

# **Investigating and Engineering Fungal Tropolone Meroterpenoid Biosynthesis**

Von der Naturwissenschaftlichen Fakultät der  
Gottfried Wilhelm Leibniz Universität Hannover

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

von  
**Lei Li, Master of Natural Science (China)**

**2022**

Referent: Prof. Dr. Russell J. Cox

Korreferent: PD Dr. rer. nat. habil. Carsten Zeilinger

Tag der Promotion: 22.02.2022

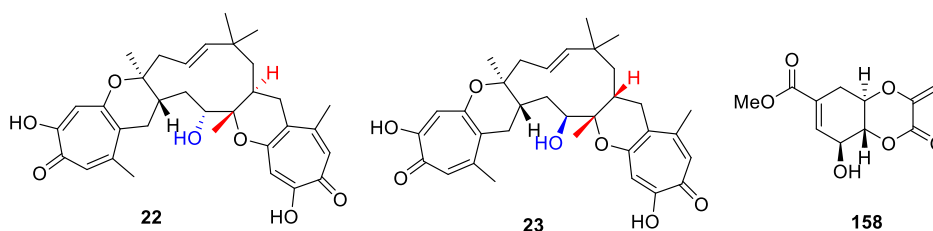
## Abstract

Keywords: tropolone sesquiterpenoid, scytolide, biosynthesis, shikimate pathway

The main focus of the presented work concentrated on understanding the biosynthetic pathways of tropolone sesquiterpenoids (TS) and scytolide in fungi.

Tropolone sesquiterpenoids, such as pycnidione **22** and eupenifeldin **23**, share the structural motif of a core 11-membered macrocycle derived from humulene, which is connected to one or two polyketide-derived tropolones *via* bridging dihydropyran rings. The whole genome of fungus *CF236968* was sequenced, which allowed *in silico* prediction of the *pyc* BGC as a good candidate for pycnidione-type TS. In parallel work from our group the *eup2* BGC was predicted as a promising candidate for eupenifeldin-type TS in the fungus *CF150626*. The *eup2* BGC was connected to eupenifeldin **23** biosynthesis based on reverse transcriptase-polymerase chain reaction (RT-PCR) experiments. Ten homologous genes between the *pyc* BGC and the *eup2* BGC were detected and five among them were shown to encode proteins homologous to the xenovulene A **24** biosynthetic pathway in *Acremonium strictum*. *In vitro* assays with Eup L4, a short chain dehydrogenase, indicated that it catalyzes the reduction of stipitaldehyde **102** and is necessary for eupenifeldin **23** production.

Scytolide **158** is a promising natural herbicide that is proposed to be derived biogenetically via lactonization of the enol pyruvate of methyl 3-epishikimate **205**. Feeding [2-<sup>13</sup>C]-glycerol to *Phyllosticta cirsii* confirmed that scytolide **158** is derived from the shikimate pathway. Looking for the homologous genes of pentafunctional aromatic polypeptide (AROM) in *P. cirsii*. and *T. reesei* *Atmus53* resulted in a promising BGC candidate for scytolide **158** and a novel biosynthetic pathway partly related to the shikimate pathway was proposed based on this BGC.



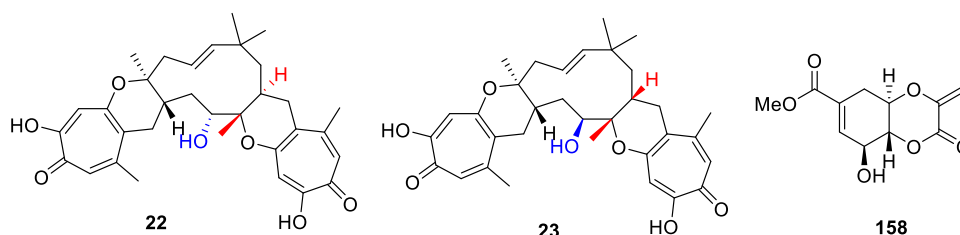
## Zusammenfassung

Schlagwörter: Tropolonesesquiterpenoid, Scytolide, Biosynthese, Shikimatweg

Der Schwerpunkt der vorliegenden Arbeit lag auf dem Verständnis des Biosynthesewegs von Tropolonesesquiterpenoiden (TS) und Scytolide in Pilzen.

Tropolonesesquiterpenoide wie Pycnidione **22** und Eupenifeldin **23** teilen sich das Strukturmotiv eines von Humulen abgeleiteten elfgliedrigen Makrocyclus. Dieser ist über Dihydropyranringe mit einem oder zwei von Polyketiden abstammenden Tropolonen verbunden. Das Genom des Pilzes *CF236968* wurde sequenziert, was eine In-silico-Vorhersage des *pyc*-BGC als einen guten Kandidaten für TS vom Pycnidione-Typ in *CF236968* ermöglichte. In parallelen Arbeiten in der Cox-Gruppe wurde der *eup2*-BGC als vielversprechende Kandidat für TS vom Eupenifeldin-Typ in *CF150626* vorhergesagt. Der *eup2*-BGC wurde mit der Biosynthese von Eupenifeldin **23** durch Reverse-Transkriptase-Polymerasekettenreaktion (RT-PCR) in Verbindung gebracht. Zehn homologe Gene zwischen dem *pyc*-BGC und dem *eup2*-BGC wurden nachgewiesen, und fünf davon kodieren für Proteine, die homolog zum Biosyntheseweg von Xenovolene A **24** in *Acremonium strictum* sind. In-vitro-Assays mit Eup L4, einer kurzkettigen Dehydrogenase, zeigten, dass es die Reduktion von Stipitaldehyde **102** katalysiert und für die Eupenifeldin **23** Produktion notwendig ist.

Scytolide **158** ist ein vielversprechendes natürliches Herbizid, das biogenetisch durch Lactonisierung des Enolpyruvats von Methyl-3-epishikimate **205** gewonnen werden könnte. Die Verfütterung von [2-<sup>13</sup>C]-Glycerin an *Phyllosticta cirsii* bestätigte, dass Scytolide **158** aus dem shikimischen Stoffwechselweg stammt. Die Suche nach homologen Genen des pentafunktionellen aromatischen Polypeptids (AROM) in *P. cirsii*. und *T. reesei* *Atmus53* führten zu einem vielversprechenden BGC Kandidaten für die Biosynthese von Scytolide **158**. Basierend auf diesem Kandidaten wurde ein neuer Biosyntheseweg vorgeschlagen, der teilweise mit dem Shikimat-Weg verwandt ist.





## Acknowledgement

First of all, I would like to thank my supervisor Prof. Russell Cox. Thank you so much for offering me this precious opportunity in studying in Leibniz University Hannover and providing two interesting projects. Thank you for your patient instructions and constructive advice for my lab work over the past four years. Thank you for your sustained help and quick feedback during my thesis writing. I am super impressed by your open-minded, positive enthusiastic attitude towards scientific research. I will always remember what you say: Make a place better than before.

I would like to thank Dr. Carsten Schotte. Thank you for showing me how to work in the BMWZ especially in microbiology and molecular biology experimental parts at the beginning of my PhD study. Working on the tropolone sesquiterpenoids project with you was a valuable experience for me. I really appreciate your kindness and practical suggestions when I got stuck with my projects. I would like to thank Dr. Elizabeth Skellam for your help in bioinformatic analysis and experimental instructions in my scytolide project. And I would like to thank Hanen Zamel for your wonderful metabolite extraction and analysis work in my scytolide project.

Thank all BMWZ media kitchen team for providing reliable support for my research. Thanks all analytical departments in OCI. Thanks for your generous help.

