and *orf*6 (Figure 3.15, Section 3.3.4). The fact that both Δ*sch*R2 and Δ*sch*R3 could not produce **169** means that SchR2 and SchR3 cannot complement one another, thus they must perform different chemical steps. We propose that SchR2 P450 uses benquoine **260** as substrate to introduce the **169** C-4-hydroxylation early in the biosynthesis, prior to the formation of the 6-membered ring, although we could not see benquoine produced by either *Phomopsis CU-LMA* WT, nor Δ*sch*R2 neither in the media, nor in the mycelia (Section 3.3.12).

3.4.5 Sch-642305 and brefeldin A

During the investigation of **169**, we realised that SchPKS is closely related to Bref-PKS, therefore we assumed that compounds **169** and **275** could share some analogies in their assembly (Figure 3.45). The biosynthesis of **275** is not yet completely understood, beside the origin of the carbon and oxygen atoms, and the chain-length control feature mediated by the *trans*-partnering Bref-TH. A homologous *trans*-thiohydrolase was also found in *sch* BGC (gene *sch*R1) with 66% identity at amino acid level (Figure 3.15, Table 3.6).
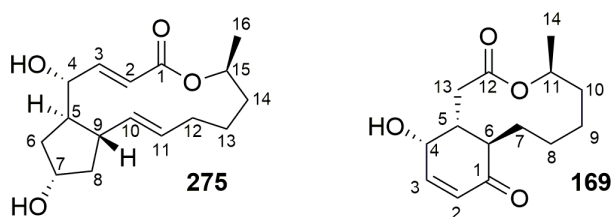
Figure 3.45: Brefeldin A and Sch-642305.

It is interesting to notice that both compounds have the similar post-elongation P450-mediated oxidation at C-7 (**275**) and C-1 (**169**) respectively. Feeding **275** precursors, with the cyclopentane in place, proved the C-7 oxidation to happen after the 5-membered ring formation (Scheme 3.3, section 3.1),[300] while in **169** the C-1 hydroxylation happens before the 6-membered ring arrangement, followed by oxidation to ketone. This suggests a different mechanism of ring formation in the two systems, with the C-1 keto group of **169** possibly involved in the cyclisation, while the OH-7 of **275** takes no part in the process.

Noteworthy is the major compound **277** yielded by expressing Bref-PKS and Bref-TH *in vitro* and in yeast.[301] We can assume **277** to be the real precursor to **275** and compare it to the macrolactone benquoine **260**, which we consider to be on the route to **169**, basing on Ouazzani's proposal (Figure 3.46).
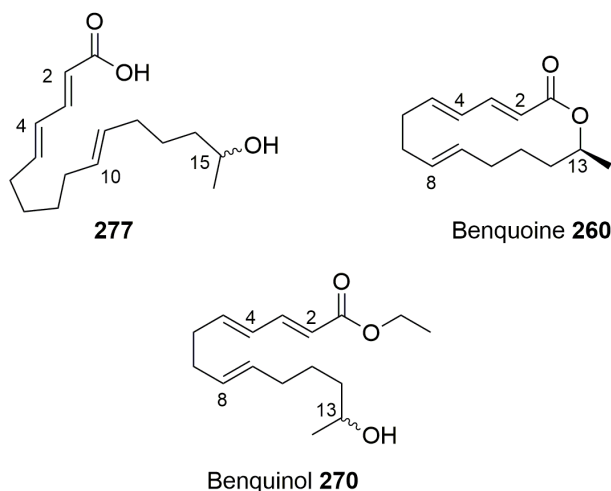
Figure 3.46: Precursor of brefeldin A **277**, and putative precursors of Sch-642305 **260**. **270** is, most likely, a shunt product.

Probably SchPKS and SchR1 produce a linear heptaketide that cyclise to yield **260**. This hypothesis is supported by the isolation of benquinol **270** from *Phomopsis CMU-LMA*.[192] Compound **270** is the ethyl ester of **260** and they match both for chain length

and β-reduction pattern, with **270** presenting the C-13 hydroxy group involved in the macrolactonisation. Another clue that SchPKS and SchR1 might synthetise an acyclic compound comes from the brefeldin A system: **277** is the early product of Bref-PKS and Bref-TH partnering, and it is in fact a linear product, which is probably cyclised after chain release by a dedicated enzyme. KO evidences in *Phomopsis CMU-LMA* showed that the macrolactonisation of **169** happens prior the 6-member ring formation, as ΔschR3 and ΔschR7 mutants yielded compounds **294** and **292** with the lactone in place. The *sch* cluster contains two hydroxylases-encoding genes: *sch*R1 and *sch*R4; the first most likely involved in the release of the chain, and the second possibly involved in the intramolecular condensation. The *Bref* BGC lacks a *sch*R4 homologous, so the macrolactonisation mechanism may differ in the two systems. Tang *et al* proposed the **275** lactone to be formed by the action of Bref-TH upon chain release, but the yield of the acyclic **277** *in vitro* and *in vivo* suggested that the 5-membered ring could be formed before lactonisation (Scheme 3.6 section 3.1). This means that a P450 in *Bref* BCG should act on the growing polyketide prior to chain release, but, since there are no known examples of tailoring happening at the same time as PKS-elongation, we rather assume that **275** lactonisation happens after chain offload.

It is intereting to notice that the β-reduction pattern of the early products of Bref-PKS and SchPKS are related to each other. Both **277** and **260/270** present a *2E, 4E* conjugated double bond system, and a third *E* olefin on C-8 (**260/270**) and C-10 (**277**). The olefin pattern might play a fundamental role in the 6/5-membered ring closing, especially the conjugated system as it is identical in both system. Interestingly, the 2*E* double bond is conserved in **275**, while it is reduced in **169**. This may be related to the fact that **275** forms the 5-membered ring between C-5 and C-9 employing the C-4 olefin, whereas **169** cyclise C-3 and C-8.

The NAD- and FAD-dependent *sch*R5 and *sch*R7 genes have no homolog in the *Bref* cluster, meaning that their products perform oxidoreductions unique to the **169** pathway, with their product possibly involved in the formation of the 6-membered ring and oxidation of the OH-7 to ketone. Unfortunately we were not able to characterise the role of these two enzymes by knockout.

3.4.6 Biosynthetic proposal

A biosynthetic pathway can be proposed considering the literature isolated compounds (Figure 3.47). In particular benquinol **270**, DHTTA **271**, benquoine **260** and LMA-P1 **268** seem to belong to the **169** pathway because they posses the same number of carbons and a similar double bond system. Another evidence is the C-13 hydroxylation involved in the macrolactonisation, which is conserved among the mentioned compounds. In particular, compound **268** has the same bicyclic system as **169**. In Ouazzani's proposal, **268** arises from a reduction of the **169** olefin and keto group to alcohol, and the isolation of compounds **294** and **292** from mutants ΔschR3 and ΔschR7 seems to confirm this hypothesis, as they posses the C-5/C-6 *Z* olefin found also in **169**.
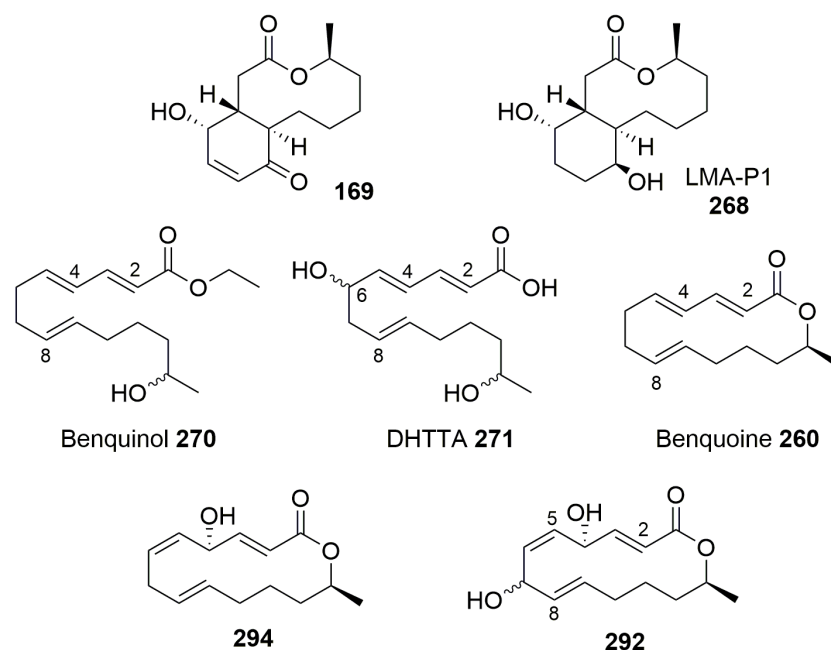
Figure 3.47: Isolated compounds from *Phomopsis CMU-LMA*[192] and from KO mutants (**292**, **294**).

It is interesting to notice how the triene $2E, 4E, 8E$ configuration of **260**, **270** and **271** is conserved, in contrast with the $2E, 5Z, 8E$ configuration of R7A and R3A. Also the position of one double bond differs: R3A and R7A have an olefin at C-5 and C-6 while in **260**, **271** and **270** the olefin is between C-4 and C-5. It is worth mentioning that the $5Z$ double bond is the same between **169** and R3A/R7A. Another difference between R7A and **260**, **271** and **270** lies in the oxygenation pattern: R7A has the -OH groups in the same position as **169** and **268** (C-4 and C-7), while DHTTA has OH-6.

Probably compound **270** is a shunt derivative of **260** or its acyclic form: isolation in acidic condition may open the ring and transesterification in ethyl acetate extraction may lead to the formation of the ethylester. The same applies to DHTTA, that may arise from a shunt C-6 hydroxylation of **270**. The **169** oxydation at C-4 and C-7 is probably inserted by P450 cytochromes, with SchR3 acting on C-7 and SchR2 probably resulting in C-4 hydroxylation, even if comparison with *P. verrucosum* cluster shows only one P450 necessary to synthetise **169**. Nevertheless, what happens in *P. verrucosum* may be different from *Phomopsis CMU-LMA*.

We propose that benquoine is the product of the highly reducing SchPKS partnering with the *trans*-thiohydrolase SchR1 in analogy with the brefeldin A system. SchPKS leaves in place the $2E, 4E, 8E$ triene. Macrolactonisation might happen upon chain release, or, more likely, might take place enzymatically post chain offload, as evidences in **275** biosynthesis suggest (**308** to **260**, Scheme 3.11). The macrolactonisation might

be catalysed by the hydrolase SchR4. Cytochrome ScyR2 probably acts on the 4*E* olefin inserting an epoxide (**260** to **309**), which can open to yield the isolated R3A intermediate with consequent double bond migration (**309** to **294**). Subsequent action of the SchR3 P450 inserts the hydroxyl on C-7 to yield the observed intermediate **292**. The flavoprotein oxidase might convert the C-7 hydroxyl to the **307** ketone and the oxidorecuctase SchR5 might perform reduction onto the C-8/C-9 olefin (**307** to **310**), to form the enolate that could perform a Michael-type cyclisation to yield the 6-membered ring (**310** to **169**). Reductive events, probably controlled by enzymes encoded outside the boundaries of the *sch* cluster, as there are no other candidate genes within the BGC, may lead to the formation of **268**.
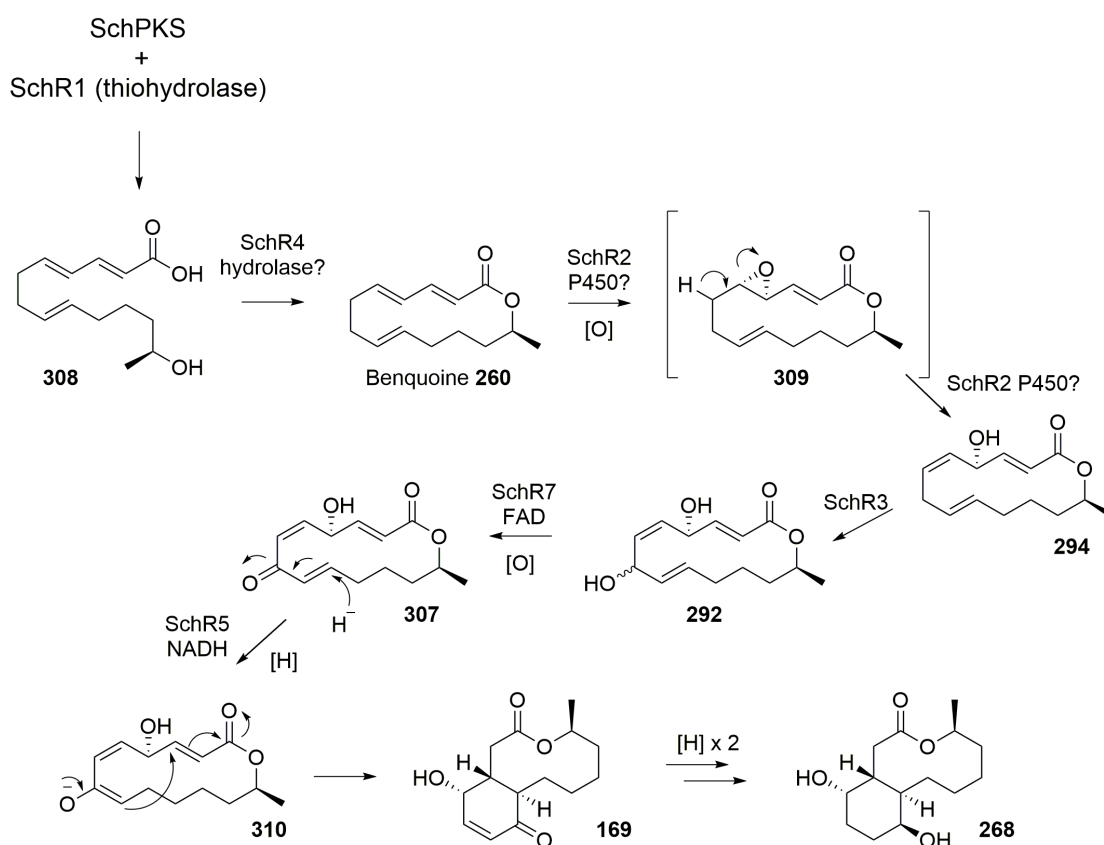
Scheme 3.11: Proposed **169** biosynthesis *via* **260**.

4 General conclusion and outlook

The unprecedented structures of phyllostictines that originally captured our interested turned out to be not the result of a new metabolic strategy in fungi, but a human error. Our investigation led to the revision of the literature structures (**168**, **173** - **175**) to a series of bicyclic 3-methylene tetramic acids (**190**, **181**, **200**, **201**) synthetised by a PKS-NRPS system. Moreover, ITS analysis of the producer revealed the fungus to be closely related to *Phaeosphaeride spp.* rather than *Phyllosticta spp.* We did not characterise the organism precisely, but deeper philogenetic analysis should be performed to classify and rename *P. cirsii.* Sequencing and rational analysis of the *P. cirsii* genome pointed at a candidate cluster (*phy* BGC), likely involved in phyllostictine biosynthesis. KO experiments proved *phy* BGC as mutant ΔphyS (PKS-NRPS) could not produce phyllostictine A, nor any related compounds. Disruption of a P450 encoding-gene (*phy*L6), not only interrupted phyllostictine production, but also generated the new compound phyllostictine E (**199**). Alignment of the release domain of the PKS-NPS with other tetramic acid-producing systems showed that *phy*S has a single mutation in the NADPH binding site GXXGXXG, resulting in a release mechanism by Dieckmann cyclisation. *In vitro* study of the release mechanism could be performed in the future by directed mutagenesis of the NADPH binding site in order to reform the cofactor binding site to observe whether the reductive release is restored. If functional, the mutated PKS-NRPS could be reintroduced in the ΔphyS mutant to observe the possible generation of new compounds *in vivo*. Fermentation of *P. cirsii* in PDB yielded the shikimate-derived scytolide **186**, future work forecasts the confirmation by KO of its putative BGC (Table 5.2 Section 5.5.3), which details and discussion of its founding has not been included in the thesis.