Thank all past and current Cox group members (Eric, Francesco, Eman, Dongsong, Jin, Haili, Sen, Chongqing, Hao, Tian, Yunlong, Jing, Mary, Lukas, Verena, Vjaceslavs, Oliver, Steffen, Katharina, Henrike, Mautice and Evgeniya.) for all the communications about work or life.

Thanks, China Scholarship Council (CSC) for the foundation

Finally, I would like to thank my parents and friends for your care, love and support in life.

## Abbreviations and Units

AROM	pentafunctional aromatic polypeptide	LCMS	liquid chromatography mass spectrometry
<i>amp</i> <sup>R</sup>	ampicillin resistance gene	L-Phe	L-phenylalanine
BLAST	basic local alignment search tool	L-Trp	L-tryptophan
bp	base pair	L	Litre
BGC	biosynthetic gene cluster	μL	microliter
CDCl <sub>3</sub>	deuterated chloroform	MeT	Methyltransferases
COSY	homonuclear correlation spectroscopy	mRNA	messenger ribonucleic acid
CD <sub>3</sub> OD	deuterated methanol	MS	mass spectrometry
cDNA	complementary DNA	MeCN	acetonitrile
<i>cam</i> <sup>R</sup>	chloramphenicol gene	mg	milligram
DNA	deoxyribonucleic acid	min	minute
d	days	mL	millilitre
DAD	diode-array detection		
DHQ	3-dehydroquinic acid	M	molar
DHS	3-dehydroshikimic acid	<i>m/z</i>	mass to charge ratio
DAHP	3-deoxy-D-arabino-heptulosonic acid 7-phosphate	NPP	nerolidyl diphosphate
ES <sup>+</sup>	positive electrospray ionization	NMR	nuclear magnetic resonance
ES <sup>-</sup>	negative electrospray ionization	nm	nanometer
EIC	extracted ion chromatogram	NADPH	Nicotinamide adenine dinucleotide phosphate
EtOAc	ethyl acetate	<i>o</i> -QM	<i>ortho</i> -quino methide
E4P	D-erythrose-4-phosphate	ORF	open reading frame
ELSD	evaporative light scattering detector	<i>O</i> -MeT	<i>O</i> -methyltransferase
EPSP	5-enolpyruvylshikimic acid 3-phosphate	PKS	polyketide synthase
FPP	farnesylpyrophosphate	PCR	polymerase chain reaction
gDNA	genomic DNA	P <sub><i>gpdA</i></sub>	<i>GpdA</i> promoter
hDA	hetero Diels-Alder	P <sub><i>amyB</i></sub>	<i>AmyB</i> promoter
hDAase	hetero Diels-Alderase	PEP	phosphoenolpyruvic acid
HMBC	heteronuclear multiple bond correlation	PEG	polyethleneglycol
HPLC	high performance liquid chromatography	rpm	revolutions per minute
HSQC	heteronuclear single quantum coherence	<i>p</i> -CA	<i>p</i> -Coumaric acid
Hz	Hertz	RT-PCR	reverse transcription polymerase chain reaction
h	hours	rRNA	ribosomal RNA
4-HBA	4-hydroxybenzoic acid	RT	retention time
ITS	internal transcribed spacer	SA	shikimic acid
<i>IC</i> <sub>50</sub>	half maximal inhibitory concentration	SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
IPTG	isopropylthio-β-galactoside	TS	tropolone sesquiterpenoids
kDa	kilo Dalton	UV	ultraviolet
KO	knockout	WT	wild-type

# Contents

<b>Abstract</b> .....	i
<b>Zusammenfassung</b> .....	ii
<b>Acknowledgement</b> .....	iii
<b>Abbreviations and Units</b> .....	iv
<b>Contents</b> .....	v
<b>1 Introduction</b> .....	1
1.1 Secondary Metabolites from Fungi.....	1
1.2 Meroterpenoids in fungi.....	2
1.2.1 Meroterpenoids Derived from Triketides.....	3
1.2.2 Meroterpenoids Derived from Tetraketides.....	4
1.2.3 Tropolone Sesquiterpenoids and Related Compounds.....	6
1.2.4 Biomimetic Synthesis of Tropolone Sesquiterpenoids.....	10
1.2.5 Biosynthesis of Tropolone Sesquiterpenoids.....	15
1.2.6 Biosynthesis of Tropolone and Humulene.....	24
1.3 Natural products derived from the shikimate pathway.....	26
1.4 Scytolide and Related Natural Products.....	31
1.5 Method for biosynthetic studies.....	32
1.5.1 Genome Sequencing.....	32
1.5.2 Fungal Transformation and Gene Knockout.....	33
1.5.3 Heterologous Expression.....	36
1.5.4 Enzyme Assay.....	38
1.5.5 Isotopic Feeding Experiments.....	38
<b>2 Biosynthesis of Tropolone Sesquiterpenoids</b> .....	40
2.1 Project Aims.....	40
2.2 Results-Analysis of Tropolone sesquiterpenoids Produced by Fungi.....	41
2.2.1 Analysis of Secondary Metabolites from <i>CF236968</i> .....	41
2.2.2 Structural Identification of Pycnidione <b>22</b> and Epolone C <b>42</b> .....	43
2.2.3 Analysis of Secondary Metabolites from <i>CF253087</i> .....	48

2.2.4	Structural Identification of Eupenifeldin <b>23</b> and Neosetophomone B <b>61</b> .....	49
2.2.5	Analysis of Secondary Metabolites from <i>Pleosporales sp 1987</i> .....	54
2.2.6	Timecourse of Metabolite Production from <i>CF236968</i> .....	56
2.2.7	Structural Identification of violaceol I <b>167</b> , cordyol C <b>168</b> and diorcinol <b>169</b> ...	58
2.3	Results-ITS Characterization of <i>CF 236968</i> and <i>CF253087</i> .....	61
2.4	Results- <i>In silico</i> Analysis of Putative BGC for Tropolone Sesquiterpenoids .....	63
2.4.1	Whole Genome Sequencing of <i>CF236968</i> .....	63
2.4.2	<i>In silico</i> Analysis of Putative BGC for Pycnidione .....	63
2.4.3	<i>In silico</i> Analysis of Putative BGC for Eupenifeldin .....	66
2.5	Results-RT-PCR Analysis for Eupenifeldin BGC in Fungus <i>CF1506262</i> .....	68
2.6	Results-Antibiotic Sensitivity of Fungi <i>CF236968</i> and <i>CF253087</i> .....	71
2.7	Results-Developing a Transformation Protocol for <i>CF150626</i> .....	72
2.8	Results-Heterologous Expression of <i>asL4</i> and <i>asL6</i> in <i>A. oryzae</i> .....	75
2.9	Results- <i>In vitro</i> Activity Tests with Eup2L4, Eup2R1 and PycR1 .....	78
2.9.1	Expression and purification of protein Eup2L4 and Eup2R1 .....	80
2.9.2	Stipitaldehyde Purification .....	82
2.9.3	<i>In vitro</i> assay of Eup2L4 .....	84
2.9.4	<i>In vitro</i> assay of Eup2R1 .....	90
2.9.5	<i>In vitro</i> Assay of PycR1.....	93
2.10	Conclusion and Discussion .....	99
<b>3</b>	<b>Biosynthesis of scytolide in fungi</b> .....	<b>111</b>
3.1	Project aims .....	111
3.2	Results-Analysis of Scytolide <b>158</b> Produced by <i>P. cirsii</i> and <i>T. reesei Δtmus53</i> .....	112
3.2.1	Analysis of Secondary Metabolites from <i>P. cirsii</i> .....	112
3.2.2	Structural Identification of Scytolide <b>158</b> .....	114
3.2.3	Structural Identification of Methylscytolide <b>198</b> .....	115
3.2.4	Analysis of Secondary Metabolites from <i>T. reesei Δtmus53</i> .....	118
3.2.5	Timecourse of Metabolite Production from <i>P. cirsii</i> .....	120
3.2.6	Timecourse of Metabolite Production from <i>T. reesei Δtmus53</i> .....	122
3.3	Results-Glycerol Feeding Experiments.....	123