The investigation of Sch-642305 (**169**) revealed the biosynthetic gene cluster in *Phomopsis CMU-LMA* (*sch* BGC), later confirmed by KO of the core hr-PKS and four tailoring genes. Disruption of two P450 encoding genes lead to the isolation of two new macro-lactones **292** and **294**, with similar olefin pattern to brefeldin A system. High homology between the *sch*PKS and *Bref*-PKS (also extended between the two clusters) suggested that the two systems share many analogies. We also found a homologous genes cluster in the **169**-producer fungus *P. verrucosum* (*PVSCH BGC*), which interestingly has a different arsenal of enzymes, suggesting a different biosynthetic pathway. It would be interesting to heterologously express combinations of *sch PVSCH* genes and observe whether they could compensate each other or yield new products. We were not able to prove the mechanism of the 6-membered ring formation of **169**, but the identification of *sch* BGC should allow a better understanding of the chemical steps involved in the biosynthesis.

Our results show that new molecules such as compounds **199**, **292** and **294** can be generated by KO from preexisting pathway. Although KO is an effective approach to link a BGC to a natural compound, it is sometimes not efficient, depending on the the species of the fungus to transform. The problem of transformation efficiency may be avoided by changing the producer organism. Moreover, often interesting intermediates are rapidly shunted or completely degraded by the cells, rendering their analysis sometimes complex. A way to bypass such eventuality is by heterologous expression of the whole or partial BGC in an expression system such as yeast, *A. oryzae* or *A. nidulans*. For example, KO of genes *sch*R2 and *sch*R5 in *Phomopsis CMU-LMA* did not yield to any new compound, although we know they are fundamental in **169** biosynthesis. In order to unravel their function, heterologous expression might be the key, or *in vitro* assay of purified SchR7 fed with **292** might reveal if this enzyme is responsible for the formation of the 6-membered ring.

5 Experimental

5.1 Fermentation

P. cirsii was grown in static M1D medium for 14 - 20 days at 28 °C to yield phyllostictine A. During the fermentation the fungus was kept in the dark, although it is not a required condition to activate secondary metabolism. Culturing in shaken PDB medium at 28 °C for 6-8 days showed different profile of secondary metabolites (scytolide **186**). *Phomopsis CMU-LMA* was principally grown in PDB medium for 5 - 10 days at 28 °C in static or shaken condition (100 - 120 rpm).

Mycelia were grew on agar from preexisting solid culture or from glycerol stocks. To inoculate a liquid culture 1 mL ddH$_2$O was used to scrape the surface of the agar plate and ∼200 μL of mycelia/spore solution injected into 100 mL of liquid medium in 500 mL flasks. Liquid and solid culturing was handled under biological hood using a flame.

5.2 Media composition

M1D medium [Ca(NO$_3$)$_2$ 1.2 mM; KNO$_3$ 0.79 mM; KCl 0.87 mM; MgSO$_4$ 3 mM; NaH$_2$PO$_4$ 0.14 mM; sucrose 87.6 mM; ammonium tartrate 27.1 mM; FeCl$_3$ 7.4 μM; MnSO$_4$ 30 μM; ZnSO$_4$ 8.7 μM; H$_3$BO$_3$ 22 μM; KI 4.5 μM; pH 5.5].

PDB [potato infusion 200 g · L^{-1}; dextrose 20 g · L^{-1}; pH: 5.1].

Czapek Dox with sorbitol, soft agar (CD+S soft agar) [Czapek Dox broth 3.5%; D-sorbitol 18.22%; Agar 0.8%].

Czapek Dox with sorbitol (CD+S) [Czapek Dox broth 3.5%; D-sorbitol 18.22%; Agar 1.5%].

CDZ/S agar [3.5% Czapek Dox broth; 1 M sorbitol; 0.05% adenine; 0.15% Methionine; 0.1% (NH$_3$)$_2$SO$_4$; 0.8% agar].

DPY [dextrin from potato starch 2%; polypeptone 1%; yeast extract 0.5%; KH$_2$PO$_4$ 0.5; MgSO$_4$ 0.05%; (agar 2.5%)].

LB agar [yeast extract 0.5%; tryptone 1%; NaCl 0.5%; agar 1.5%].

Malt extract (ME) [malt extract 1.28%; peptone *ex* soya 0.08%; glycerol 0.24%; dextrin from potato starch 0.28%; (agar 1.5%)].

Supplement mixture minus uracil (SM-URA) [yeast nitrogen base 0.17%; $(NH_4)_2SO_4$ 0.5%; D(+)-glucose monohydrate 2%; complete supplement mixture minus uracil 0.077%; agar 1.5%].

YPAD [yeast extract 1%; tryptone 2%; D(+)-glucose monohydrate 2%; adenine 0.03%].

5.3 Natural products extraction

The fermentation was homogenized, the cellular debris removed by paper filtration under vacuum. The supernatant was extracted twice with 1.5 volumes of ethyl acetate. The organic extract was dried ($MgSO_4$) and evaporated *in vacuo*. The crude extract was either stored at -20 °C, or dissolved in acetonitrile or methanol (10 mg $\cdot$ mL^{-1}) for LCMS analysis and purification.

In case of mycelia extraction, the cells were collected by filtration and stirred in acetone or ethyl acetate for 1-2 h. The mycelia were removed by filtration and the acetone/EtAc evaporated *in vacuo*. The crude extract was either stored at -20 °C, or dissolved in acetonitrile or methanol (10 mg $\cdot$ mL^{-1}) for LCMS analysis and purification.

Solid extraction from agar plate was performed as described for mycelia extraction, for homogenised agar.

5.4 Analytical

5.4.1 Analytical LCMS

LCMS data were obtained with using a Waters LCMS system comprising of a Waters 2767 autosampler, Waters 2545 pump system, a Phenomenex Kinetex column (2.6 μ, C-18, 100 Å, 4.6 × 100 mm) equipped with a Phenomenex Security Guard precolumn (Luna C-5 300 Å) eluted at 1 mL$\cdot$min^{-1}. Detection was by Waters 2998 Diode Array detector between 200 and 600 nm; Waters 2424 ELSD and Waters Single Quadrupole detector SQD-2 mass detector operating simultaneously in ES+ and ES- modes between 100 m/z and 650 m/z. Solvents were: A, HPLC grade H_2O containing 0.05% formic acid; and B, HPLC grade acetonitrile containing 0.045% formic acid. Gradients were as follows. Method 1. Kinetex/CH_3CN: 0 min, 10% B; 10 min, 90% B; 12 min, 90% B; 13 min, 10% B; 15 min, 10% B.

5.4.2 Semi-Preparative LCMS

Purification of compounds was generally achieved using a Waters mass-directed autopurification system comprising of a Waters 2767 autosampler, Waters 2545 pump system, a Phenomenex Kinetex Axia column (5μ, C-18, 100 Å, 21.2×250 mm) equipped with a Phenomenex Security Guard precolumn (Luna C-5 300 Å) eluted at 20 mL·min^{-1} at ambient temperature. Solvent A, HPLC grade H_2O + 0.05% formic acid; Solvent B, HPLC grade CH_3CN + 0.045% formic acid. The post-column flow was split (100:1) and the minority flow was made up with HPLC grade MeOH + 0.045% formic acid to 1 mL·min^{-1} for simultaneous analysis by diode array (Waters 2998), evaporative light scattering (Waters 2424) and ESI mass spectrometry in positive and negative modes (Waters SQD-2). Detected peaks were collected into glass test tubes. Combined tubes were evaporated by vacuum concentrator (SPEEDVAC) or freeze dried, weighed, and residues dissolved directly in NMR solvent for NMR analysis.

5.4.3 NMR Analysis

The ^{1}H-NMR analysis was performed using Brucker DPX 200, Avance 400, DPX 400 and DRX 500 spectrometers. Resonances were assigned using two dimensional NMR ^{1}H,^{1}H-COSY, ^{1}H, ^{13}C-HSQC and ^{1}H, ^{13}C-HMBC experiments. Deuterated DMSO (ref. 2.50 ppm / 39.5 ppm) or $CDCl_3$ (ref 7.26 ppm / 77.4 ppm) were used as solvents and reference. The ^{13}C-NMR analysis was performed using Brucker Avance 400, DPX 400 and DRX 500 spectrometers. NMR data was processed using Topspin and MestReNova software packages.

5.4.4 DNA Gel Electrophoresis

DNA samples such as PCR products, plasmids and gDNA were evaluated by agarose gel electrophoresis. Depending on the samples the agarose concentration was 0.8 - 2.0% (w/v) in 1x TAE buffer (Tris acetate 200 mM; EDTA 50 mM; pH 8.3). Large DNA such as gDNA required lower agarose concentration, while smaller DNA fragments (1 - 5 KBp) had a better separation at higher agarose concentration. Agarose was molten in 1x TAE buffer and gel stain (2 μL/25 mL, Roti-GelStain, Roth) was added and poured on a tray with comb fixed in a gel caster system (Bio-Rad). After solidifying of the agarose, combs were removed and the tray transferred to a Sub-Cell GT agarose gel electrophoresis system (Bio-Rad). 6x DNA loading dye was added to samples to a final concentration of 1x and loaded in the gel wells. 1 μL of 1 KBp DNA ladder was loaded for size comparison and the gel was run for 30-40 min at 100-120 V using 1x TAE as running buffer. DNA bands were analysed under UV light with a gel documentation system (Gel Doc XR+, Bio-Rad) and the software Image Lab (Bio-Rad).

5.5 Bioinformatics

5.5.1 Fungal gDNA isolation and sequencing

To obtain fungal gDNA, cells were grown in 200 mL DPY medium (Section 5.2) for 5 days. The liquid medium was removed and mycelia were freeze-dried to remove all the water. Liquid nitrogen was used to grind the freeze-dried cells with mortar and pestle. gDNA purification was performed using the Kit GenElute Plant Genomic DNA miniPrep (G2N350, Sigma-Aldrich). Genomic DNA was sequenced on the MiSeq system (Illumina) in a paired-end sequencing run. Raw data was processed by an in house software platform (GenDBE). After assembly of all sequence reads by applying the GS/De Novo/Assembler version 2.8 software with default settings. The annotation of the draft genome was made within the GenDBE platform including AntiSMASH, SoftBerry FGENESH and AUGUSTUS 3.0.3 for gene prediction. Gene cluster analysis was performed using browser NCBI, CDD, Clustal Omega, Artemis and Artemis comparison tool.

5.5.2 KO cassettes designing

Design of primers, plasmids and sequences management was perfermed with Geneious v4. Oligonucletides were purchased at Sigma Aldrich. The KO cassettes were designed against *P. cirsii* genes (*phy*S, *phy*L1, *phy*L5, *phy*L6, *phy*L8, *phy*L9) and *Phomopsis CMU-LMA* genes (*sch*PKS, *sch*R2, *sch*R3, *sch*R4 *sch*R5, *sch*R7). KO cassette for gene *sch*R4 was not achieved. The KO cassettes were cloned into pE-YA vector (Figure 5.1 left). Plasmid pE-YA contains a gene for uracil production (*URA3*); a kanamycin resistance gene (*KanR*); the origin of replication *pUC ori*; and a Not-I restriction site between two sequences for homologous recombination (attL1 and attL2).
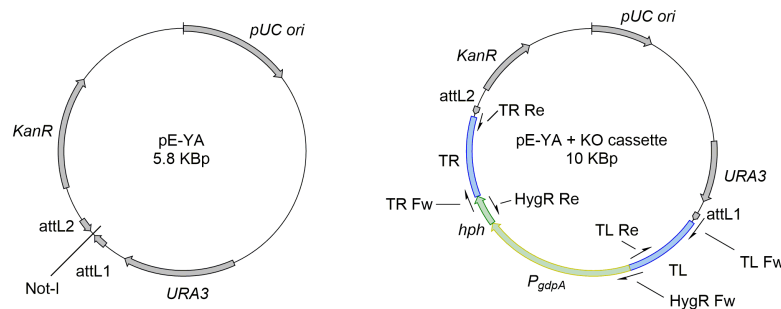
Figure 5.1: Map of pE-YA vector (left) and pE-YA + generic KO cassette (right). Primers show overlap tails.

KO cassettes consisting of hygromycin resistance flanked by 0.5-1 KBp of the target gene (Target Left and Target Right, T_L and T_R) was cloned into pE-YA (Figure 5.1 right)

exploiting homologous recombination in *S. cerevisiae PK*. The hygromycin B resistance gene *hph* was driven by the *Aspergillus nidulans gdpA* promoter (P_{gdpA}); HygR refers to the *hph*-P_{gdpA} system. Primers were designed to contain 30 Bp-long tail homologous to the adjacent segment of DNA. The tails mediate the homologous recombination in yeast. Amplification of the HygR was performed with high fidelity Q5 polymerase to preserve the resistance activity using pTH-GS-eGFP as template. T_L and T_R were amplified by Taq polymerase, using fungal genomic DNA as template. The primers used for the construction of the KO cassettes (Table 5.1) were designed with 30 nt overlapping tails to mediate homologous recombination in yeast. Prior ordering the primers were checked for annealing temperature and secondary structure formation using Sigma OligoEvaluator.

Primer	Sequence 5'- 3'
HygR Re	CTATTCCTTTGCCCTCGGAC

	Phyllosticta cirsii
*phy*S T$_L$ Fw	GCCAACTTTGTACAAAAAAGCAGGCTCCGCCGTTGTGCAGGAGATCCCTA
*phy*S T$_L$ Re	GTTTTGAAAGTCCATACCGA
*phy*S HygR Fw	TATACGTGGGATCGGTATGGACTTTCAAAACTCTAGTGGATCTTTCGACA
*phy*S T$_R$ Fw	CCCAGCACTCGTCCGAGGGCAAAGGAATAGATCGATTGGGATCAAGAGAC
*phy*S T$_R$ Re	TGCCAACTTTGTACAAGAAAGCTGGGTCGGCTAAGGACGCCAAGATATAC
*phy*L1 T$_L$ Fw	GCCAACTTTGTACAAAAAAGCAGGCTCCGCGTCTCTCTTCCGTGTCATGG
*phy*L1 T$_L$ Re	GCGAGATCGAATCTGGTAATG
*phy*L1 HygR Fw	CTTTGGCATCATTACCAGATTCGATCTCGCTCTAGTGGATCTTTCGACAC
*phy*L1 T$_R$ Fw	CCCAGCACTCGTCCGAGGGCAAAGGAATAGGTATGAACCAGCTCAACTTC
*phy*L1 T$_R$ Re	TGCCAACTTTGTACAAGAAAGCTGGGTCGGGGCATAGTTGAGGTACTTGA
*phy*L5 T$_L$ Fw	GCCAACTTTGTACAAAAAAGCAGGCTCCGCCACTCATCCACTGGCTCTAC
*phy*L5 T$_L$ Re	GACGTTGAGCGAGAACAAGG
*phy*L5 HygR Fw	GAGATCCAGGCCTTGTTCTCGCTCAACGTCTCTAGTGGATCTTTCGACAC
*phy*L5 T$_R$ Fw	GCCCCAGCACTCGTCCGAGGGCAAAGGAATAGAAAGCCCACGCTGGTCAT
*phy*L5 T$_R$ Re	TGCCAACTTTGTACAAGAAAGCTGGGTCGGGTTCTTTCTTCGGCTCCACC
*phy*L6 T$_L$ Fw	GCCAACTTTGTACAAAAAAGCAGGCTCCGCCACTCAGGTCGCGAAGAACT
*phy*L6 T$_L$ Re	GGCACTCTTCCGTCTGTATT
*phy*L6 HygR Fw	CCCGTCCTTCAATACAGACGGAAGAGTGCCTCTAGTGGATCTTTCGACAC
*phy*L6 T$_R$ Fw	CCCAGCACTCGTCCGAGGGCAAAGGAATAGGTACCGTTCTGCACTGGCGA
*phy*L6 T$_R$ Re	TGCCAACTTTGTACAAGAAAGCTGGGTCGGCCAACTTCAAAAACCCACCC
*phy*L8 T$_L$ Fw	GCCAACTTTGTACAAAAAAGCAGGCTCCGCCAACCTTGAACTTCCCATCA
*phy*L8 T$_L$ Re	CTTCTGGTAGTGTTCCCAGA
*phy*L8 HygR Fw	TCCGCACCCCCGTCTGGGAACACTACCAGAAGTCTAGTGGATCTTTCGAC
*phy*L8 T$_R$ Fw	CCCAGCACTCGTCCGAGGGCAAAGGAATAGCCATGAAGATGATCCGCTAC
*phy*L8 T$_R$ Re	TGCCAACTTTGTACAAGAAAGCTGGGTCGGCGATAATGTCCTTCTCGAAT
*phy*L9 T$_L$ Fw	GCCAACTTTGTACAAAAAAGCAGGCTCCGCCAAGCAAGGTGTCTCTTTCG
*phy*L9 T$_L$ Re	GTTCCTTTGATGTACGATGG
*phy*L9 HygR Fw	AGCATCTCATCCATCGTACATCAAAGGAACTCTAGTGGATCTTTCGACAC
*phy*L9 T$_R$ Fw	CCCAGCACTCGTCCGAGGGCAAAGGAATAGCAGCGGAGTAATGTGCATTAG
*phy*L9 T$_R$ Re	TGCCAACTTTGTACAAGAAAGCTGGGTCGGGTCATCAGTGTTGGAGCATTT

	Phomosis CMU-LMA
*sch*PKS T$_L$ Fw	GCCAACTTTGTACAAAAAAGCAGGCTCCGCATGCCTTCGTCATACATACA
*sch*PKS T$_L$ Re	AAACGAAAGGCACAGAACAT
*sch*PKS HygR Fw	CTCCTTTTCTATGTTCTGTGCCTTTCGTTTTCTAGTGGATCTTTCGACAC
*sch*PKS T$_R$ Fw	CCCAGCACTCGTCCGAGGGCAAAGGAATAGGGGCGACGTCGCCAAGGAAG
*sch*PKS T$_R$ Re	TGCCAACTTTGTACAAGAAAGCTGGGTCGGTCACGTAACAACCATCTTGG
*sch*R2 T$_L$ Fw	GCCAACTTTGTACAAAAAAGCAGGCTCCGCAACCGGACATCTCGCACACTC
*sch*R2 T$_L$ Re	GGAACTCCGACTTCAGTCTG
*sch*R2 HygR Fw	AGCTATGTCACAGACTGAAGTCGGAGTTCCTCTAGTGGATCTTTCGACAC
*sch*R2 T$_R$ Fw	CCCAGCACTCGTCCGAGGGCAAAGGAATAGGACCCAGATAGAATTCCTGC
*sch*R2 T$_R$ Re	TGCCAACTTTGTACAAGAAAGCTGGGTCGGCCAACTCCACATCCGACTTG
*sch*R3 T$_L$ Fw	GCCAACTTTGTACAAAAAAGCAGGCTCCGCCAATATTCATGCTCTTGGGC
*sch*R3 T$_L$ Re	AATCAGGGTATTCGAGAAAG
*sch*R3 HygR Fw	TATCATATGACTTTCTCGAATACCCTGATTTCTAGTGGATCTTTCGACAC
*sch*R3 T$_R$ Fw	CCCAGCACTCGTCCGAGGGCAAAGGAATAGGACCCCGTTCACGACATCTA
*sch*R3 T$_R$ Re	TGCCAACTTTGTACAAGAAAGCTGGGTCGGCGGGGTCAAATTTACAATAC
*sch*R4 T$_L$ Fw	GCCAACTTTGTACAAAAAAGCAGGCTCCGCCAAATGATCTACCTCGGGAG
*sch*R4 T$_L$ Re	CAAGCTCATGGTGCAGTATC
*sch*R4 HygR Fw	CAGCTTCCATGATACTGCACCATGAGCTTGCTATTCCTTTGCCCTCGGAC
*sch*R4 T$_R$ Fw	ACGTATTTCAGTGTCGAAAGATCCACTAGACTCTGTTCCAGTTTGCCATT
*sch*R4 T$_R$ Re	TGCCAACTTTGTACAAGAAAGCTGGGTCGGCAGGTCTCAGTATCACTTTTC
*sch*R5 T$_L$ Fw	GCCAACTTTGTACAAAAAAGCAGGCTCCGCGAGCAACAAACCATTCTGGT
*sch*R5 T$_L$ Re	GTCCAGTTGGGAATGCCACT
*sch*R5 HygR Fw	CCTGAGGACAAGTGGCATTCCCAACTGGACTCTAGTGGATCTTTCGACAC
*sch*R5 T$_R$ Fw	CCCAGCACTCGTCCGAGGGCAAAGGAATAGCGTCCTTGCACATTCATGTC
*sch*R5 T$_R$ Re	TGCCAACTTTGTACAAGAAAGCTGGGTCGGGATTGCCTAACTTCGTCTTC
*sch*R7 T$_L$ Fw	GCCAACTTTGTACAAAAAAGCAGGCTCCGCGAACCCTTACTTTCAAAGCC
*sch*R7 T$_L$ Re	GTAGAGGTCCTCGTGCTCAG
*sch*R7 HygR Fw	GCATCTCCAACTGAGCACGAGGACCTCTACTCTAGTGGATCTTTCGACAC
*sch*R7 T$_R$ Fw	CCCAGCACTCGTCCGAGGGCAAAGGAATAGGAGACATATACCGGGCCACT
*sch*R7 T$_R$ Re	TGCCAACTTTGTACAAGAAAGCTGGGTCGGCATTTATTCGGCAAAGAGTC

Table 5.1: Primers used for KO cassettes construction.

5.5.3 Scytolide putative cluster

Scytolide **186** BGC (*P. cirsii*)

Name	Annotation	Cofactor
*scy*L3	Dehydratase	-
*scy*L2	Phosphopyruvate hydratase	Metal ions
*scy*L1	*N*-terminal nucleophile aminohydrolase/DHQS	-
*scy*DHQS	Dihydroquinate synthase	$NADP^+$; Fe^{2+}
*scy*R1	*O*-MeT	SAM

Table 5.2: Annotation of the putative scytolide **186** BGC in *P. cirsii*.

5.6 Molecular biology

5.6.1 PCR

DNA amplification was performed with two kinds of polymerases: proofreading Q5 polymrase; and Taq polymerase (Table 5.3). Q5 High Fidelity 2x Master Mix was used when sequence identity was crucial, such as maintaining *hph* gene functionality; One Taq Quick-Load 2x Master Mix was sufficient in any other scenario. PCR were carried out in volumes of 10 - 100 μL depending on the goal. Template DNA was either genomic DNA or plasmid. For *E. coli* colony screening the PCR template was obtained by picking part of a single colony with a toothpick and transferring 10 μL ddH$_2$O, boil for 5 min and add 2-5 μL to the PCR reaction. Annealing temperatures were optimised by running gradient PCR of 10 μL each. For all PCR a Bio-Rad T100 thermal cycler was used (Table 5.4).

Component	Final concentration
Q5 High Fidelity 2x Master Mix / One Taq Quick-Load 2x Master Mix	1x
10 μM Fw primer	0.5 μM
10 μM Re primer	0.5 μM
template DNA	> 1 μg
ddH$_2$O	to volume

Table 5.3: PCR composition.

	Step	Temperature	Time
Q5 High Fidelity 2x Master Mix	Denaturation	98 °C	30 sec
		98 °C	10 - 30 sec
	30 cycles	50 - 70 °C	10 - 30 sec
		72 °C	20 - 30 sec/KBp
	Final extension	72 °C	2 min
	Hold	4 °C	∞
One Taq Quick-Load 2x Master Mix	Denaturation	94 °C	30 sec
		94 °C	10 - 30 sec
	30 cycles	50 - 70 °C	10 - 30 sec
		68 °C	60 sec/KBp
	Final extension	68 °C	5 min
	Hold	4 °C	∞

Table 5.4: PCR programs.

5.6.2 ITS sequencing

The internal transcribed spacer 5.8S (ITS 5.8S) was amplified from genomic DNA of *Phyllosticta cirsii* using ITS1 (TCCGTAGGTGAACCTGCGG) and ITS4 (TCCTCCGCT-TATTGATATGC) primers. The PCR products were sequenced by Sanger method by the Thermofin service facility.

The *P. cirsii* ITS sequence was used to search the NCBI database (BLAST). Clustal Omega with default settings was used to align *P. cirsii* ITS within a collection of 20 ITSs sequences from different organisms, counting *Phyllosticta spp*; the closest organisms found by NCBI blasting (including: *Phoma spp*; *Phaeosphaeria spp*; and *Leptosphaeria spp*) and also the distantly related *Aspergillus oryzae* and *Saccaromyces cerevisiae* as outliers.

5.6.3 Yeast recombination

Saccharomyces cerevisiae PKAJNJA (auxotrophic for uracil) was incubated on YPAD agar at 28 °C for up to one week or at 4 °C for up to 1 month. For stock preparation 1 mL of a ddH$_2$O, 25% glycerol was used to suspend and kept at -80 °C. Vector pE-YA allows the yeast to regain autotrophy and the uracil-free medium SM-URA (Section 5.2) is used to select the revertant cells. Vector pE-YA was linearized using the restriction enzyme *Not*-I prior to transformation. The yeast was transformed with ∼10 μg of each fragment. A control consisting of the sole linearized pE-YA was generated to observe the background given by uptake of DNA without the correct recombination of the KO cassette. Consistently, yeast transformed with the fragments mixture displays a larger number of colonies compared to the control. The whole plate was utilized to extract the plasmid by Zymoprep Yeast Plasmid Miniprep.[319] *E. coli* TOP10 was transformed to amplify the number of vector copies and selected through Kanamycin 50 μg/ml. Single colony PCR

was performed targeting a region spanning T_R and part of hygR. The positive colonies were amplified overnight and the DNA purified by mini prep precipitation. Plasmids were screened by enzymatic digestion to confirm the success of the recombination of the KO cassette into pE-YA.

5.6.4 *E. coli* TOP10 transformation

Frozen stocks of *E. coli* TOP10 chemically competent cells were thawed on ice and 5 μg of vector were added. The solution was mixed gently and incubated on ice for 30 min. Cells were subjected to heat shock at 42 °C for 30 sec and put back in ice for 2 min. 250 mL of SOC medium was added and cells were incubated at 37 °C in shacking conditions at 220 rpm for 1 h. The selection was performed overnight on LB agar plates with antibiotic 50 μg/mL at 37 °C. Transformant colonies were picked with a sterile toothpick and grown overnight in 2 mL LB broth containing 50 μg/mL antibiotic.

5.6.5 Miniprep DNA purification

Cells were collected by centrifugation at max potency for 2 minutes. 100 μL of Solution I [5 mM glucose; 25 mM Tris pH 8; 10 mM EDTA pH8] with RNase 0.5 μg/mL were added to the pellet and incubate on ice for 5 min. 200 μL of solution II [0.2 M NaOH; 1% SDS additional after autoclave] were added to the mixture and mixed by inversion followed by incubation on ice for 5 min. Additon of 150 μL of Solution III at 4 °C [50 mL 5 M KAc; 11.5 mL glacial acid acetic; 20 mL H_2O) and put in ice for 5 min. Vials were spinned at maximum speed for 15 min at 4 °C to get rid of all cellular debris. The supernatant was transferred in a fresh 1.5 mL Eppendorf. 1 mL concentrated ethanol cold was added along with 50 μL of 3 M KAc. The solution was mixed by inverting and placed on ice for 30 min (max 2-3 h). The sample was spinned at max speed for 15 min, the supernatant was discarded and the pellet dissolved in 20 μL of nuclease-free ddH_2O. The concentration of the plasmid was assessed by Nanodrop and agarose gel elecrophoresis.

5.6.6 Fungal transformation

A protocol for *P. cirsii* and *Phomopsis CMU-LMA* transformation was developed basing on the transformation protocol adopted for *Aspergillus oryzae*. The vector pTH-GS-eGFP (Figure 5.2) was used to determine whether the fungi are transformable. pTH-GS-eGFP contains the hygromycin resistance *hph* under the control of the fungal promoter P_{gdpA}; the enhanced green fluorescent protein gene (*eGFP*) controlled by the inducible promoter P_{amyB}; and an ampicillin resistance (*AmpR*). We chose pTH-GS-eGFP to test whether is possible to transform the fungi; to test the effectiveness of the P_{gdpA}-*hph* hygromycin B resistance system; and to test whether P_{amyB} is recognised by *P. cirsii* and *Phomopsis CMU-LMA*.
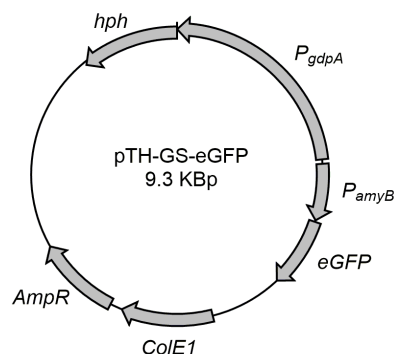
Figure 5.2: Map of pTH-GS-eGFP vector.

Mycelia from growing plates (4-5 days old) were inoculated into 50 mL PDB medium and incubated for 4-5 days at 28 °C in static condition to avoid clumping. Germinated cells were collected by Myracloth filtration. 10 mL of 10 mg/mL filter sterilized *Trichoderma harzianum* lysing enzyme (TLE) was used to resuspend the cellular pellet. TLE was dissolved in NaCl (0.8 M for *Phompsis CMU-LMA* transformation and 1.3 M for *P. cirsii* transformation). The tube was incubated at 30 °C with gentle mixing for 2-3 hours. The protoplasts were released from hyphal strands by gentle pipetting with a wide-bore pipette and filtered by Miracloth. The filtrate was centrifugated at 3000 x g for 5 minutes to pellet the protoplasts. Protoplasts were resuspended in 100 μL of Solution I [0.8/1.3 M NaCl; 10 mM CaCl$_2$; 50 mM Tris-HCl pH 7.5]. The concentration of the protoplast was assessed microspically by hemocytometer. Around 10 μg of pTH-GS-eGFP were added to the solution and incubated on ice for 2 minutes. 1 mL of Solution II [60% (w/v) PEG 3350; 0.8/1.3 M NaCl; 10 mM Tris-HCl pH 7.5] was added in the transformation mixture and incubated at RT for 20 minutes. 5 mL of molten CDZ/S agar was added to the transformation mixture and laid into Petri dish after mixing. Plates were incubated overnight at 28 °C to let the protoplasts grow back the cell wall. The next day 10 mL of CD agar with Hygromycin B 100 μg/mL were spread onto the plates in order to select the transformants. Single colonies were picked onto secondary plates with 100 μg/mL hygromycin B. The positive colonies were grown in liquid with no antibiotics.

To achieve a KO of a specific gene, bipartite transformation was adopted in order to reduce the number of false positive transformants: the specific pE-YA + KO cassette was used as PCR template to originate two overlapping fragments (α and β) resulting in a defective HygR, so that random insertion of one fragment cannot lead to antibiotic resistance (Figure 1.25, Section 1.7.3). Primers used to split the antibiotic resistance were HygRP7.2 Fw (GCTTTCAGCTTCGATGTAGG) and HygRP6.2 Re (CGTCAGGA-CATTGTTGGAG) coupled with the target-specific T_L Fw and T_R Re respectively (Figure 5.3). HygRP7.2 Fw + T_L Fw yielded the α fragment, while HygRP6.2 Re and T_R Re yielded the β fragments. Both α and β fragments were transformed simultaneously (5-10 μg each) into the fungal cells.
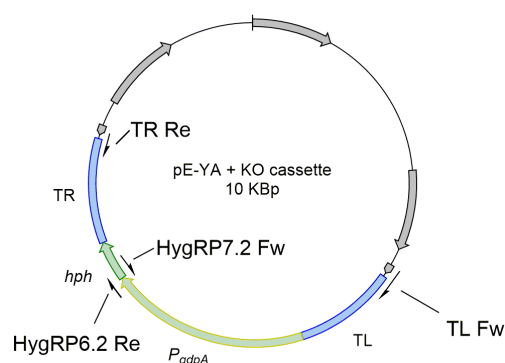
Figure 5.3: Map of pE-YA + KO cassette with primers HygRP7.2 Fw + T_L Fw and Hy-gRP6.2 Re + T_R Re.