3.4	Results-Genome of <i>P. cirsii</i> and <i>T. reesei Atmus53</i> .....	127
3.4.1	Shikimic Acid Pathway in Fungi .....	127
3.4.2	<i>In silico</i> Analysis of Putative BGC for Scytolide .....	127
3.4.3	<i>In silico</i> Analysis of Methyltransferase .....	135
3.5	Results-RT-PCR Analysis for the Interesting Genes from <i>P. cirsii</i> .....	136
3.6	Results-Heterologous Expression of <i>PA4</i> in <i>A. oryzae</i> .....	139
3.7	Results-Targeted Gene Knockout in <i>P. cirsii</i> .....	142
3.7.1	Plasmid Construction .....	142
3.7.2	Gene Knockout Targeting <i>PA4</i> and <i>PA6</i> .....	143
3.8	Results-Targeted Gene Knockout in <i>T. reesei Atmus53</i> .....	148
3.8.1	Transformation Protocol.....	149
3.8.2	Gene Knockout Targeting <i>TA1</i> , <i>TA4</i> , and <i>TA5</i> .....	151
3.9	Conclusion and Outlook.....	153
<b>4</b>	<b>Experimental</b> .....	158
4.1	Growth Media, Buffers and Solutions .....	158
4.2	Enzymes and Antibiotics .....	161
4.3	Microbiology methods .....	161
4.3.1	<i>E. coli</i> growth and transformation .....	162
4.3.2	<i>Saccharomyces cerevisiae</i> growth and transformation.....	162
4.3.3	<i>Aspergillus oryzae</i> NSAR1 growth and transformation .....	163
4.3.4	<i>CF150626</i> growth and transformation .....	164
4.3.5	<i>P. cirsii</i> growth and transformation .....	165
4.3.6	<i>T. reesei Atmus53</i> growth and transformation .....	166
4.4	Molecular biology methods.....	167
4.4.1	Oligonucleotides.....	167
4.4.2	Obtained and constructed vectors in this thesis.....	171
4.4.3	Polymerase chain reaction (PCR).....	172
4.4.4	RNA Extraction and cDNA Preparation.....	172
4.4.5	Agarose gel electrophoresis.....	172
4.4.6	DNA purification from gel or PCR .....	173

4.4.7	Isolation of plasmid DNA from <i>E. coli</i> .....	173
4.4.8	Isolation of plasmid DNA from <i>S. cerevisiae</i> .....	173
4.4.9	Isolation of genomic DNA from Fungi .....	173
4.4.10	DNA sequencing .....	173
4.4.11	Determination of DNA Concentration.....	173
4.5	Heterologous Protein Production and Purification.....	174
4.5.1	Production of Protein in <i>E. coli</i> .....	174
4.5.2	Cell Lysis and Protein Extraction.....	174
4.5.3	Histidine-tagged Protein Purification by Ni-NTA.....	174
4.5.4	Protein Separation Using SDS - Polyacrylamide Gel Electrophoresis.....	175
4.5.5	Determination of Protein Concentration .....	176
4.6	Chemical Analysis.....	176
4.6.1	Extraction of <i>A. oryzae</i> cultures .....	176
4.6.2	Extraction of <i>CF236968</i> and <i>CF253087</i> .....	176
4.6.3	Extraction of <i>CF150626</i> .....	176
4.6.4	Extraction of <i>P. cirsii</i> and <i>T. reesei Atmus53</i> .....	177
4.6.5	Extraction of <i>T. stipitatus AtsL2</i> .....	177
4.6.6	Liquid Chromatography Mass Spectrometry (LCMS) Analysis .....	177
4.6.7	Nuclear Magnetic Resonance (NMR) Analysis.....	178
<b>References</b> .....		179
<b>Appendix</b> .....		187
<b>Curriculum vitae</b> .....		ix
<b>List of Publications</b> .....		ix

# 1 Introduction

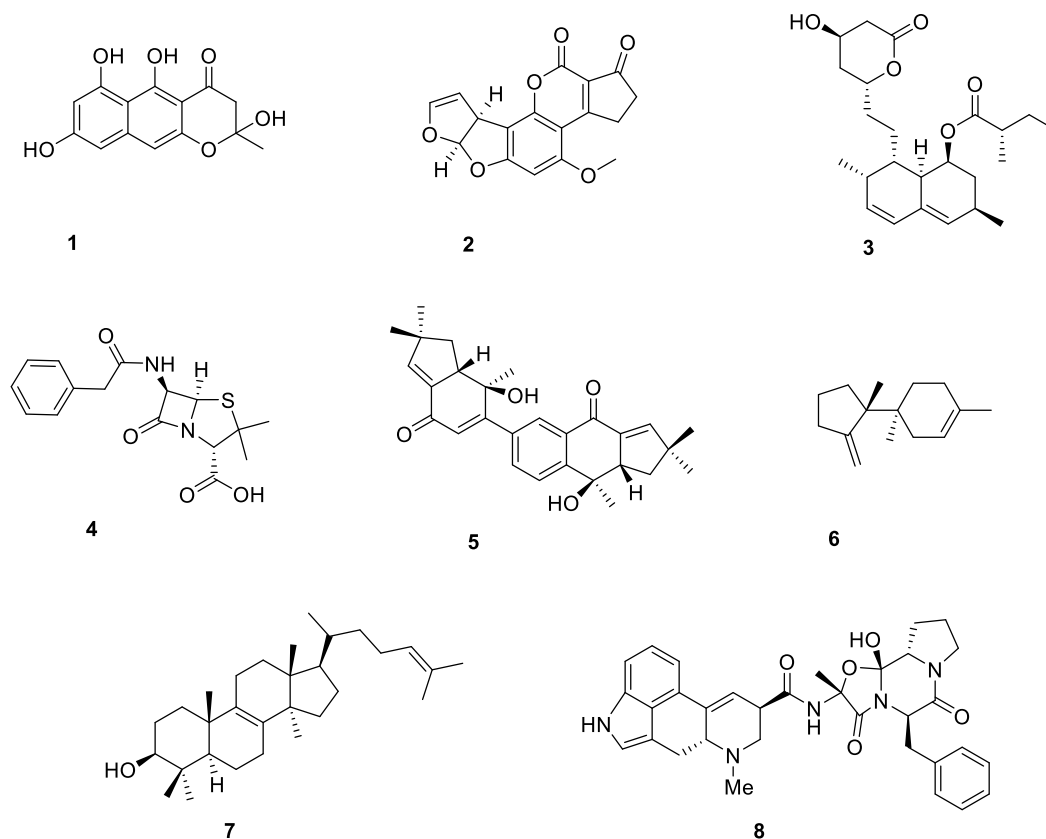
Natural products are a considerably abundant library of compounds with diverse chemical structures and biological activities derived from nature. They are usually divided into two different classes: primary metabolites and secondary metabolites. Primary metabolites (*e.g.* amino acids or sugars) are necessary for the growth, development and reproduction of living organism. Secondary metabolites are not essential for life, and are usually thought to provide some form of advantage or protection to the producers. Diversity in chemical structures and biological activities make secondary metabolites a valuable source for the pharmaceutical industry to find the new molecules for medicine.<sup>1</sup> The genes encoding the biosynthesis of secondary metabolites tend to gather contiguously on the chromosome in coregulated gene clusters, which are easily identified with the development in genome sequencing, computational tools, and analytical chemistry.<sup>2</sup>

## 1.1 Secondary Metabolites from Fungi

The systematic study of fungal secondary metabolites started from 1922 under the leadership of Harold Raistrick.<sup>3</sup> It was not until the discovery and development of penicillin that widespread attention was focused on fungal metabolites.<sup>4</sup> Fungal secondary metabolism encompasses a large group of structurally diverse natural products and it is a rich source of biologically active compounds, which are classified to four main types according to their different biosynthetic precursors and involved enzymes.<sup>5</sup>

Polyketides, synthesized by iterative type I polyketide synthases (PKS), are among the most abundant fungal secondary metabolites.<sup>6</sup> Spore-pigment intermediate naphthopyrone **1**, the carcinogen aflatoxin B1 **2** and cholesterol-lowering compound lovastatin **3** are representative examples. Non-ribosomal peptides are derived from both proteinogenic amino acids and non-proteinogenic amino acids by multidomain, multi modular nonribosomal peptide synthetases (NRPS).<sup>7</sup> The most famous example is penicillin G **4** that is derived from a tripeptide. Terpenes produced by fungi, including sterostrein A **5**, trichodiene **6**, and lanosterol **7**, are generated from different poly-isoprene diphosphates. Their backbones are built up with different numbers of isoprene units.<sup>8</sup> Indole alkaloids are usually derived from tryptophan **131** and dimethylallyl

diphosphate, although sometimes amino acids other than tryptophan are used as precursors. One classic example is ergotamine **8**, which has biological activity as a vasoconstrictor and is used medicinally for treatment of acute migraine attacks (Figure 1.1).<sup>9</sup>



**Figure 1.1** Representative compounds of fungal natural products from different classes (polyketides, nonribosomal peptides, terpenoids and indole alkaloids).

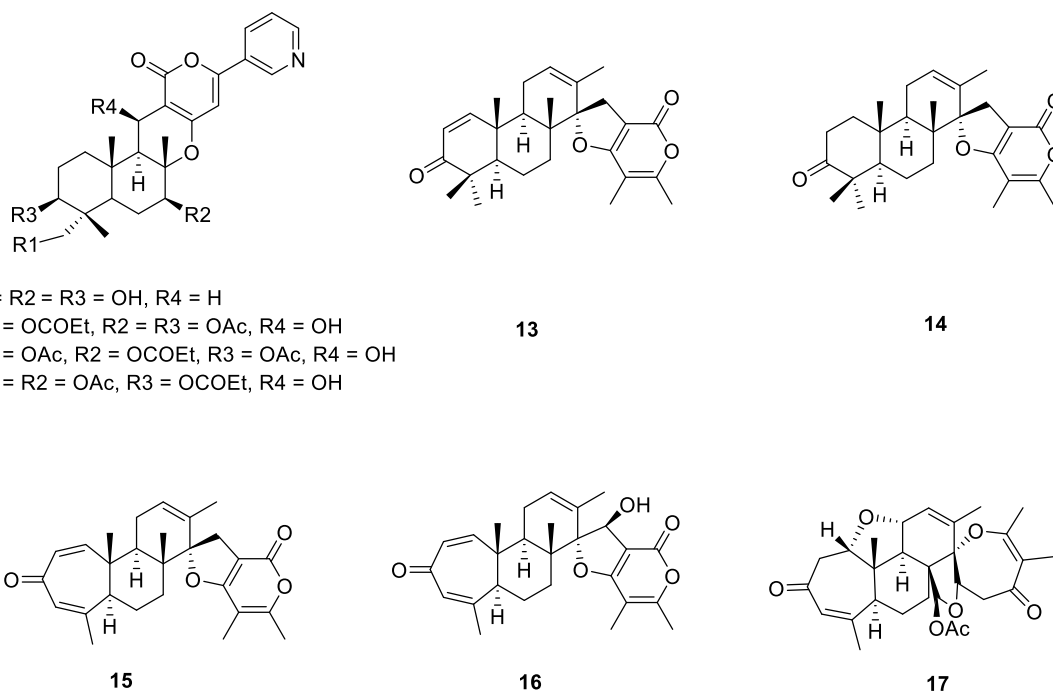
## 1.2 Meroterpenoids in fungi

Meroterpenoids are natural products of mixed-biogenesis, and typically originate from the terpenoid and polyketide/shikimate/indole pathways.<sup>10</sup> They are widely distributed in nature and have been obtained from fungi, marine organisms and higher plants. Meroterpenoids display a huge range of structural diversity, and most of them possess abundant and important biological activities. Based on their biosynthetic origins, meroterpenoids can be grouped into two major classes: polyketide-terpenoids and non-polyketide-terpenoids. Polyketide meroterpenoids are further classified according to the number of acyl units which have contributed to the polyketide chain.<sup>11</sup>



### 1.2.1 Meroterpenoids Derived from Triketides

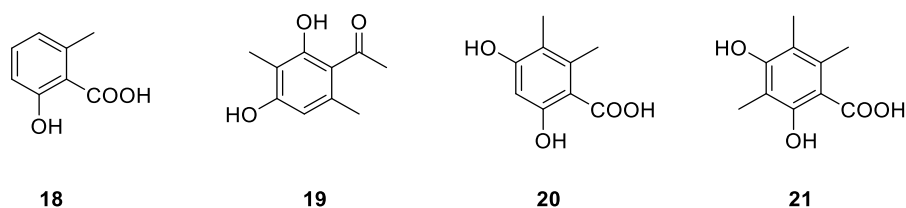
Pyripyropenes are meroterpenoids containing triketide–terpenoid moieties, which have a common skeleton consisting of a polyoxygenated sesquiterpene,  $\alpha$ -pyrone and pyridine moieties.<sup>12</sup> Pyripyropenes A–D **9–12** isolated from the fermentation broth of *Aspergillus fumigatus* FO-1289 show very potent Acyl-CoA:cholesterol acyltransferase inhibitory activity.<sup>13,14</sup> Other compounds of the same type include the naturally occurring bioactive spiro-diterpenoids (+)-breviones A–E (**13–17**). They were isolated from extracts of semi-solid fermented *Penicillium brevicompactum* Dierckx. These compounds inhibit wheat coleoptile growth, indicating their potential use as new agrochemicals (Figure 1.2.1).<sup>15,16</sup>