5.6.7 Fungal transformants screening

Transformed protoplasts were plated on primary Petri dish with HygB 100-150 μg/mL and incubated at 28 °C for ~10 days. Single colonies that grew though the antibiotic were picked and plated on a secondary plate with antibiotic of the same concentration. Colonies that could grow easily on secondary plate were picked and their gDNA extracted directly from a tiny piece of agar using the GenElute Plant Genomic DNA Miniprep Kit (Aldrich).[240] The agar was smashed with pestle in the Eppendorf prior to gDNA extraction. gDNA was eluted in a final volume of 30 μL ddH$_2$O.

gDNA was used as template for two PCR: whole PCR (wPCR) and external PCR (extPCR). Primers used for wPCR were T_L Fw and T_R Re. Primers for extPCR (Table 5.5) were designed to be on gDNA outside the gene of interest and were coupled with either HygRP7.2 Fw or HygRP6.2 Re accordingly to the gene orientation.

Primer	Sequence 5'- 3'
*phy*S ext	GGCGAGAGTGACGAGATGAG
*phy*L1 ext	CATGAAAATAACAATGTTAATC
*phy*L5 ext	CTTGTATTACAGCGCTGTAG
*phy*L6 ext	CAGTGGGCTTAATAGTATAG
*phy*L8 ext	AGCCACGACAGCACTTCTTG
*phy*L9 ext	GAAATTCACGTCATGAACAC
*sch*PKS ext	CTAGGCCTAGACTACTTGAC
*sch*R2 ext	GAGAAAAGGCGATTTGAGTC
*sch*R3 ext	GAATCACAAAGCCACCGTAATG
*sch*R5 ext	GCATTAGCTCAACGGACC
*sch*R7 ext	GCATTAGCTCAACGGACC

Table 5.5: Primers used for transformant screening.

Bibliography

[1] Michael Wink. Evolution of secondary metabolites from an ecological and molecular phylogenetic perspective. *Phytochemistry*, 64(1):3–19, 2003.

[2] Richard N Bennett and Roger M Wallsgrove. Secondary metabolites in plant defence mechanisms. *New phytologist*, 127(4):617–633, 1994.

[3] Barbara Schulz, Christine Boyle, Siegfried Draeger, Anne-Katrin Römmert, and Karsten Krohn. Endophytic fungi: a source of novel biologically active secondary metabolites. *Mycological Research*, 106(9):996–1004, 2002.

[4] Brian P Akers, Juan Francisco Ruiz, Alan Piper, and Carl AP Ruck. A prehistoric mural in spain depicting neurotropic psilocybe mushrooms? *Economic Botany*, 65(2):121–128, 2011.

[5] Rainer W Bussmann and Douglas Sharon. Traditional medicinal plant use in northern peru: tracking two thousand years of healing culture. *Journal of Ethnobiology and Ethnomedicine*, 2(1):47, 2006.

[6] David E Jones. *Poison arrows: North American Indian hunting and warfare.* University of Texas Press, 2007.

[7] Bhuwan Khatri Chhetri, Serge Lavoie, Anne Marie Sweeney-Jones, and Julia Kubanek. Recent trends in the structural revision of natural products. *Nat. Prod. Rep.*, 00:1–18, 2018.

[8] David J. Newman and Gordon M. Cragg. Natural products as sources of new drugs over the last 25 years. *J. Nat. Prod.*, 70(3):461–477, 2007.

[9] Armin Bauer and Mark Brönstrup. Industrial natural product chemistry for drug discovery and development. *Nat. Prod. Rep.*, 31(1):35–60, 2014.

[10] Jassem G Mahdi. Medicinal potential of willow: A chemical perspective of aspirin discovery. *Journal of Saudi Chemical Society*, 14(3):317–322, 2010.

[11] János Fischer, C Robin Ganellin, A Ganesan, and J Proudfoot. *Analogue-based drug discovery.* Wiley-VCH, 2010.

[12] O Lafont. Clarification on publications concerning the synthesis of acetylsalicylic acid, 1996.

[13] Satoshi Omura. Microbial metabolites: 45 years of wandering, wondering and discovering. *Tetrahedron*, 67(35):6420–6459, 2011.

[14] Raoul Herbrecht, David W Denning, Thomas F Patterson, John E Bennett, Reginald E Greene, Jörg-W Oestmann, Winfried V Kern, Kieren A Marr, Patricia Ribaud, Olivier Lortholary, et al. Voriconazole versus amphotericin b for primary therapy of invasive aspergillosis. *New England Journal of Medicine*, 347(6):408–415, 2002.

[15] Greg Keller, Lori Spatola, Dennis McCabe, Brian Martinell, William Swain, and Maliyakal E John. Transgenic cotton resistant to herbicide bialaphos. *Transgenic Research*, 6(6):385–392, 1997.

[16] William A Craig. Interrelationship between pharmacokinetics and pharmacodynamics in determining dosage regimens for broad-spectrum cephalosporins. *Diagnostic microbiology and infectious disease*, 22(1):89–96, 1995.

[17] Stuart L Schreiber and Gerald R Crabtree. The mechanism of action of cyclosporin a and fk506. *Immunology today*, 13(4):136–142, 1992.

[18] Christer Paul, Jan Liliemark, Ulf Tidefelt, Gösta Gahrton, and Curt Peterson. Pharmacokinetics of daunorubicin and doxorubicin in plasma and leukemic cells from patients with acute nonlymphoblastic leukemia. *Therapeutic drug monitoring*, 11(2):140–148, 1989.

[19] Armando Santoro, Thomas Tursz, Henning Mouridsen, Jaap Verweij, William P Steward, Reiner Somers, Jose Buesa, Paolo Casali, David Spooner, and Elaine M Rankin. Doxorubicin versus cyvadic versus doxorubicin plus ifosfamide in first-line treatment of advanced soft tissue sarcomas: a randomized study of the european organization for research and treatment of cancer soft tissue and bone sarcoma group. *Journal of Clinical Oncology*, 1995.

[20] Kazuei Igarashi, Hideo Ishitsuka, and Akira Kaji. Comparative studies on the mechanism of action of lincomycin, streptomycin, and erythromycin. *Biochemical and biophysical research communications*, 37(3):499–504, 1969.

[21] Fred E Hahn and Stefan G Sarre. Mechanism of action of gentamicin. *The Journal of infectious diseases*, 119(4/5):364–369, 1969.

[22] Hamao Umezawa, Masahiro Ueda, Kenji Maeda, Koki Yagishita, Shinichi Kondo, Yoshiro Okami, Ryozo Utahara, Yasuke Osato, Kazuo Nitta, and Tomio Takeuchi. Production and isolation of a new antibiotic: kanamycin. *The Journal of antibiotics*, 10(5):181, 1957.

[23] Christopher G Sotiriou and Judy WM Cheng. Beneficial effects of statins in coronary artery disease - beyond lowering cholesterol. *Annals of Pharmacotherapy*, 34(12):1432–1439, 2000.

[24] Masanao Matsui, Yasuhiro Yamada, Keizo Uzu, and Tadashi Hirata. Studies on mitomycins. iii. *The Journal of antibiotics*, 21(3):189–198, 1968.

[25] Min Chen and MJ Wolin. Effect of monensin and lasalocid-sodium on the growth of methanogenic and rumen saccharolytic bacteria. *Applied and Environmental Microbiology*, 38(1):72–77, 1979.

[26] Hamao Umezawa, Satoshi Mizuno, Hisaji Yamazaki, and Kazuo Nitta. Inhibition of dna-dependent rna synthesis by rifamycins. *The Journal of antibiotics*, 21(3):234–236, 1968.

[27] Haruyasu Kinashi, Noboru Ōtake, Hiroshi Yonehara, Shoichi Sato, and Yoshihiko Saito. The structure of salinomycin, a new member of the polyether antibiotics. *Tetrahedron Letters*, 14(49):4955–4958, 1973.

[28] John D Pirsch, Joshua Miller, Mark H Deierhoi, Flavio Vincenti, and Ronald S Filo. A comparison of tacrolimus (fk506) and cyclosporine for immunosuppression after cadaveric renal transplantation1. *Transplantation*, 63(7):977–983, 1997.

[29] I Chopra, PM Hawkey, and M Hinton. Tetracyclines, molecular and clinical aspects. *Journal of Antimicrobial Chemotherapy*, 29(3):245–277, 1992.

[30] Dudley H Williams and Ben Bardsley. The vancomycin group of antibiotics and the fight against resistant bacteria. *Angewandte Chemie International Edition*, 38(9):1172–1193, 1999.

[31] Sascha Baumann, Jennifer Herrmann, Ritesh Raju, Heinrich Steinmetz, Kathrin I Mohr, Stephan Hüttel, Kirsten Harmrolfs, Marc Stadler, and Rolf Müller. Cystobactamids: myxobacterial topoisomerase inhibitors exhibiting potent antibacterial activity. *Angewandte Chemie International Edition*, 53(52):14605–14609, 2014.

[32] Juliane Korp, María S Vela Gurovic, and Markus Nett. Antibiotics from predatory bacteria. *Beilstein journal of organic chemistry*, 12:594, 2016.

[33] Daniel L Klayman. Qinghaosu (artemisinin): an anti-malarial drug from china. *Science*, 228(4703):1049–1055, 1985.

[34] Steven R Meshnick. Artemisinin: mechanisms of action, resistance and toxicity. *International journal for parasitology*, 32(13):1655–1660, 2002.

[35] R Bryan Yamasaki, James A Klocke, S Mark Lee, Gregory A Stone, and Mark V Darlington. Isolation and purification of azadirachtin from neem (azadirachta indica) seeds using flash chromatography and high-performance liquid chromatography. *Journal of Chromatography A*, 356:220–226, 1986.

[36] J Hrnorgan Butterworth and ED Morgan. Isolation of a substance that suppresses feeding in locusts. *Chemical Communications (London)*, (1):23–24, 1968.

[37] KC Nicolaou, Z Yang, JJ Liu, H Ueno, PG Nantermet, RK Guy, CF Claiborne, J Renaud, EA Couladouros, K Paulvannan, et al. Total synthesis of taxol. *Nature*, 367(6464):630, 1994.

[38] Teruaki Mukaiyama, Isamu Shiina, Hayato Iwadare, Masahiro Saitoh, Toshihiro Nishimura, Naoto Ohkawa, Hiroki Sakoh, Koji Nishimura, Yu-ichirou Tani, Masatoshi Hasegawa, et al. Asymmetric total synthesis of taxol\r. *Chemistry-A European Journal*, 5(1):121–161, 1999.

[39] Eric WP Damen, Lesly Braamer, and Hans W Scheeren. Lanthanide trifluoromethanesulfonate catalysed selective acylation of 10-deacetylbaccatin iii. *Tetrahedron letters*, 39(33):6081–6082, 1998.

[40] M. P. Caulfield. Muscarinic Receptors-Characterization, coupling and function. *Pharmacol. Ther.*, 58(3):319–379, 1993.

[41] Raissa Schor and Russell Cox. Classic fungal natural products in the genomic age: the molecular legacy of Harold Raistrick. *Nat. Prod. Rep.*, 00:1–27, 2018.

[42] J. Davison, A. al Fahad, M. Cai, Z. Song, S. Y. Yehia, C. M. Lazarus, A. M. Bailey, T. J. Simpson, and R. J. Cox. Genetic, molecular, and biochemical basis of fungal tropolone biosynthesis. *Proc. Natl. Acad. Sci.*, 109(20):7642–7647, 2012.

[43] James Staunton and Kira J. Weissman. Polyketide biosynthesis: a millennium review. *Nat. Prod. Rep.*, 18(4):380–416, 2001.

[44] Timm Maier, Timm Maier, Marc Leibundgut, Marc Leibundgut, Nenad Ban, and Nenad Ban. The crystal structure of a mammalian fatty acid synthase. *Science*, 321(September):1315–22, 2008.

[45] C Nguyen, R W Haushalter, D J Lee, P R Markwick, J Bruegger, G Caldara-Festin, K Finzel, D R Jackson, F Ishikawa, B O'Dowd, J A McCammon, S J Opella, S C Tsai, and M D Burkart. Trapping the dynamic acyl carrier protein in fatty acid biosynthesis. *Nature*, 505(7483):427–431, 2014.

[46] Ban Nenad Simon Jenni, Leibundgut Marc, Boehringer Daniel, Frick Christian, Mikolásek Bohdan. Structure of Fungal Fatty Acid Synthase and Implication for Iterative Substrate Shuttling. *Science*, 316(80):254–261, 2007.

[47] Stuart Smith and Shiou-Chuan Tsai. The type I fatty acid and polyketide synthases: a tale of two megasynthases. *Nat. Prod. Rep.*, 24(5):1041, 2007.

[48] Russell J. Cox. Polyketides, proteins and genes in fungi: programmed nano-machines begin to reveal their secrets. *Org. Biomol. Chem.*, 5(13):2010, 2007.

[49] Michael A Fischbach and Christopher T Walsh. Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: logic, machinery, and mechanisms. *Chemical reviews*, 106(8):3468–3496, 2006.

[50] Bernard J Rawlings. Type i polyketide biosynthesis in bacteria (part a-erythromycin biosynthesis). *Natural product reports*, 18(2):190–227, 2001.

[51] R Stanzak, P Matsushima, RH Baltz, and RN Rao. Cloning and expression in streptomyces lividans of clustered erythromycin biosynthesis genes from streptomyces erythreus. *Nature Biotechnology*, 4(3):229, 1986.

[52] David E. Cane. Programming of erythromycin biosynthesis by a modular polyketide synthase. *J. Biol. Chem.*, 285(36):27517–27523, 2010.

[53] Chantel D Campbell and John C Vederas. Biosynthesis of lovastatin and related metabolites formed by fungal iterative pks enzymes. *Biopolymers*, 93(9):755–763, 2010.

[54] Leibniz Hang, Nicholas Liu, and Yi Tang. Coordinated and Iterative Enzyme Catalysis in Fungal Polyketide Biosynthesis. *ACS Catal.*, 6(9):5935–5945, 2016.

[55] Brian AM Rudd and David A Hopwood. Genetics of actinorhodin biosynthesis by streptomyces coelicolor a3 (2). *Microbiology*, 114(1):35–43, 1979.

[56] Christian Hertweck, Andriy Luzhetskyy, Yuri Rebets, and Andreas Bechthold. Type ii polyketide synthases: gaining a deeper insight into enzymatic teamwork. *Natural product reports*, 24(1):162–190, 2007.

[57] Christopher J. Arthur, Anna Szafranska, Simon E. Evans, Stuart C. Findlow, Steven G. Burston, Philip Owen, Ian Clark-Lewis, Thomas J. Simpson, John Crosby, and Matthew P. Crump. Self-malonylation is an intrinsic property of a chemically synthesized type II polyketide synthase acyl carrier protein. *Biochemistry*, 44(46):15414–15421, 2005.

[58] Sabine Grüschow, Tonia J. Buchholz, Wolfgang Seufert, Jonathan S. Dordick, and David H. Sherman. Substrate profile analysis and ACP-mediated acyl transfer in Streptomyces coelicolor type III polyketide synthases. *ChemBioChem*, 8(8):863–868, 2007.

[59] Ikuro Abe, Tsuyoshi Abe, Kiyofumi Wanibuchi, and Hiroshi Noguchi. Enzymatic formation of quinolone alkaloids by a plant type iii polyketide synthase. *Organic letters*, 8(26):6063–6065, 2006.

[60] Nobutaka Funa, Takayoshi Awakawa, and Sueharu Horinouchi. Pentaketide resorcylic acid synthesis by type iii polyketide synthase from neurospora crassa. *Journal of Biological Chemistry*, 282(19):14476–14481, 2007.

[61] Yit Heng Chooi and Yi Tang. Navigating the fungal polyketide chemical space: From genes to molecules. *J. Org. Chem.*, 77(22):9933–9953, 2012.

[62] Lan-Qing Ma, Yan-Wu Guo, Dong-Yao Gao, Dong-Ming Ma, You-Nian Wang, Guo-Feng Li, Ben-Ye Liu, Hong Wang, and He-Chun Ye. Identification of a polygonum cuspidatum three-intron gene encoding a type iii polyketide synthase producing both naringenin and p-hydroxybenzalacetone. *Planta*, 229(5):1077–1086, 2009.

[63] Desen Zheng and Geza Hrazdina. Molecular and biochemical characterization of benzalacetone synthase and chalcone synthase genes and their proteins from raspberry (rubus idaeus l.). *Archives of biochemistry and biophysics*, 470(2):139–145, 2008.

[64] Mohamed A Marahiel, Torsten Stachelhaus, and Henning D Mootz. Modular peptide synthetases involved in nonribosomal peptide synthesis. *Chemical reviews*, 97(7):2651–2674, 1997.

[65] Dirk Schwarzer, Robert Finking, and Mohamed A. Marahiel. Nonribosomal peptides: from genes to products. *Nat. Prod. Rep.*, 20(3):275, 2003.

[66] Dirk Schwarzer, Robert Finking, and Mohamed A Marahiel. Nonribosomal peptides: from genes to products. *Natural product reports*, 20(3):275–287, 2003.

177

[67] Liangcheng Du and Lili Lou. Pks and nrps release mechanisms. *Natural product reports*, 27(2):255–278, 2010.

[68] Robert Finking and Mohamed A Marahiel. Biosynthesis of nonribosomal peptides. *Annu. Rev. Microbiol.*, 58:453–488, 2004.

[69] Nicole M Gaudelli, Darcie H Long, and Craig A Townsend. β-lactam formation by a non-ribosomal peptide synthetase during antibiotic biosynthesis. *Nature*, 520(7547):383, 2015.

[70] Jack Ho Wong. Fungal toxins. In *Handbook of Biologically Active Peptides (Second Edition)*, pages 166–168. Elsevier, 2013.

[71] Katja Maria Fisch. Biosynthesis of natural products by microbial iterative hybrid pks–nrps. *RSC Advances*, 3(40):18228–18247, 2013.

[72] Carsten D Richter, Daniel Nietlispach, R William Broadhurst, and Kira J Weissman. Multienzyme docking in hybrid megasynthetases. *Nature chemical biology*, 4(1):75, 2008.

[73] Kira J Weissman. The structural biology of biosynthetic megaenzymes. *Nature chemical biology*, 11(9):660, 2015.

[74] Eva Maria Niehaus, Katharina W. Von Bargen, José J. Espino, Andreas Pfannmüller, Hans Ulrich Humpf, and Bettina Tudzynski. Characterization of the fusaric acid gene cluster in Fusarium fujikuroi. *Appl. Microbiol. Biotechnol.*, 98(4):1749–1762, 2014.

[75] Rainer Schobert and Andrea Schlenk. Tetramic and tetronic acids: An update on new derivatives and biological aspects. *Bioorganic Med. Chem.*, 16(8):4203–4221, 2008.

[76] Werner Graf, Jean-louis Robert, John C Vederas, and Christoph Tamm. Biosynthesis of the Cytochalasans. Part III. ^{13}C-NMR. *Helv. Chim. Acta*, 57(190):1801–1815, 1974.

[77] Christopher T Walsh, Huawei Chen, Thomas A Keating, Brian K Hubbard, Heather C Losey, Lusong Luo, C Gary Marshall, Deborah Ann Miller, and Hiten M Patel. Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on nrps assembly lines. *Current opinion in chemical biology*, 5(5):525–534, 2001.

[78] DJG White and RY Calne. The use of cyclosporin a immunosuppression in organ grafting. *Immunological reviews*, 65(1):115–131, 1982.

[79] Zhongshu Song, Russell J. Cox, Colin M. Lazarus, and Thomas J. Simpson. Fusarin C biosynthesis in Fusarium moniliforme and Fusarium venenatum. *ChemBioChem*, 5(9):1196–1203, 2004.

[80] S Yu Yunusov et al. Alkaloids. *Alkaloids.*, 1981.

[81] Michael Wink. Chemical ecology of alkaloids. In *Alkaloids*, pages 265–300. Springer, 1998.

[82] Roland R Griffiths, William A Richards, Una McCann, and Robert Jesse. Psilocybin can occasion mystical-type experiences having substantial and sustained personal meaning and spiritual significance. *Psychopharmacology*, 187(3):268–283, 2006.

[83] ED Levin, SJ Briggs, NC Christopher, and JE Rose. Chronic nicotine stimulation and blockade effects on working memory. *Behavioural pharmacology*, 1993.

[84] Jane Achan, Ambrose O Talisuna, Annette Erhart, Adoke Yeka, James K Tibenderana, Frederick N Baliraine, Philip J Rosenthal, and Umberto D'Alessandro. Quinine, an old anti-malarial drug in a modern world: role in the treatment of malaria. *Malaria journal*, 10(1):144, 2011.

[85] Sérgio H Ferreira, Igor DG Duarte, and Berenice B Lorenzetti. The molecular mechanism of action of peripheral morphine analgesia: stimulation of the cgmp system via nitric oxide release. *European journal of pharmacology*, 201(1):121–122, 1991.

[86] Brent A Smith. Strychnine poisoning. *Journal of Emergency Medicine*, 8(3):321–325, 1990.

[87] WR Martin, JW Sloan, JD Sapira, and DR Jasinski. Physiologic, subjective, and behavioral effects of amphetamine, methamphetamine, ephedrine, phenmetrazine, and methylphenidate in man. *Clinical Pharmacology & Therapeutics*, 12(2part1):245–258, 1971.

[88] Margaret F Roberts. *Alkaloids: biochemistry, ecology, and medicinal applications*. Springer Science & Business Media, 2013.

[89] Philipp Christen and Perdeep K Mehta. From cofactor to enzymes. the molecular evolution of pyridoxal-5'-phosphate-dependent enzymes. *The Chemical Record*, 1(6):436–447, 2001.

[90] Jonathan S Nishimura, James M Manning, and Alton Meister. Studies on the mechanism of activation of aspartic acid β-decarboxylase by α-keto acids and pyridoxal 5'-phosphate. *Biochemistry*, 1(3):442–447, 1962.

[91] Hideyuki Hayashi, Hiroshi Wada, Tohru Yoshimura, Nobuyoshi Esaki, and Kenji Soda. Recent topics in pyridoxal 5'-phosphate enzyme studies. *Annual review of biochemistry*, 59(1):87–110, 1990.

[92] Mario A Rivera, Arthur C Aufderheide, Larry W Cartmell, Constantino M Torres, and Odin Langsjoen. Antiquity of coca-leaf chewing in the south central andes: a 3,000 year archaeological record of coca-leaf chewing from northern chile. *Journal of psychoactive drugs*, 37(4):455–458, 2005.

[93] Thomas R Hoye, Jeffrey A Bjorklund, Dmitry O Koltun, and Matthew K Renner. N-methylputrescine oxidation during cocaine biosynthesis: Study of prochiral methylene hydrogen discrimination using the remote isotope method. *Organic letters*, 2(1):3–5, 2000.

[94] Gregor Wolfgang Schmidt, Jan Jirschitzka, Tiffany Porta, Michael Reichelt, Katrin Luck, José Carlos Pardo Torre, Franziska Dolke, Emmanuel Varesio, Gérard Hopfgartner, Jonathan Gershenzon, et al. The last step in cocaine biosynthesis is catalyzed by a bahd acyltransferase. *Plant physiology*, 167(1):89–101, 2015.

[95] Andrew J Humphrey and David O'Hagan. Tropane alkaloid biosynthesis. a century old problem unresolved. *Natural product reports*, 18(5):494–502, 2001.

[96] Teresa Docimo, Michael Reichelt, Bernd Schneider, Marco Kai, Grit Kunert, Jonathan Gershenzon, and John C D Auria. The first step in the biosynthesis of cocaine in erythroxylum coca: the characterization of arginine and ornithine decarboxylases. *Plant molecular biology*, 78(6):599–615, 2012.

[97] Eberhard Breitmaier. *Terpenes: flavors, fragrances, pharmaca, pheromones*. John Wiley & Sons, 2006.

[98] Eric Oldfield and Fu-Yang Lin. Terpene biosynthesis: modularity rules. *Angewandte Chemie International Edition*, 51(5):1124–1137, 2012.

[99] Vincent JJ Martin, Douglas J Pitera, Sydnor T Withers, Jack D Newman, and Jay D Keasling. Engineering a mevalonate pathway in escherichia coli for production of terpenoids. *Nature biotechnology*, 21(7):796, 2003.

[100] Wolfgang Eisenreich, Felix Rohdich, and Adelbert Bacher. Deoxyxylulose phosphate pathway to terpenoids. *Trends in plant science*, 6(2):78–84, 2001.

[101] Courtney M Starks, Kyoungwhan Back, Joseph Chappell, and Joseph P Noel. Structural basis for cyclic terpene biosynthesis by tobacco 5-epi-aristolochene synthase. *Science*, 277(5333):1815–1820, 1997.

[102] Herve Gambliel and Rodney Croteau. Pinene cyclases i and ii. two enzymes from sage (salvia officinalis) which catalyze stereospecific cyclizations of geranyl pyrophosphate to monoterpene olefins of opposite configuration. *Journal of Biological Chemistry*, 259(2):740–748, 1984.

[103] Takayuki Itoh, Kinya Tokunaga, Yudai Matsuda, Isao Fujii, Ikuro Abe, Yutaka Ebizuka, and Tetsuo Kushiro. Reconstitution of a fungal meroterpenoid biosynthesis reveals the involvement of a novel family of terpene cyclases. *Nature chemistry*, 2(10):858–864, 2010.

[104] Mairi E Raggatt, Thomas J Simpson, and M Inês Chicarelli-Robinson. Biosynthesis of xenovulene a®: formation of a cyclopentenone via a unique ring expansion–ring contraction mechanism. *Chemical Communications*, (22):2245–2247, 1997.

[105] M David Marks, Li Tian, Jonathan P Wenger, Stephanie N Omburo, Wilfredo Soto-Fuentes, Ji He, David R Gang, George D Weiblen, and Richard A Dixon. Identification of candidate genes affecting $\delta 9$-tetrahydrocannabinol biosynthesis in cannabis sativa. *Journal of experimental botany*, 60(13):3715–3726, 2009.

[106] Ute Metzger. The structure of dimethylallyl tryptophan synthase reveals a common architecture of aromatic prenyltransferases in fungi and bacteria.

[107] Ole Rigbers. Ergot alkaloid biosynthesis in Aspergillus fumigatus: overproduction and biochemical characterization of a 4-dimethylallyltryptophan N-methyl transferase.

[108] Nicole Lorenz, Jana Olšovská, Miroslav Šulc, and Paul Tudzynski. Alkaloid cluster gene ccsA of the ergot fungus Claviceps purpurea encodes chanoclavine I synthase, a flavin adenine dinucleotide-containing oxidoreductase mediating the transformation of N-methyl-dimethylallyltryptophan to chanoclavine I. *Appl. Environ. Microbiol.*, 76(6):1822–1830, 2010.

[109] D P Dowling, N A Bruender, A P Young, R M McCarty, V Bandarian, and C L Drennan. Radical SAM enzyme QueE defines a new minimal core fold and metal-dependent mechanism. *Nat. Chem. Biol.*, 10(2):106–112, 2014.

[110] Bradley J Landgraf and Squire J Booker. The ylide has landed. *Nature*, 498(May):45–47, 2013.

[111] John G Scandalios. Oxygen stress and superoxide dismutases. *Plant physiology*, 101(1):7, 1993.

[112] John Emsley. *Nature's building blocks: an A-Z guide to the elements*. Oxford University Press, 2011.

[113] Robert P Hausinger. Fe (ii)/α-ketoglutarate-dependent hydroxylases and related enzymes. *Critical reviews in biochemistry and molecular biology*, 39(1):21–68, 2004.

[114] Paul R Ortiz De Montellano. *Cytochrome P450: structure, mechanism, and biochemistry*. Springer Science & Business Media, 2005.

[115] Christopher J Schofield and Zhihong Zhang. Structural and mechanistic studies on 2-oxoglutarate-dependent oxygenases and related enzymes. *Current opinion in structural biology*, 9(6):722–731, 1999.

[116] V Massey. Introduction: flavoprotein structure and mechanism. *The FASEB journal*, 9(7):473–475, 1995.

[117] Gerard J Bishop. Refining the plant steroid hormone biosynthesis pathway. *Trends in plant science*, 12(9):377–380, 2007.

[118] Paul R Ortiz de Montellano. Substrate oxidation by cytochrome p450 enzymes. In *Cytochrome P450*, pages 111–176. Springer, 2015.

[119] Debashree Basudhar, Yarrow Madrona, Sylvie Kandel, Jed N Lampe, Clinton R Nishida, and Paul R Ortiz de Montellano. Analysis of cytochrome p450 cyp119 ligand-dependent conformational dynamics by two-dimensional nmr and x-ray crystallography. *Journal of Biological Chemistry*, 290(16):10000–10017, 2015.

[120] Matthew J Ryle and Robert P Hausinger. Non-heme iron oxygenases. *Current opinion in chemical biology*, 6(2):193–201, 2002.

[121] Eric L Hegg and Lawrence Que Jr. The 2-his-1-carboxylate facial triadâĂŤan emerging structural motif in mononuclear non-heme iron (ii) enzymes. *European Journal of Biochemistry*, 250(3):625–629, 1997.

[122] Giulio Magni, Giuseppe Orsomando, Nadia Raffelli, and Silverio Ruggieri. Enzymology of mammalian nad metabolism in health and disease. *Front Biosci*, 13:6135–6154, 2008.

[123] Richard S Rivlin. Riboflavin metabolism. *New England Journal of Medicine*, 283(9):463–472, 1970.

[124] Christopher T Walsh and Timothy A Wencewicz. Flavoenzymes: versatile catalysts in biosynthetic pathways. *Natural product reports*, 30(1):175–200, 2013.

[125] Marc Fontecave, Mohamed Atta, and Etienne Mulliez. S-adenosylmethionine: nothing goes to waste. *Trends in biochemical sciences*, 29(5):243–249, 2004.

[126] Jürg Müller, Craig M Hart, Nicole J Francis, Marcus L Vargas, Aditya Sengupta, Brigitte Wild, Ellen L Miller, Michael B O'Connor, Robert E Kingston, and Jeffrey A Simon. Histone methyltransferase activity of a drosophila polycomb group repressor complex. *Cell*, 111(2):197–208, 2002.

[127] François Fuks, Paul J Hurd, Rachel Deplus, and Tony Kouzarides. The dna methyltransferases associate with hp1 and the suv39h1 histone methyltransferase. *Nucleic acids research*, 31(9):2305–2312, 2003.

[128] Thomas Ernst, Andrew J Chase, Joannah Score, Claire E Hidalgo-Curtis, Catherine Bryant, Amy V Jones, Katherine Waghorn, Katerina Zoi, Fiona M Ross, Andreas Reiter, et al. Inactivating mutations of the histone methyltransferase gene ezh2 in myeloid disorders. *Nature genetics*, 42(8):722, 2010.

[129] M José Fernández Lozano, Lily L Remsing, Luis M Quirós, Alfredo F Braña, Ernestina Fernández, César Sánchez, Carmen Méndez, Jürgen Rohr, and José A Salas. Characterization of two polyketide methyltransferases involved in the biosynthesis of the antitumor drug mithramycin bystreptomyces argillaceus. *Journal of Biological Chemistry*, 275(5):3065–3074, 2000.

[130] Chloe Zubieta, Xian-Zhi He, Richard A Dixon, and Joseph P Noel. Structures of two natural product methyltransferases reveal the basis for substrate specificity in plant o-methyltransferases. *Nature Structural and Molecular Biology*, 8(3):271, 2001.

[131] Emily M Drabant, Ahmad R Hariri, Andreas Meyer-Lindenberg, Karen E Munoz, Venkata S Mattay, Bhaskar S Kolachana, Michael F Egan, and Daniel R Weinberger. Catechol o-methyltransferase val158met genotype and neural mechanisms related to affective arousal and regulation. *Archives of general psychiatry*, 63(12):1396–1406, 2006.