**Figure 1.2.1** Representative meroterpenoids derived from triketide.

## 1.2.2 Meroterpenoids Derived from Tetraketides

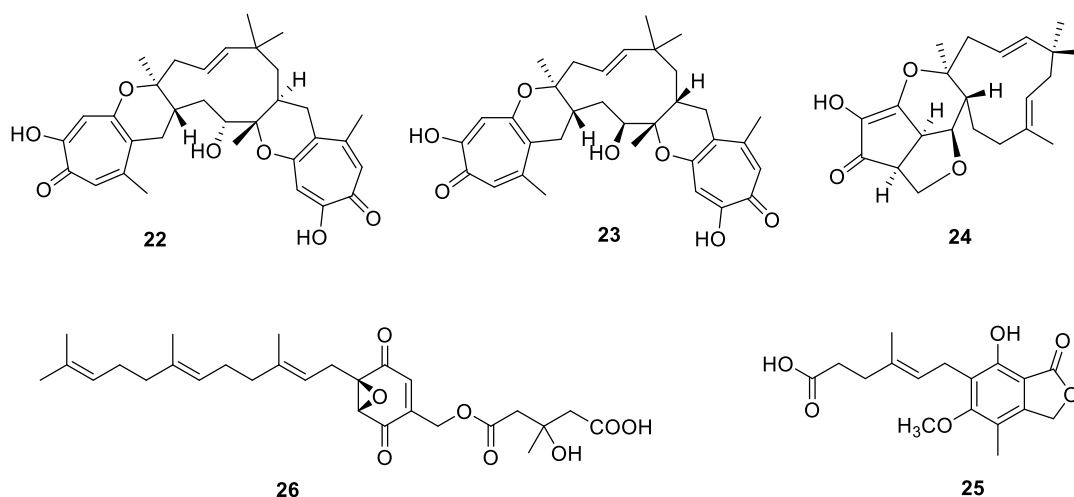
Tetraketide-terpenoid derived compounds constitute the largest class of fungal meroterpenoids. They can be conveniently subdivided according to different tetraketide precursor, which include 6-methylsalicylic acid **18**, 3-methylorcinolaldehyde **19**, 5-methylorsellinic acid **20**, and 3,5-dimethylorsellinic acid **21**. The methyl groups of compounds **19**, **20**, **21** are derived from *S*-adenosylmethionine (SAM) (Figure 1.2.2.1).<sup>10</sup>



**Figure 1.2.2.1** Different tetraketide precursors for meroterpenoids.

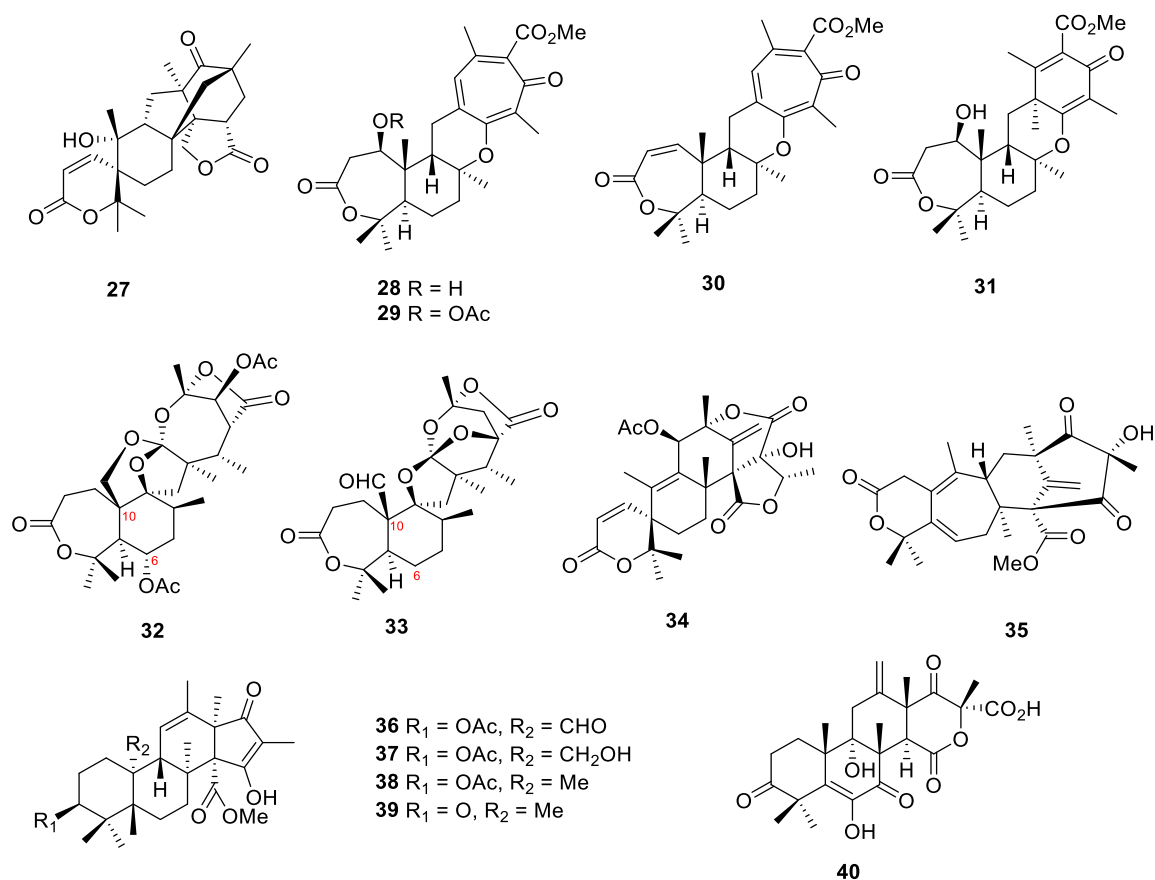
Pycnidione **22** was first isolated from cultures of *Phoma* sp. as a stromelysin inhibitor. Eupenifeldin **23** was purified from *Eupenicillium brefeldianum* and it shows cytotoxic activity against the HCT-116 cell line as well as *in vivo* antitumor activity in the P388 leukemia model.<sup>17,18</sup> The tropolone rings in these meroterpenoids are derived from the polyketide precursor 3-methylorcinolaldehyde **19**.<sup>19</sup> In addition, xenovulene A **24** isolated from *Acremonium strictum* contains a cyclopentenone moiety that is proved to be derived from 3-methylorcinolaldehyde **19** by a ring-expansion/ring-contraction mechanism.<sup>20</sup> It is an inhibitor of benzodiazepine binding to the GABA-benzodiazepine receptor and has potential use as an antidepressant with reduced addictive properties.

Mycophenolic acid **25** is reported from several species of *Penicillium*, which have been demonstrated to consist of an acetate-derived aromatic nucleus and a linear sesquiterpenoid side-chain. Due to extensive biological properties such as antiviral, antifungal, antibacterial, antitumor, and immunosuppressive, several efforts have been made to investigate its biosynthesis. 5-Methylorsellinic acid **20** produced by nr-PKS MpaC, is identified to be the polyketide precursor of mycophenolic acid **25**.<sup>21</sup> Yanuthone D **26**, isolated from *Aspergillus niger*, displays potent antifungal activity, its polyketide precursor is 6-methylsalicylic acid **18** produced by PR PKS YanA (Figure 1.2.2.2).<sup>22-23</sup>



**Figure 1.2.2.2** Representative meroterpenoids derived from 3-methylorcinaldehyde **19** and 6-methylsalicylic acid **18**.

Fungal meroterpenoids derived from 3,5-dimethylorsellinic acid **21** display huge structural diversity. It is noteworthy that they possess complex and intriguing skeletons. Andibenin **27** is the first compound belong to this group, discovered in 1976.<sup>24</sup> It is unique in the position of attachment of the sesquiterpene unit and the introduction of two alkyl substituents on the same carbon of the polyketide unit, and in having more than one carbon-carbon bond between the terpenoid and polyketide units. Four cytotoxic meroterpenoids tropolactones A–D **28–31**, containing an interesting substituted 2,4,6-cycloheptatriene (tropone) ring, were isolated from the whole broth extract of a marine-derived fungus of the genus *Aspergillus*.<sup>25</sup> Fumigatonin **32** and novofumigatonin **33** are isolated from *Aspergillus fumigatus* and *Aspergillus novofumigatus*, respectively.<sup>26,27</sup> Both of them contains the unusual orthoester moiety, while the major differences between them are the presence of an aldehyde substituent on C-10 in **33** (where **32** has an oxygen-bound methylene group) and the lack of an acetate group at C-6 in **33**. In addition, unique and abundant skeletons of the meroterpenoids are also exemplified by the structure of austin **34**,<sup>28</sup> berkeleydione **35**,<sup>29</sup> andrastins A–D **36–39**,<sup>30–31</sup> and terretonin **40** (Figure 1.2.2.3).<sup>32</sup> Interestingly, all these compounds share the same intermediate which is formed by the alkylation of a *bis-C*-methylated 3,5-dimethylorsellinic acid **21**. So, it is the presence of many cyclization patterns of the terpenoid moiety and a variety of tailoring reactions that results in the large variety of compounds in this family.



**Figure 1.2.2.3** Representative meroterpenoids derived from 3, 5-dimethylorsellinic acid **21**.

### 1.2.3 Tropolone Sesquiterpenoids and Related Compounds

Tropolone sesquiterpenoids, such as eupenifeldin **23** and pycnidione **22**, share the structural motif of a core 11-membered macrocycle derived from humulene, which is connected to one or two polyketide-derived tropolones *via* dihydropyran rings. Eupenifeldin **23** and pycnidione **22** were first isolated from *Phoma sp* and *Eupenicillium brefeldianum ATCC741* respectively in 1993. After that, a few analogues have also been reported as the co-metabolites from a series of different fungi.

As a class, natural tropolone sesquiterpenoids (TS) have been reported mainly from fungi and can be grouped into two major subclasses based on their stereochemistry: Pycnidione-type TS and Eupenifeldin-type TS. Within the pycnidione series, both pycnidione **22** and epolone A **41** are bistropolone sesquiterpenoids containing a 4*R* hydroxyl group and *anti*-relationship at the eastern dihydropyran ring junction, while epolone B **44** is a mono-tropolone sesquiterpenoid that bears a 4*R*

hydroxyl group and *2E*-alkene without the eastern dihydropyran ring junction. The diastereomeric eupenifeldin-type TS maintain an identical trend among compounds but are differentiated from the former by *4S* hydroxyl group, the *syn*-relationship at the eastern dihydropyran ring junction in bistropolone sesquiterpenoids and *2Z*-alkene in the mono-tropolone sesquiterpenoids.

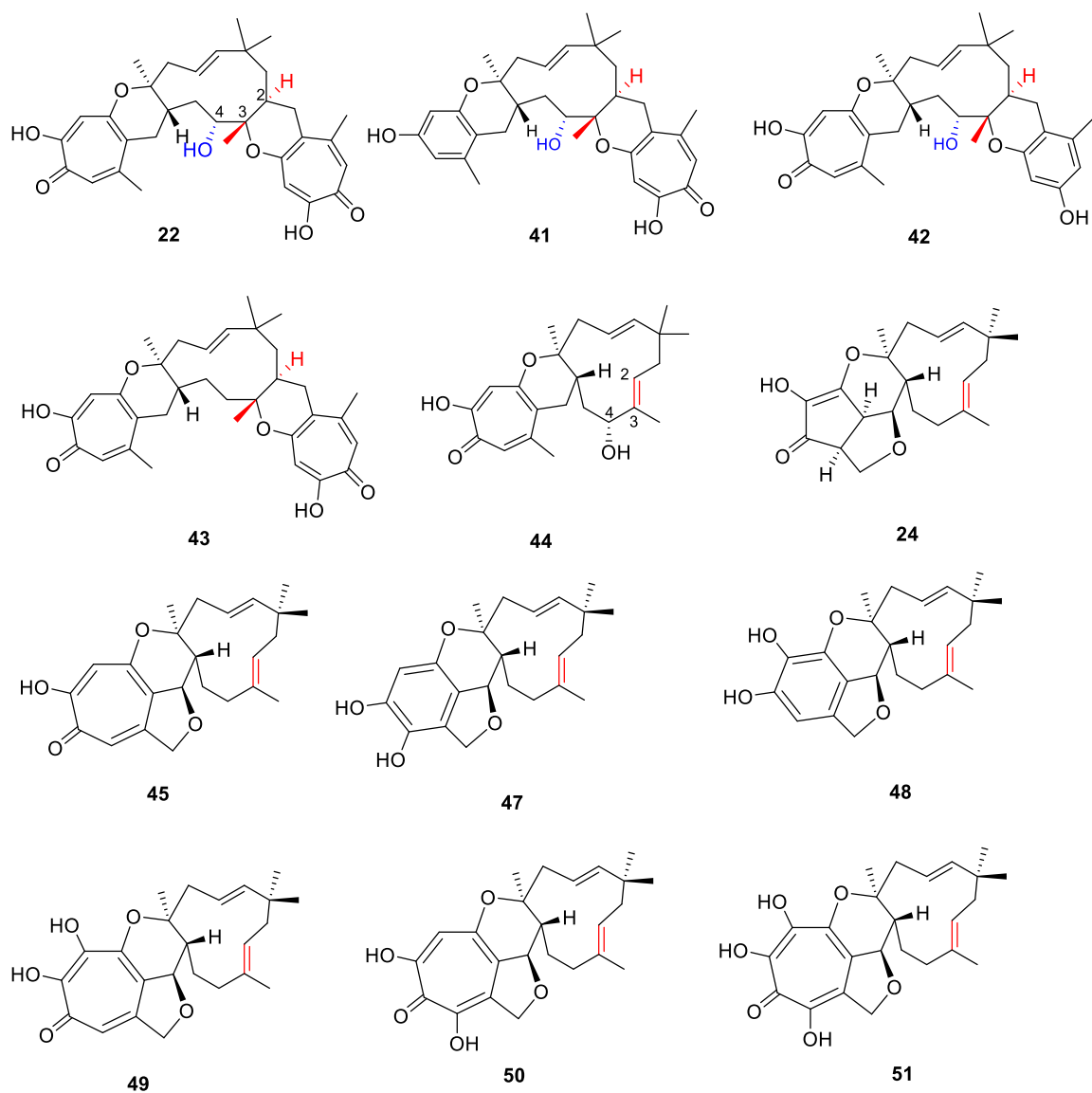


Figure 1.2.3.1 Pycnidione-type TS isolated from different fungi.

In the course of screening for small molecule modulators of erythropoietin gene expression, epolones A **41** and B **44** together with the known pycnidone were discovered from a culture of *OS-F69284*.<sup>33</sup> In a study investigating the production of antimicrobial natural products from fungi associated with sea foam, epolones A **41** and B **44**, dehydroxypycnidione **43**, together with pycnidione **22** were purified from a strain of *Neosetophoma samarorum*.<sup>34</sup> Additionally, xenovulene A **24** and xenovulene B **45** together with a series of related TS compounds **47-51** were isolated from fungi *Acremonium strictum* IMI 501407.<sup>20</sup> The cyclopentenone moiety of **24** is biologically derived from the ring-contraction of a tropolone. Since xenovulene A **24** contains a core 11-membered macrocycle with a 2*E*-alkene, we propose that it belongs to the pycnidione-type TS. In 2020, epolone C **42** and pycnidione **22** were reported from the undescribed fungus *CF-236968* together with their biosynthetic gene cluster (Figure 1.2.3.1).<sup>35</sup>