[132] Janis Fricke, Felix Blei, and Dirk Hoffmeister. Enzymatic Synthesis of Psilocybin. *Angew. Chemie - Int. Ed.*, 2017.

[133] Chie Koshiishi, Ayako Kato, Sachiko Yama, Alan Crozier, and Hiroshi Ashihara. A new caffeine biosynthetic pathway in tea leaves: utilisation of adenosine released from the s-adenosyl-l-methionine cycle. *FEBS letters*, 499(1-2):50–54, 2001.

[134] Shelly C Lu. S-adenosylmethionine. *The international journal of biochemistry & cell biology*, 32(4):391–395, 2000.

[135] Rachel U Hutcheson and Joan B Broderick. Radical sam enzymes in methylation and methylthiolation. *Metallomics*, 4(11):1149–1154, 2012.

[136] Tyler L. Grove Booker, Squire J., Robert M. Cicchillo. Self Sacrifice in Radical S-adenosylmethionine Proteins. *Biochemistry*, 11(5):543–552, 2009.

[137] Joan B Broderick, Benjamin R Duffus, Kaitlin S Duschene, and Eric M Shepard. Radical s-adenosylmethionine enzymes. *Chemical reviews*, 114(8):4229–4317, 2014.

[138] Jianping Xu. *Next-generation Sequencing*. Caister Academic Press, 2014.

[139] Stephan C Schuster. Next-generation sequencing transforms today's biology. *Nature methods*, 5(1):16, 2007.

[140] Elaine R Mardis. The impact of next-generation sequencing technology on genetics. *Trends in genetics*, 24(3):133–141, 2008.

[141] Michael L Metzker. Sequencing technologies - the next generation. *Nature reviews genetics*, 11(1):31, 2010.

[142] Lin Liu, Yinhu Li, Siliang Li, Ni Hu, Yimin He, Ray Pong, Danni Lin, Lihua Lu, and Maggie Law. Comparison of next-generation sequencing systems. *BioMed Research International*, 2012, 2012.

[143] Barry Merriman, Ion Torrent, Jonathan M Rothberg, R&D Team, et al. Progress in ion torrent semiconductor chip based sequencing. *Electrophoresis*, 33(23):3397–3417, 2012.

[144] Jan Berka, Zhoutao Chen, Michael Egholm, Brian C Godwin, Stephen Kyle Hutchison, John Harris Leamon, Gary James Sarkis, and Jan Fredrik Simons. Paired end sequencing, October 13 2009. US Patent 7,601,499.

[145] Zhoutao Chen, Brian Christopher Godwin, Gianni Calogero Ferreri, and David Roderick Riches. Paired end sequencing, September 17 2009. US Patent App. 12/322,119.

[146] Ting Ni, David L Corcoran, Elizabeth A Rach, Shen Song, Eric P Spana, Yuan Gao, Uwe Ohler, and Jun Zhu. A paired-end sequencing strategy to map the complex landscape of transcription initiation. *Nature methods*, 7(7):521, 2010.

[147] Jose Luis Goicoechea, Jetty Siva S Ammiraju, Pradeep Reddy Marri, Mingsheng Chen, Scott Jackson, Yeisoo Yu, Steve Rounsley, and Rod A Wing. The future of rice genomics: sequencing the collective oryza genome. *Rice*, 3(2-3):89–97, 2010.

[148] Jeremy Schmutz, Phillip E McClean, Sujan Mamidi, G Albert Wu, Steven B Cannon, Jane Grimwood, Jerry Jenkins, Shengqiang Shu, Qijian Song, Carolina Chavarro, et al. A reference genome for common bean and genome-wide analysis of dual domestications. *Nature genetics*, 46(7):707, 2014.

[149] Kai Blin, Marnix H. Medema, Daniyal Kazempour, Michael A. Fischbach, Rainer Breitling, Eriko Takano, and Tilmann Weber. antiSMASH 2.0–a versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Res.*, 41(Web Server issue):204–212, 2013.

[150] Tilmann Weber, Kai Blin, Srikanth Duddela, Daniel Krug, Hyun Uk Kim, Robert Bruccoleri, Sang Yup Lee, Michael A Fischbach, Rolf Müller, Wolfgang Wohlleben, et al. antismash 3.0âĂŤa comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic acids research*, 43(W1):W237–W243, 2015.

[151] Inga Bach. Biosynthetic gene clusters. 2012.

[152] Jason C Slot and Antonis Rokas. Horizontal transfer of a large and highly toxic secondary metabolic gene cluster between fungi. *Current biology*, 21(2):134–139, 2011.

[153] Norio Kimura and Takashi Tsuge. Gene cluster involved in melanin biosynthesis of the filamentous fungus alternaria alternata. *Journal of Bacteriology*, 175(14):4427–4435, 1993.

[154] Nora Khaldi, Jérôme Collemare, Marc-Henri Lebrun, and Kenneth H Wolfe. Evidence for horizontal transfer of a secondary metabolite gene cluster between fungi. *Genome biology*, 9(1):R18, 2008.

[155] Axel A Brakhage and Volker Schroeckh. Fungal secondary metabolites–strategies to activate silent gene clusters. *Fungal Genetics and Biology*, 48(1):15–22, 2011.

[156] Ben Field and Anne E Osbourn. Metabolic diversification - independent assembly of operon-like gene clusters in different plants. *Science*, 320(5875):543–547, 2008.

[157] X Qi, S Bakht, M Leggett, C Maxwell, R Melton, and A Osbourn. A gene cluster for secondary metabolism in oat: implications for the evolution of metabolic diversity in plants. *Proceedings of the National Academy of Sciences of the United States of America*, 101(21):8233–8238, 2004.

[158] Michael L Nielsen, Line Albertsen, Gaëlle Lettier, Jakob B Nielsen, and Uffe H Mortensen. Efficient pcr-based gene targeting with a recyclable marker for aspergillus nidulans. *Fungal Genetics and Biology*, 43(1):54–64, 2006.

[159] Xiuren Zhang, Rossana Henriques, Shih-Shun Lin, Qi-Wen Niu, and Nam-Hai Chua. Agrobacterium-mediated transformation of arabidopsis thaliana using the floral dip method. *Nature protocols*, 1(2):641, 2006.

[160] André Hoekema, Penny R Hirsch, Paul JJ Hooykaas, and Rob A Schilperoort. A binary plant vector strategy based on separation of vir-and t-region of the agrobacterium tumefaciens ti-plasmid. *Nature*, 303(5913):179, 1983.

[161] Le Cong, F Ann Ran, David Cox, Shuailiang Lin, Robert Barretto, Naomi Habib, Patrick D Hsu, Xuebing Wu, Wenyan Jiang, Luciano Marraffini, et al. Multiplex genome engineering using crispr/cas systems. *Science*, page 1231143, 2013.

[162] Philippe Horvath and Rodolphe Barrangou. Crispr/cas, the immune system of bacteria and archaea. *Science*, 327(5962):167–170, 2010.

[163] Marcel JA De Groot, Paul Bundock, Paul JJ Hooykaas, and Alice GM Beijersbergen. Agrobacterium tumefaciens-mediated transformation of filamentous fungi. *Nature biotechnology*, 16(9):839, 1998.

[164] E D Mullins, X Chen, P Romaine, R Raina, DM Geiser, and S Kang. Agrobacterium-mediated transformation of fusarium oxysporum: an efficient tool for insertional mutagenesis and gene transfer. *Phytopathology*, 91(2):173–180, 2001.

[165] Rui Liu, Ling Chen, Yanping Jiang, Zhihua Zhou, and Gen Zou. Efficient genome editing in filamentous fungus trichoderma reesei using the crispr/cas9 system. *Cell Discovery*, 1:15007, 2015.

[166] Takayuki Arazoe, Kennosuke Miyoshi, Tohru Yamato, Tetsuo Ogawa, Shuichi Ohsato, Tsutomu Arie, and Shigeru Kuwata. Tailor-made crispr/cas system for highly efficient targeted gene replacement in the rice blast fungus. *Biotechnology and bioengineering*, 112(12):2543–2549, 2015.

[167] Christina S Nødvig, Jakob B Nielsen, Martin E Kogle, and Uffe H Mortensen. A crispr-cas9 system for genetic engineering of filamentous fungi. *PLoS One*, 10(7):e0133085, 2015.

[168] Mariana Schuster, Gabriel Schweizer, Stefanie Reissmann, and Regine Kahmann. Genome editing in ustilago maydis using the crispr–cas system. *Fungal Genetics and Biology*, 89:3–9, 2016.

[169] Jean Paul Latgé. The cell wall: A carbohydrate armour for the fungal cell. *Mol. Microbiol.*, 66(2):279–290, 2007.

[170] Shaun M. Bowman and Stephen J. Free. The structure and synthesis of the fungal cell wall. *BioEssays*, 28(8):799–808, 2006.

[171] David J. Adams. Fungal cell wall chitinases and glucanases. *Microbiology*, 150(7):2029–2035, 2004.

[172] Wolf B Frommer and Olaf Ninnemann. Heterologous expression of genes in bacterial, fungal, animal, and plant cells. *Annual review of plant biology*, 46(1):419–444, 1995.

[173] Mary N Heneghan, Ahmed A Yakasai, Laura M Halo, Zhongshu Song, Andrew M Bailey, Thomas J Simpson, Russell J Cox, and Colin M Lazarus. First heterologous reconstruction of a complete functional fungal biosynthetic multigene cluster. *ChemBioChem*, 11(11):1508–1512, 2010.

[174] Julia Schümann and Christian Hertweck. Advances in cloning, functional analysis and heterologous expression of fungal polyketide synthase genes. *Journal of biotechnology*, 124(4):690–703, 2006.

[175] Yasushi Sugano, Ryosuke Nakano, Katsuya Sasaki, and Makoto Shoda. Efficient heterologous expression inaspergillus oryzae of a unique dye-decolorizing peroxidase, dyp, of geotrichum candidum dec 1. *Applied and environmental microbiology*, 66(4):1754–1758, 2000.

[176] Yi-Ming Chiang, C Elizabeth Oakley, Manmeet Ahuja, Ruth Entwistle, Aric Schultz, Shu-Lin Chang, Calvin T Sung, Clay CC Wang, and Berl R Oakley. An efficient system for heterologous expression of secondary metabolite genes in aspergillus nidulans. *Journal of the American Chemical Society*, 135(20):7720–7731, 2013.

[177] Dominik Mumberg, Rolf Müller, and Martin Funk. Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene*, 156(1):119–122, 1995.

[178] CM Lazarus, K Williams, and AM Bailey. Reconstructing fungal natural product biosynthetic pathways. *Natural product reports*, 31(10):1339–1347, 2014.

[179] Khomaizon A K Pahirulzaman, Katherine Williams, and Colin M. Lazarus. A toolkit for heterologous expression of metabolic pathways in aspergillus oryzae. *Methods Enzymol.*, 517:241–260, 2012.

[180] Anja Schueffler and Timm Anke. Fungal natural products in research and development. *Natural product reports*, 31(10):1425–1448, 2014.

[181] David L Hawksworth. The magnitude of fungal diversity: the 1.5 million species estimate revisited. *Mycological research*, 105(12):1422–1432, 2001.

[182] Nancy P Keller, Geoffrey Turner, and Joan W Bennett. Fungal secondary metabolism - from biochemistry to genomics. *Nature Reviews Microbiology*, 3(12):937, 2005.

[183] Daniel S Heckman, David M Geiser, Brooke R Eidell, Rebecca L Stauffer, Natalie L Kardos, and S Blair Hedges. Molecular evidence for the early colonization of land by fungi and plants. *Science*, 293(5532):1129–1133, 2001.

[184] Yusuf Abou-Jawdah, Hana Sobh, and Abdu Salameh. Antimycotic activities of selected plant flora, growing wild in lebanon, against phytopathogenic fungi. *Journal of agricultural and food chemistry*, 50(11):3208–3213, 2002.

[185] Gary A Strobel, R Vincent Miller, Concepcion Martinez-Miller, Margaret M Condron, David B Teplow, and WM Hess. Cryptocandin, a potent antimycotic from the endophytic fungus cryptosporiopsis cf. quercina. *Microbiology*, 145(8):1919–1926, 1999.

[186] Dave W Bartlett, John M Clough, Jeremy R Godwin, Alison A Hall, Mick Hamer, and Bob Parr-Dobrzanski. The strobilurin fungicides. *Pest management science*, 58(7):649–662, 2002.

[187] Herman Goossens, Matus Ferech, Robert Vander Stichele, Monique Elseviers, ESAC Project Group, et al. Outpatient antibiotic use in europe and association with resistance: a cross-national database study. *The Lancet*, 365(9459):579–587, 2005.

[188] Ramanan Laxminarayan, Adriano Duse, Chand Wattal, Anita KM Zaidi, Heiman FL Wertheim, Nithima Sumpradit, Erika Vlieghe, Gabriel Levy Hara, Ian M Gould, Herman Goossens, et al. Antibiotic resistanceâĂŤthe need for global solutions. *The Lancet infectious diseases*, 13(12):1057–1098, 2013.

[189] F Baquero, J Martínez Beltrén, and E Loza. A review of antibiotic resistance patterns of streptococcus pneumoniae in europe. *Journal of Antimicrobial Chemotherapy*, 28(suppl_C):31–38, 1991.

[190] Ester Boserup. *The conditions of agricultural growth: The economics of agrarian change under population pressure*. Routledge, 2017.

[191] Antonio Evidente, Cimmino Alessio, Anna Andolfi, Vurro Maurizio, Maria Chiara Zonno, Charles L Cantrell, and Andrea Motta. Phyllostictines A-D, oxazatricycloalkenones produced by Phyllosticta cirsii, a potential mycoherbicide for Cirsium arvense biocontrol. *Tetrahedron*, 64:1612–1619, 2008.

[192] Emilie Adelin, Claudine Servy, Sylvie Cortial, Hélne Lévaique, Marie Thérse Martin, Pascal Retailleau, Géraldine Le Goff, Boonsom Bussaban, Saisamorn Lumyong, and Jamal Ouazzani. Isolation, structure elucidation and biological activity of metabolites from Sch-642305-producing endophytic fungus Phomopsis sp. CMU-LMA. *Phytochemistry*, 72(18):2406–2412, 2011.

[193] Barry B. Snider and Jingye Zhou. Synthesis of (+)-Sch 642305 by a biomimetic transannular Michael reaction. *Org. Lett.*, 8(7):1283–1286, 2006.

[194] Björn Kusebauch, Benjamin Busch, Kirstin Scherlach, Martin Roth, and Christian Hertweck. Polyketide-chain branching by an enzymatic Michael addition. *Angew. Chemie - Int. Ed.*, 48(27):5001–5004, 2009.

[195] Esther Müller and Wolfgang Nentwig. Plant pathogens as biocontrol agents of Cirsium arvense - an overestimated approach? *NeoBiota*, 11:1–24, 2011.

[196] Gordon E. D. Tiley. Biological Flora of the British Isles: Cirsium arvense (L.) Scop. *J. Ecol.*, 98(4):938–983, 2010.

[197] R Moore. The biology of canadian weed. *Can. J. Plant Sci*, 55(May):1033 − 1048, 1975.

[198] Kerri Skinner, Lincoln Smith, and Peter Rice. Using noxious weed lists to prioritize targets for developing weed management strategies. *Weed Sci.*, 48(5):640–644, 2000.

[199] J Stach. Länge und gewicht der sprosse und der wurzelausläufer von acker-kratzdistel (cirsium arvense (l.) scop.) in mais. *Zeitschrift für Pflanzenkrankheiten und Pflanzenschutz, Sonderheft XV*, pages 87–90, 1996.

[200] William W Donald. The biology of canada thistle (cirsium arvense), 1994.

[201] A Hugh Gourlay et al. Classical biological control of californian thistle: the new zealand story. In *Weed management: balancing people, planet, profit. 14th Australian Weeds Conference, Wagga Wagga, New South Wales, Australia, 6-9 September 2004: papers and proceedings*, pages 374–377. Weed Society of New South Wales, 2004.