In 2002, pughinin A **52** and eupenifeldin **23** were isolated from the seed fungus *Kionochaeta pughii* BCC 3878. Both **52** and **23** exhibit *in vitro* antiplasmodial activity against *Plasmodium falciparum* (K1 strain) with *IC*<sub>50</sub> values of 2.4-0.28 µg/ml. In addition, eupenifeldin **23** also shows anti-cancer activity against human epidermoid carcinoma and human breast cancer cells.<sup>36</sup> In the process of discovering new anthelmintics, Sloan and co-workers identified eupenifeldin **23** and its monotropolone derivative noreupenifeldin **53** from an unidentified *ascomycete* by bioassay-guided fractionation, in which eupenifeldin **23** shows modest *in vitro* anthelmintic activity.<sup>37</sup> In the same year, a new bisphenol meroterpenoid ramiferin **54**, together with two known compound pughinin A **52** and eupenifeldin **23**, were isolated from the fungus *Kionochaeta ramifera* BCC 7585. Crude extracts of the fungus exhibit strong antimalarial activity with an *IC*<sub>50</sub> value of 0.6 µg/ml. Ramiferin **54** and eupenifeldin **23** proved to be the bioactive components.<sup>38</sup> Next, in a search for natural products effective against human glioma and tumor cells, Che and co-workers investigated a strain of *Phoma* sp. isolated from a soil sample collected from the Qinghai-Tibetan plateau, Tibet, P. R. China. Bioassay guided separation of the crude extracts afforded the active compound eupenifeldin **23**, which shows potent antiproliferative effects against three human glioma cell lines, with *IC*<sub>50</sub> values of 0.08-0.13 µM. Additionally, two new related meroterpenoids phomanolides A **55** and B **56** were isolated from the solid substrate fermentation cultures of the fungus together.<sup>39</sup>

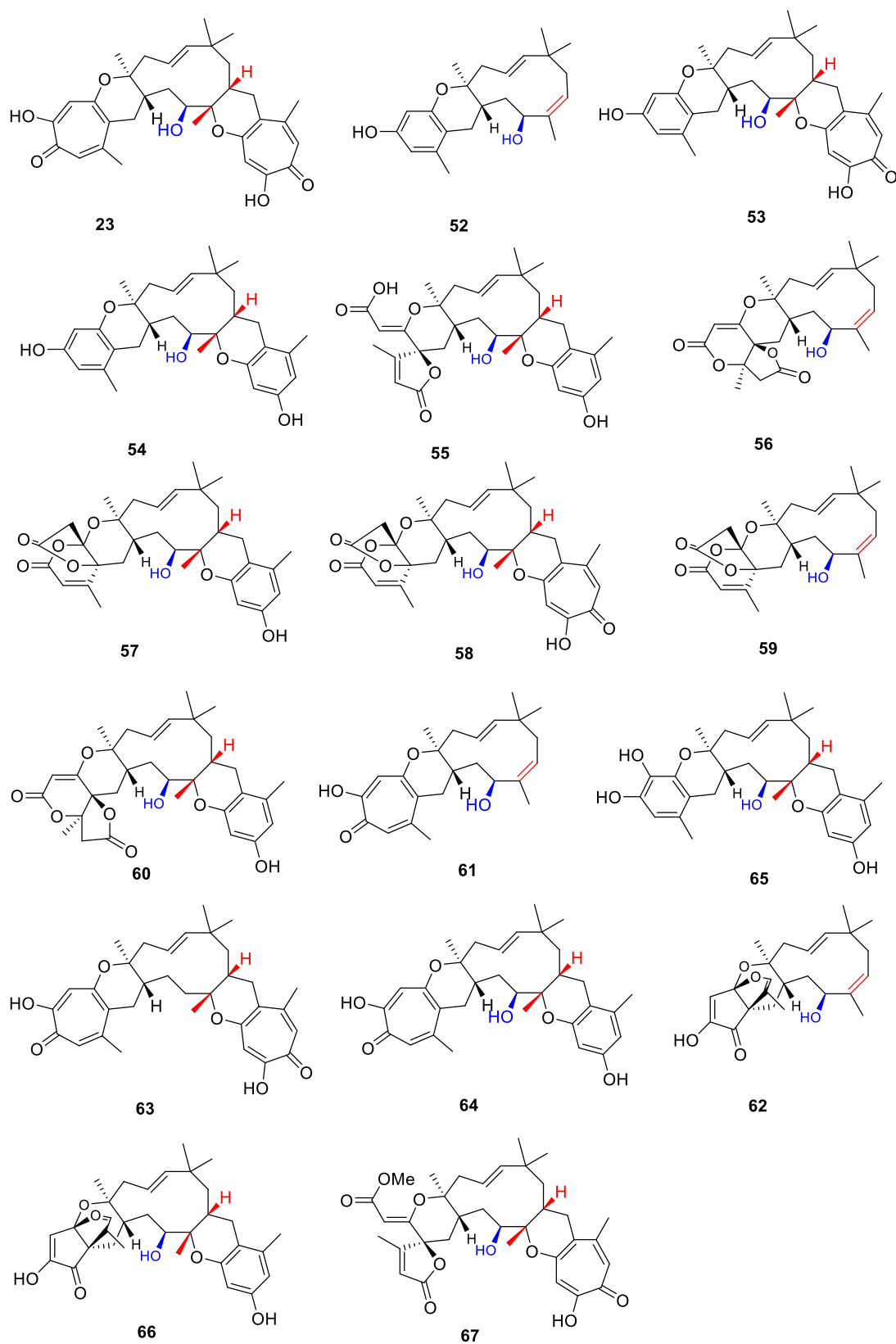


Figure 1.2.3.2 Eupenifeldin-type TS isolated from different fungi.

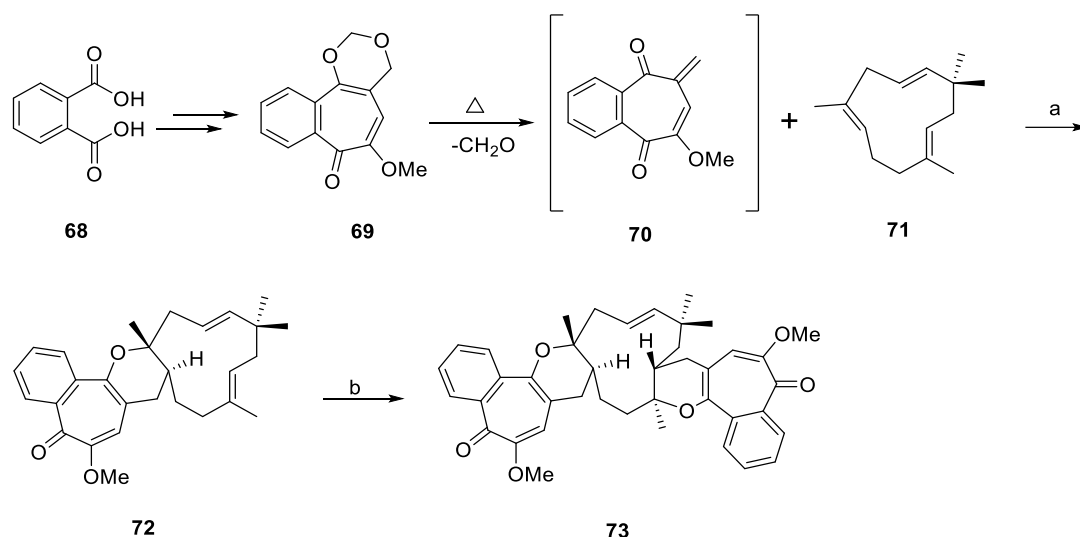
An ongoing search for new tropolone sesquiterpenoids from *Phoma* sp. led to the production of four new meroterpenoids phomanolides C–F **57-60** and neosetophomone B **61**.<sup>40</sup> In 2019, six meroterpenoids including neosetophomone A **62**, neosetophomone B **61**, dehydroxyeupenifeldin **63**, noreupenifeldin B **64**, 22-hydroxyramiferin **65** and eupenifeldin **23**, were isolated and elucidated from *Neosetophoma* sp. Cytotoxicity assays reveal eupenifeldin **23**, possessing two tropolone moieties, to be the most potent against breast, ovarian, mesothelioma, and lung cancer cells.<sup>41</sup> In 2020, fungus *Phomopsis tersa* FS441 originating from a deep-sea sediment sample was tentatively investigated for discovering structurally unique and biologically effective fungal metabolites. As a result, two unusual meroterpenoids phomeroids A **66** and B **67**, together with eupenifeldin **23**, were obtained and identified. And phomeroid B **67** and eupenifeldin **23** show potent antiproliferative effects against four human tumor cell lines (Figure 1.2.3.2).<sup>42</sup>

#### 1.2.4 Biomimetic Synthesis of Tropolone Sesquiterpenoids

Due to fascinating structures and important biological activities, tropolone sesquiterpenoids (TS) and related compounds from different fungi have attracted broad interest from synthetic chemists since 1999. Biosynthetically, these fungal metabolites can be viewed as addition products of tropolones or benzenes to a sesquiterpene such as humulene, which may be hydroxylated prior or after addition.