[202] Olga Gimeno, Pawel Plucinski, Stan T. Kolaczkowski, Francisco J. Rivas, and Pedro M. Alvarez. Removal of the herbicide MCPA by commercial activated carbons: Equilibrium, kinetics, and reversibility. *Ind. Eng. Chem. Res.*, 42(5):1076–1086, 2003.

[203] Robert Nicholson, Anthony Navoy, Ph D, and Laura Nicholson. W Ork P Lan T He K Irkwood -C Ohansey P Roject W Ork P Lan. 34(1):10–22, 2003.

[204] J Kovach, C Petzoldt, J Degni, and J Tette. A Method to Measure The Environmental Impact of Pesticides. *New York's Food Life Sci. Bull.*, 139(139):1–8, 1992.

[205] J. Kobylecka, B. Ptaszynski, R. Rogaczewski, and A. Turek. Phenoxyalkanoic acid complexes. Part I. Complexes of lead(II), cadmium(II) and copper(II) with 4-chloro-2-methylphenoxyacetic acid (MCPA). *Thermochim. Acta*, 407(1-2):25–31, 2003.

[206] Stephen D. Strachan, Sergio C. Nanita, Marc Ruggiero, Mark S. Casini, Kathleen M. Heldreth, Larry H. Hageman, Helen A. Flanigan, Nancy M. Ferry, and Anne M. Pentz. Correlation of Chemical Analysis of Residual Levels of Aminocyclopyrachlor in Soil to Biological Responses of Alfalfa, Cotton, Soybean, and Sunflower. *Weed Technol.*, 25(2):239–244, 2011.

[207] Rachel E. Cruttwell McFadyen. Biological Control of Weeds. *Annu. Rev. Entomol.*, 43(1):369–393, 1998.

[208] Dragana Rancic, Branka Stevanovic, Radmila Petanović, Biljana Magud, Ivo Tosevski, and André Gassmann. Anatomical injury induced by the eriophyid mite Aceria anthocoptes on the leaves of Cirsium arvense. *Exp. Appl. Acarol.*, 38(4):243–253, 2006.

[209] J Frantzen. The Role of Clonal Growth in the Pathosystem Cirsium-Arvense Puccinia-Punctiformis. *Can. J. Bot. Can. Bot.*, 72(6):832–836, 1994.

[210] S. Kluth, A. Kruess, and T. Tscharntke. Effects of two pathogens on the performance of Cirsium arvense in a successional fallow. *Weed Res.*, 45(4):261–269, 2005.

[211] R.C. French and A.R. Lightfield. Induction of systemic aecial infection in Canada Thistle (Cirsium arvense) by Teliospores of Puccinia punctiformis, 1990.

[212] Vibeke Leth, J Netland, and Christian Andreasen. Phomopsis cirsii: a potential biocontrol agent of cirsium arvense. *Weed research*, 48(6):533–541, 2008.

[213] V Leth and C Andreasen. Alginate helps phomopsis cirsii grove to overcome short dew periods. In *BIO-Y2K Combined Millenium Meeting, Grahamstown*, pages 23–28, 2000.

[214] Brenda S Brosten and David C Sands. Field trials of sclerotinia sclerotiorum to control canada thistle (cirsium arvense). *Weed Science*, 34(3):377–380, 1986.

[215] GW Bourdôt, GA Hurrell, DJ Saville, and DM Leathwick. Impacts of applied sclerotinia sclerotiorum on the dynamics of a cirsium arvense population. *Weed research*, 46(1):61–72, 2006.

[216] AO Berestetskii, OS Yuzikhin, AS Katkova, AV Dobrodumov, DE Sivogrivov, and LV Kolombet. Isolation, identification, and characteristics of the phytotoxin produced by the fungus alternaria cirsinoxia. *Applied biochemistry and microbiology*, 46(1):75–79, 2010.

[217] S Green and KL Bailey. Effects of leaf maturity, infection site, and application rate of alternaria cirsinoxia conidia on infection of canada thistle (cirsium arvense). *Biological Control*, 19(2):167–174, 2000.

[218] S Guske, B Schulz, and C Boyle. Biocontrol options for cirsium arvense with indigenous fungal pathogens. *Weed research*, 44(2):107–116, 2004.

[219] Andreas Kruess. Indirect interaction between a fungal plant pathogen and a herbivorous beetle of the weed cirsium arvense. *Oecologia*, 130(4):563–569, 2002.

[220] SL Bithell, A Stewart, SM Zydenbos, et al. Evaluation of the pathogenicity of phoma exigua var. exigua on californian thistle. In *New Zealand Plant Protection Volume 54, 2001. Proceedings of a conference, Quality Hotel, Palmerston North, New Zealand, 14-16 August 2001.*, pages 179–183. New Zealand Plant Protection Society, 2001.

[221] Health Protection Branch and Welfare Canada. Cytochalasins A and B from Strains of Phoma exigua var. exigua and Formation of Cytochalasin B in Potato Gangrene. (1975):77–80, 1975.

[222] EL Gasich and AO Berestetskiy. Studies on soil application of stagonospora cirsii, a candidate for biological control of cirsium arvense. *Agronomy Research*, 4(Special Issue):171–175, 2006.

[223] Oleg Yuzikhin, Galina Mitina, and Alexander Berestetskiy. Herbicidal potential of stagonolide, a new phytotoxic nonenolide from Stagonospora cirsii. *J. Agric. Food Chem.*, 55(19):7707–7711, 2007.

[224] V Leth. Inventor of patent âĂIJherbicide containing phytotoxic fungal material from phomopsis cirsii or septoria cirsii, especially for control of compositaeâĂİ. *Patent no. EP A*, 13685, 1985.

[225] Mauricio Vurro. Natural metabolites for parasitic weed management. *PestManag Sci*, 65:566–571, 2009.

[226] Maurizio Vurro and Maria Chiara Zonno. Phyllostictine A, a potential natural herbicide produced by Phyllosticta cirsii: In vitro production and toxicity. *Plant Sci.*, 175:818–825, 2008.

[227] B a Bailey, J C Jennings, and J D Anderson. The 24-kDa protein from Fusarium oxysporum f.sp. erythroxyli: occurrence in related fungi and the effect of growth medium on its production. *Can. J. Microbiol.*, 43(1):45–55, 1997.

[228] John W. Gronwald, Kathryn L. Plaisance, and Bryan a. Bailey. Effects of the fungal protein Nep1 and Pseudomonas syringae on growth of Canada thistle (Cirsium arvense), common ragweed (Ambrosia artemisiifolia), and common dandelion (Taraxacum officinale). *Weed Sci.*, 52(1):98–104, 2004.

[229] John W Gronwald, Kathryn L Plaisance, Donald A Ide, and Donald L Wyse. Assessment of pseudomonas syringae pv. tagetis as a biocontrol agent for canada thistle. *Weed science*, 50(3):397–404, 2002.

[230] Samuel Coe, Nicole Pereira, Joanna V Geden, Guy J Clarkson, David J Fox, Richard M Napier, Paul Neve, and Michael Shipman. Ring closing metathesis reactions of α-methylene-β-lactams: application to the synthesis of a simplified phyllostictine analogue with herbicidal activity. *Org. {&} Biomol. Chem.*, 13(28):7655–7663, 2015.

[231] Axel Dalhoff, Nebojsa Janjic, and Roger Echols. Redefining penems. *Biochemical pharmacology*, 71(7):1085–1095, 2006.

182

[232] A Salamov and V Solovyev. Fgenesh multiple gene prediction program, 1998.

[233] Arlie J Rinaldi, Paul E Lund, Mario R Blanco, and Nils G Walter. The shine-dalgarno sequence of riboswitch-regulated single mrnas shows ligand-dependent accessibility bursts. *Nature communications*, 7:8976, 2016.

[234] Mark Johnson, Irena Zaretskaya, Yan Raytselis, Yuri Merezhuk, Scott McGinnis, and Thomas L Madden. Ncbi blast: a better web interface. *Nucleic acids research*, 36(suppl_2):W5–W9, 2008.

[235] Thomas Madden. The blast sequence analysis tool. 2013.

[236] Tim Carver, Simon R. Harris, Matthew Berriman, Julian Parkhill, and Jacqueline A. McQuillan. Artemis: An integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics*, 28(4):464–469, 2012.

[237] Mario Stanke, Rasmus Steinkamp, Stephan Waack, and Burkhard Morgenstern. Augustus: a web server for gene finding in eukaryotes. *Nucleic acids research*, 32(suppl_2):W309–W312, 2004.

[238] Fabian Sievers, Andreas Wilm, David Dineen, Toby J Gibson, Kevin Karplus, Weizhong Li, Rodrigo Lopez, Hamish McWilliam, Michael Remmert, Johannes Söding, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using clustal omega. *Molecular systems biology*, 7(1):539, 2011.

[239] R H Nilsson, E Kristiansson, M Ryberg, and N Hallenberg. Intraspecific ITS variability in the kigdom Fungi as expressed in the internacional sequence databases and its implications for molecular species identification. *Evol. Bioinforma.*, 4:193–201, 2008.

[240] Sigma. GenElute Plant Genomic DNA Miniprep Kit. page 12, 2012.

[241] Katherine N Maloney, Wenshan Hao, Jun Xu, Jay Gibbons, John Hucul, Sean F Brady, Frank C Schroeder, and Jon Clardy. NIH Public Access. 8(18):4067–4070, 2008.

[242] Katherine N. Maloney, Wenshan Hao, Jun Xu, Jay Gibbons, John Hucul, Deborah Roll, Sean F. Brady, Frank C. Schroeder, and Jon Clardy. Phaeosphaeride A, an inhibitor of STAT3-dependent signaling isolated from an endophytic fungus. *Org. Lett.*, 8(18):4067–4070, 2006.

[243] Daniel Wibberg, Louise Andersson, Georgios Tzelepis, Oliver Rupp, Jochen Blom, Lukas Jelonek, Alfred Pühler, Johan Fogelqvist, Mark Varrelmann, Andreas Schlüter, and Christina Dixelius. Genome analysis of the sugar beet pathogen Rhizoctonia solani AG2-2IIIB revealed high numbers in secreted proteins and cell wall degrading enzymes. *BMC Genomics*, 17(1):245, 2016.

[244] Oliver Rupp, Jennifer Becker, Karina Brinkrolf, Christina Timmermann, Nicole Borth, Alfred Pühler, Thomas Noll, and Alexander Goesmann. Construction of a public CHO cell line transcript database using versatile bioinformatics analysis pipelines. *PLoS One*, 9(1), 2014.

[245] Mario Stanke, Oliver Schöffmann, Burkhard Morgenstern, and Stephan Waack. Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. *BMC Bioinformatics*, 7(1):62, 2006.

[246] Brian O Bachmann and Jacques Ravel. Methods for in silico prediction of microbial polyketide and nonribosomal peptide biosynthetic pathways from dna sequence data. *Methods in enzymology*, 458:181–217, 2009.

[247] Laura M. Halo, James W. Marshall, Ahmed A. Yakasai, Zhongshu Song, Craig P. Butts, Matthew P. Crump, Mary Heneghan, Andrew M. Bailey, Thomas J. Simpson, Colin M. Lazarus, and Russell J. Cox. Authentic heterologous expression of the tenellin iterative polyketide nonribosomal peptide synthetase requires cooexpression with an enoyl reductase. *ChemBioChem*, 9(4):585–594, 2008.

[248] Laura M. Halo, Mary N. Heneghan, Ahmed A. Yakasai, Zhongshu Song, Katherine Williams, Andrew M. Bailey, Russell J. Cox, Colin M. Lazarus, and Thomas J. Simpson. Late stage oxidations during the biosynthesis of the 2-pyridone tenellin in the entomopathogenic fungus beauveria bassiana. *J. Am. Chem. Soc.*, 130(52):17988–17996, 2008.

[249] Igor V. Grigoriev, Roman Nikitin, Sajeet Haridas, Alan Kuo, Robin Ohm, Robert Otillar, Robert Riley, Asaf Salamov, Xueling Zhao, Frank Korzeniewski, Tatyana Smirnova, Henrik Nordberg, Inna Dubchak, and Igor Shabalov. MycoCosm portal: Gearing up for 1000 fungal genomes. *Nucleic Acids Res.*, 42(D1):699–704, 2014.

[250] Igor V. Grigoriev, Daniel Cullen, Stephen B. Goodwin, David Hibbett, Thomas W. Jeffries, Christian P. Kubicek, Cheryl Kuske, Jon K. Magnuson, Francis Martin, Joseph W. Spatafora, Adrian Tsang, and Scott E. Baker. Fueling the future with fungal genomics. *Mycology*, 2(3):192–209, 2011.

[251] Simon V S Ipcho, James K. Hane, Eva A. Antoni, Dag Ahren, Bernard Henrissat, Timothy L. Friesen, Peter S. Solomon, and Richard P. Oliver. Transcriptome analysis of Stagonospora nodorum: Gene models, effectors, metabolism and pantothenate dispensability. *Mol. Plant Pathol.*, 13(6):531–545, 2012.

[252] Takayoshi Awakawa, Xiao Long Yang, Toshiyuki Wakimoto, and Ikuro Abe. Pyranonigrin E: A PKS-NRPS hybrid metabolite from aspergillus niger identified by genome mining. *ChemBioChem*, 14(16):2095–2099, 2013.

[253] Tsuyoshi Yamamoto, Yuta Tsunematsu, Hiroshi Noguchi, Kinya Hotta, and Kenji Watanabe. Elucidation of Pyranonigrin Biosynthetic Pathway Reveals a Mode of Tetramic Acid, Fused γ-Pyrone, and exo-Methylene Formation. *Org. Lett.*, 17(20):4992–4995, 2015.

[254] Linda Gritz and Julian Davies. Plasmid-encoded hygromycin b resistance: the sequence of hygromycin b phosphotransferase gene and its expression in escherichia coli and saccharomyces cerevisiae. *Gene*, 25(2):179–188, 1983.

[255] R N Rao, N E Allen, and J N Hobbs. Genetic and enzymatic basis of hygromycin B resistance in Escherichia coli. *Antimicrob. Agents Chemother.*, 24(5):689–695, 1983.

[256] Sigma-Aldrich. Trichoderma Lysing Enzymes Product Specification. *Build. Res. Inf.*, 21(1):21–22, 1993.

[257] Sigma-Aldrich. Driselase Product Specification. *Build. Res. Inf.*, 21(1):21–22, 1993.

[258] Francesco Trenti and Russell J. Cox. Structural Revision and Biosynthesis of the Fungal Phytotoxins Phyllosticines A and B. *J. Nat. Prod.*, 80(5):1235–1240, 2017.

[259] Victoria V. Abzianidze, Ekaterina V. Poluektova, Ksenia P. Bolshakova, Taras L. Panikorovskii, Alexander S. Bogachenkov, and Alexander O. Berestetskiy. Crystal structure of natural phaeosphaeride A. *Acta Crystallogr. Sect. E Crystallogr. Commun.*, 71(1975):o625–o626, 2015.

[260] Chimica Acta, Organische Chemie, St Johanns-ring, and C H Base. 86. Studies on the Biosynthesis of Spirostaphylotrichin A. 1989.

[261] Rainer Schobert. Domino syntheses of bioactive tetronic and tetramic acids. *Naturwissenschaften*, 94(1):1–11, 2007.

[262] Hengyu Xu, Babak Andi, Jinghua Qian, Ann H. West, and Paul F. Cook. The α-aminoadipate pathway for lysine biosynthesis in fungi. *Cell Biochem. Biophys.*, 46(1):43–64, 2006.

[263] Henry J. Vogel. Distribution of Lysine Pathways Among Fungi: Evolutionary Implications. *Am. Nat.*, 98(903):435–446, 1964.

183

[264] T Mark Zabriskie and Michael D Jackson. Lysine biosynthesis and metabolism in fungi. *Nat. Prod. Rep.*, 17(July 1999):85–97, 2000.

[265] Xinyu Liu and Christopher T. Walsh. Cyclopiazonic acid biosynthesis in Aspergillus sp.: Characterization of a reductase-like R* domain in cyclopiazonate synthetase that forms and releases cyclo-acetoacetyl-L-tryptophan. *Biochemistry*, 48(36):8746–8757, 2009.

[266] Yit Heng Chooi, Mariano Jordi Muria-Gonzalez, and Peter S. Solomon. A genome-wide survey of the secondary metabolite biosynthesis genes in the wheat pathogen Parastagonospora nodorum. *Mycology*, 5(3):192–206, 2014.

[267] Paul H. M. Harrison, Petar A. Duspara, Stephen I. Jenkins, Salima A. Kassam, David K. Liscombe, and Donald W. Hughes. The biosynthesis of pramanicin in Stagonospora sp. ATCC 74235: a modified acyltetramic acidâĂă. *J. Chem. Soc. Perkin Trans. 1*, (24):4390–4402, 2000.

[268] P Duspara, S I Jenkins, D W Hughes, and P H M Harrison. The biosynthesis of pramanicin: intact incorporation of serine and absolute configuration of the antibiotic. *Chem. Commun.*, (23):2643–2644, 1998.

[269] Roy Garcia, C L Yu, a Hudnall, Robyn Catlett, K L Nelson, T Smithgall, D J Fujita, S P Ethier, and R Jove. Constitutive activation of Stat3 in fibroblasts transformed by diverse oncoproteins and in breast carcinoma cells. *Cell Growth Differ.*, 8(12):1267–76, 1997.

[270] Debra L Silver, Honami Naora, Jinsong Liu, Debra L Silver, Honami Naora, Jinsong Liu, Wenjun Cheng, and Denise J Montell. Activated Signal Transducer and Activator of Transcription (STAT) 3 : Localization in Focal Adhesions and Function in Ovarian Cancer Cell Motility Activated Signal Transducer and Activator of Transcription (STAT) 3 : Localization in Focal Adhesions an. pages 3550–3558, 2004.

[271] a Pansky, P Hildebrand, E Fasler-Kan, L Baselgia, S Ketterer, C Beglinger, and M H Heim. Defective Jak-STAT signal transduction pathway in melanoma cells resistant to growth inhibition by interferon-alpha. *Int. J. cancer*, 85(5):720–5, 2000.

[272] Guilian Niu, Kenneth L. Wright, Mei Huang, Lanxi Song, Eric Haura, James Turkson, Shumin Zhang, Tianhong Wang, Dominic Sinibaldi, Domenico Coppola, Richard Heller, Lee M. Ellis, James Karras, Jacqueline Bromberg, Drew Pardoll, Richard Jove, and Hua Yu. Constitutive Stat3 activity up-regulates VEGF expression and tumor angiogenesis. *Oncogene*, 21(13):2000–2008, 2002.

[273] Yves-René Naves and AV Grampoloff. Etudes sur les matières végétales volatiles xx. sur la composition de l'extrait éthéro-pétrolique (essence concrète). de la fleur de jasmin. *Helvetica Chimica Acta*, 25(7):1500–1514, 1942.

[274] Anokha S Ratnayake, Wesley Y Yoshida, Susan L Mooberry, and Thomas Hemscheidt. The structure of microcarpalide, a microfilament disrupting agent from an endophytic fungus. *Organic letters*, 3(22):3479–3481, 2001.

[275] Masashi Tsuda, Takao Mugishima, Kazusei Komatsu, Teruo Sone, Michiko Tanaka, Yuzuru Mikami, and Jun'ichi Kobayashi. Modiolides a and b, two new 10-membered macrolides from a marine-derived fungus. *Journal of natural products*, 66(3):412–415, 2003.

[276] Dachepally Aravind Kumar and Harshadas Mitaram Meshram. Concise stereoselective total synthesis of (+)-mueggelone. *Synthetic Communications*, 43(8):1145–1154, 2013.

[277] Antonio Evidente, Alessio Cimmino, Alexander Berestetskiy, Anna Andolfi, and Andrea Motta. Stagonolides g- i and modiolide a, nonenolides produced by stagonospora cirsii, a potential mycoherbicide for cirsium arvense. *Journal of natural products*, 71(11):1897–1901, 2008.

[278] Xue Song Chen, Shi Jun Da, Li Hong Yang, Bo Yan Xu, Zhi Xiang Xie, and Ying Li. A new convenient asymmetric approach to herbarumin iii. *Chinese Chemical Letters*, 18(3):255–257, 2007.

[279] José Fausto Rivero-Cruz, Martha Macías, Carlos M Cerda-García-Rojas, and Rachel Mata. A new phytotoxic nonenolide from phoma h erbarum. *Journal of natural products*, 66(4):511–514, 2003.

[280] Debendra K Mohapatra, Gokarneswar Sahoo, Dhondi K Ramesh, J Srinivasa Rao, and G Narahari Sastry. Total syntheses and absolute stereochemistry of decarestrictines c1 and c2. *Tetrahedron Letters*, 50(40):5636–5639, 2009.

[281] John W Blunt, Brent Copp, Robert A Keyzers, Murray HG Munro, and Michele R Prinsep. Marine natural products. 2014.

[282] Min Chu, Ronald Mierzwa, Ling Xu, Ling He, Joseph Terracciano, Mahesh Patel, Vincent Gullo, Todd Black, Wenjun Zhao, Tze Ming Chan, and Andrew T. McPhail. Isolation and Structure Elucidation of Sch 642305, a Novel Bacterial DNA Primase Inhibitor Produced by Penicillium verrucosum. *J. Nat. Prod.*, 66(12):1527–1530, 2003.

[283] Hiranthi Jayasuriya, Deborah L. Zink, Jon D. Polishook, Gerald F. Bills, Anne W. Dombrowski, Olga Genilloud, Fernando F. Pelaez, Lucia Herranz, Donette Quamina, Russell B. Lingham, Renee Danzeizen, Pia L. Graham, Joanne E. Tomassini, and Sheo B. Singh. Identification of diverse microbial metabolites as potent inhibitors of HIV-1 Tat transactivation. *Chem. Biodivers.*, 2(1):112–122, 2005.

[284] Erica M. Wilson and Dirk Trauner. Concise synthesis of the bacterial DNA primase inhibitor (+)-Sch 642305. *Org. Lett.*, 9(7):1327–1329, 2007.

[285] Goverdhan Mehta and Harish M Shinde. Enantioselective synthesis of epi-(+)-sch 642305: observation of an interesting diastereoselection during rcm. *Tetrahedron letters*, 46(39):6633–6636, 2005.

[286] Goverdhan Mehta and Harish M Shinde. Enantioselective total synthesis of bioactive natural product (+)-sch 642305: a structurally novel inhibitor of bacterial dna primase and hiv-1 tat transactivation. *Chemical Communications*, (29):3703–3705, 2005.

[287] S Thongkantha, R Jeewon, D Vijaykrishna, S Lumyong, E H C Mc Kenzie, and K D Hyde. Molecular phylogeny of Magnaporthaceae (Sordariomycetes) with a new species Ophioceras chiangdaoense from Dracaena loureiroi in Thailand. *Fungal Divers.*, 34:157–173, 2009.

[288] KD Bussaban, B; Lumyong, S; Lumyong, P; McKenzie, EHC; Hyde. Endophytic fungi from Amomum siamense. *Can. J. Microbiol.*, 47(10):943–948, 2001.

[289] B. Bussaban, S. Lumyong, P. Lumyong, T. Seelanan, D.C. Park, E.H.C. McKenzie, and K.D. Hyde. Molecular and morphological characterization of Pyricularia and allied genera. *Mycologia*, 97(5):1002–1011, 2005.

[290] De B Scott and Amelia C Stolk. Studies on the genus eupenicillium ludwig ii. perfect states of some penicillia. *Antonie van Leeuwenhoek*, 33(1):297–314, 1967.

[291] VL Singleton, N Bohonos, and AJ Ullstrup. Decumbin, a new compound from a species of penicillium. *Nature*, 181(4615):1072, 1958.

[292] Vl Betina, P Nemec, J Dobias, and Z Barath. Cyanein, a new antibiotic from penicillium cyaneum. *Folia microbiologica*, 7(6):353, 1962.

[293] Yoshiyuki Suzuki, Hiroshi Tanaka, Hiroo Aoki, and Teiichi Tamura. Ascotoxin (decumbin), a metabolite of ascochyta imperfecta peck. *Agricultural and Biological Chemistry*, 34(3):395–413, 1970.

[294] Bu-Zhu Yu, Na Zhu, and Zhi-Zhi Du. Two new 7-dehydrobrefeldin a acids from cylindrocarpon obtusisporum, an endophytic fungus of trewia nudiflora. *Helvetica Chimica Acta*, 93(2):324–328, 2010.

[295] J Bernd Helms and James E Rothman. Inhibition by brefeldin a of a golgi membrane enzyme that catalyses exchange of guanine nucleotide bound to arf. *Nature*, 360(6402):352, 1992.

[296] V Betina. Biological effects of the antibiotic brefeldin a (decumbin, cyanein, ascotoxin, synergisidin): a retrospective. *Folia microbiologica*, 37(1):3–11, 1992.

[297] RG Coombe, PS Foss, JJ Jacobs, and TR Watson. The biosynthesis of brefeldin a. *Australian Journal of Chemistry*, 22(9):1943–1950, 1969.

[298] Clayton T Mabuni, Luigi Garlaschelli, Robert A Ellison, and C Richard Hutchinson. Biosynthetic origin of the oxygen atoms in the c15 macrolide antibiotic brefeldin a. *Journal of the American Chemical Society*, 99(23):7718–7720, 1977.

[299] Mario Gonzalez de la Parra and C Richard Hutchinson. Macrolide biosynthesis: stereochemistry of the hydroxylation of brefeldin c. *The Journal of antibiotics*, 40(8):1170–1174, 1987.

[300] Yoshikuni Yamamoto, Akira Hori, and C. Richard Hutchinson. Biosynthesis of Macrolide Antibiotics. 6. Late Steps in Brefeldin A Biosynthesis. *J. Am. Chem. Soc.*, 107(8):2471–2474, 1985.

[301] Angelica O. Zabala, Yit Heng Chooi, Moon Seok Choi, Hsiao Ching Lin, and Yi Tang. Fungal polyketide synthase product chain-length control by partnering thiohydrolase. *ACS Chem. Biol.*, 9(7):1576–1586, 2014.

[302] Emilie Adelin, Marie Thérèse Martin, Sylvie Cortial, Pascal Retailleau, Saisamorn Lumyong, and Jamal Ouazzani. Bioactive polyketides isolated from agar-supported fermentation of Phomopsis sp. CMU-LMA, taking advantage of the scale-up device, Platotex. *Phytochemistry*, 93:170–175, 2013.

[303] Emilie Adelin, Claudine Servy, Sylvie Cortial, Hélne Lévaique, Jean Franois Gallard, Marie Thérse Martin, Pascal Retailleau, Boonsom Bussaban, Saisamorn Lumyong, and Jamal Ouazzani. Biotransformation of natural compounds. Oxido-reduction of Sch-642305 by Aspergillus ochraceus ATCC 1009. *Bioorganic Med. Chem. Lett.*, 21(8):2456–2459, 2011.

[304] Robert D Finn, Teresa K Attwood, Patricia C Babbitt, Alex Bateman, Peer Bork, Alan J Bridge, Hsin-Yu Chang, Zsuzsanna Dosztányi, Sara El-Gebali, Matthew Fraser, et al. Interpro in 2017âĂŤbeyond protein family and domain annotations. *Nucleic acids research*, 45(D1):D190–D199, 2016.

[305] Philip Jones, David Binns, Hsin-Yu Chang, Matthew Fraser, Weizhong Li, Craig McAnulla, Hamish McWilliam, John Maslen, Alex Mitchell, Gift Nuka, et al. Interproscan 5: genome-scale protein function classification. *Bioinformatics*, 30(9):1236–1240, 2014.

[306] Ditlev E. Brodersen, William M. Clemons, Andrew P. Carter, Robert J. Morgan-Warren, Brian T. Wimberly, and V. Ramakrishnan. The structural basis for the action of the antibiotics tetracycline, pactamycin, and hygromycin B, on the 30S ribosomal subunit. *Cell*, 103(7):1143–1154, 2000.

[307] María Jesús Cabañas, David Vázquez, and Juan Modolell. Inhibition of ribosomal translocation by aminoglycoside antibiotics. *Biochemical and biophysical research communications*, 83(3):991–997, 1978.

[308] A Gonzalez, A Jimenez, D Vazquez, JE Davies, and D Schindler. Studies on the mode of action of hygromycin b, an inhibitor of translocation in eukaryotes. *Biochimica et Biophysica Acta (BBA)-Nucleic Acids and Protein Synthesis*, 521(2):459–469, 1978.

[309] Helge Bjorn Bode, Martina Walker, and Axel. Zeeck. Secondary metabolites by chemical screening, 41 structure and biosynthesis of mutolide, a novel macrolide from a UV mutant of the fungus F-24'707. *European J. Org. Chem.*, (8):1451–1456, 2000.

[310] John S Harwooda, Horace G Cutler, and John M Jacyno. Nigrosporolide, a plant growth-inhibiting macrolide from the mould nigrospora sphaerica. *Natural Product Letters*, 6(3):181–185, 1995.

[311] Yuquan Xu, Patricia Espinosa-Artiles, Vivien Schubert, Ya-ming Xu, Wei Zhang, Min Lin, AA Leslie Gunatilaka, Roderich Süssmuth, and István Molnár. Characterization of the biosynthetic genes for 10, 11-dehydrocurvularin, a heat shock response-modulating anticancer fungal polyketide from aspergillus terreus. *Applied and environmental microbiology*, 79(6):2038–2047, 2013.

[312] N Mahanti, D Bhatnagar, JW Cary, J Joubran, and JE Linz. Structure and function of fas-1a, a gene encoding a putative fatty acid synthetase directly involved in aflatoxin biosynthesis in aspergillus parasiticus. *Applied and environmental microbiology*, 62(1):191–195, 1996.

[313] DW Brown, TH Adams, and NP Keller. Aspergillus has distinct fatty acid synthases for primary and secondary metabolism. *Proceedings of the National Academy of Sciences*, 93(25):14873–14877, 1996.

[314] David H Kwan and Frank Schulz. The stereochemistry of complex polyketide biosynthesis by modular polyketide synthases. *Molecules*, 16(7):6092–6115, 2011.

[315] Christian Hertweck. The biosynthetic logic of polyketide diversity. *Angewandte Chemie International Edition*, 48(26):4688–4716, 2009.

[316] Kanichiro Ishiuchi, Takehito Nakazawa, Fumitoshi Yagishita, Takashi Mino, Hiroshi Noguchi, Kinya Hotta, and Kenji Watanabe. Combinatorial generation of complexity by redox enzymes in the chaetoglobosin a biosynthesis. *Journal of the American Chemical Society*, 135(19):7371–7377, 2013.

[317] Dirk W Heinz, Walter A Baase, Frederick W Dahlquist, and Brian W Matthews. How amino-acid insertions are allowed in an α-helix of t4 lysozyme. *Nature*, 361(6412):561, 1993.

[318] Jeffrey B Bonanno, Steven C Almo, Anne Bresnick, Mark R Chance, Andras Fiser, S Swaminathan, J Jiang, F William Studier, Lawrence Shapiro, Christopher D Lima, et al. New york-structural genomix research consortium (nysgxrc): a large scale center for the protein structure initiative. *Journal of structural and functional genomics*, 6(2-3):225–232, 2005.

[319] Yeast Plasmid Miniprep. Zymoprep - 96 Yeast Plasmid Miniprep. (949).

185

6 *Curriculum Vitae*

Oct 2014 - May 2018	**PhD student** in Organic Chemistry, Gottfried Wilhelm Leibniz Universität Hannover, Germany
Sep 2011 - Dec 2013	**Master degree** in Molecular Biology, *Alma Mater Studiorum*, Bologna, Italy
Sep 2008 - Jul 2011	**Bachelor degree**, Università di Trento, Facoltà di Scienze MM. FF. NN., Italy
Jun 2008	**Diploma**, Liceo Scientifico Galilei Galileo, Trento, Italy