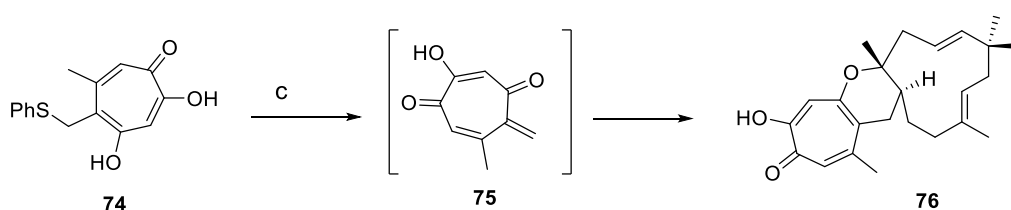
To determine whether a hetero Diels-Alder (hDA) reaction is a chemically feasible biomimetic strategy for the synthesis of TS, benzotropolone **69** was prepared starting from commercially available phthalic acid **68** and used as biomimetic precursor. Thermolysis of compound **69** generates the required tropolone ortho-quinone methide (*o*-QM) **70** *in situ*, this undergoes an hDA reaction with 1.5 equivalents of 2*E*-humulene **71** in *p*-xylene at 150 °C to afford the least hindered monosubstituted TS product **72** with correct regiochemistry in good yield. Addition of the second tropolone was proved to be difficult but possible. Heating compound **72** and 4 equivalents of **69** in xylene gives di-substituted TS **73** (1:1 mixture of diastereomers) in 20 % yield (Scheme 1.2.4.1).<sup>43</sup>





**Scheme 1.2.4.1** Biomimetic synthesis of tropolone natural products *via* an hDA reaction. Reagents and conditions: (a) sealed tube, *2E*-humulene **71** (1.5 equivalents), *p*-xylene, 150 °C, 24 h, 60 %; (b) sealed tube, benzotropolone **69** (4 equivalents), *p*-xylene, 150 °C, 24 h, 20 %.

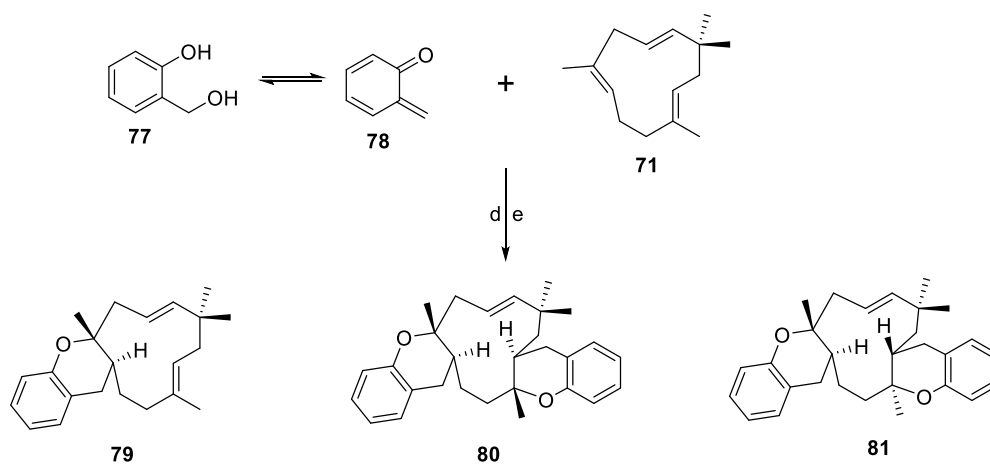
Three years later, the research toward a possible biomimetic hDA reaction between *2E*-humulene **71** and a novel tropolone *o*-QM **75** was reported. This tropolone *o*-QM was a completely novel reactive intermediate at that time, whose precursor is tropolone **74**. Heating compound **74** and *2E*-humulene **71** in *p*-xylene at 150 °C for 24 h yields dehydroxyepolone B **76** in an encouraging yield (Scheme 1.2.4.2), thus offering experimental support to a possible hDA biosynthesis of epolone B **44** and pycnidione **22**.<sup>44</sup>



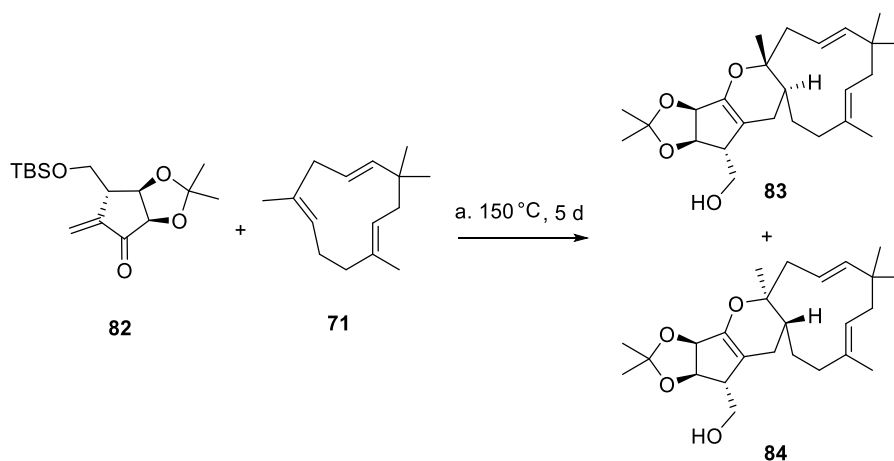
**Scheme 1.2.4.2** Biomimetic hDA reaction between *2E*-humulene **71** and tropolone *o*-QM. Reagents and conditions: (c) sealed tube, *2E*-humulene **71**, *p*-xylene, 150 °C, 24 h, 22 %.

There is also one study related to the biomimetic synthesis of phenol-meroterpenoids. Lucidene **80** is a benzopyran derived natural product isolated from plants, which is the analogue of pughiinin A **52** and ramiferin **54**. The co-isolation of *2E*-humulene **71** from the same species leads to the speculation

that lucidene is the product of two consecutive hDA cycloadditions of compound **71** and two equivalents of benzene *o*-QM **78** that are synthesized from compound **77**. When 2*E*-humulene **71** is heated in the presence of 2.05 equivalents of compound **77**, monoadduct **79** is obtained in 28 % yield, together with a mixture of *bis*-adducts lucidene **80** and its diastereomer **81**. The same reaction using 6 equivalents of precursor **77** gives only an inseparable mixture of lucidene **80** and its diastereomer **81** with an encouraging 45 % yield. In this case, no monoadduct **79** is observed (Scheme 1.2.4.3).<sup>45</sup>

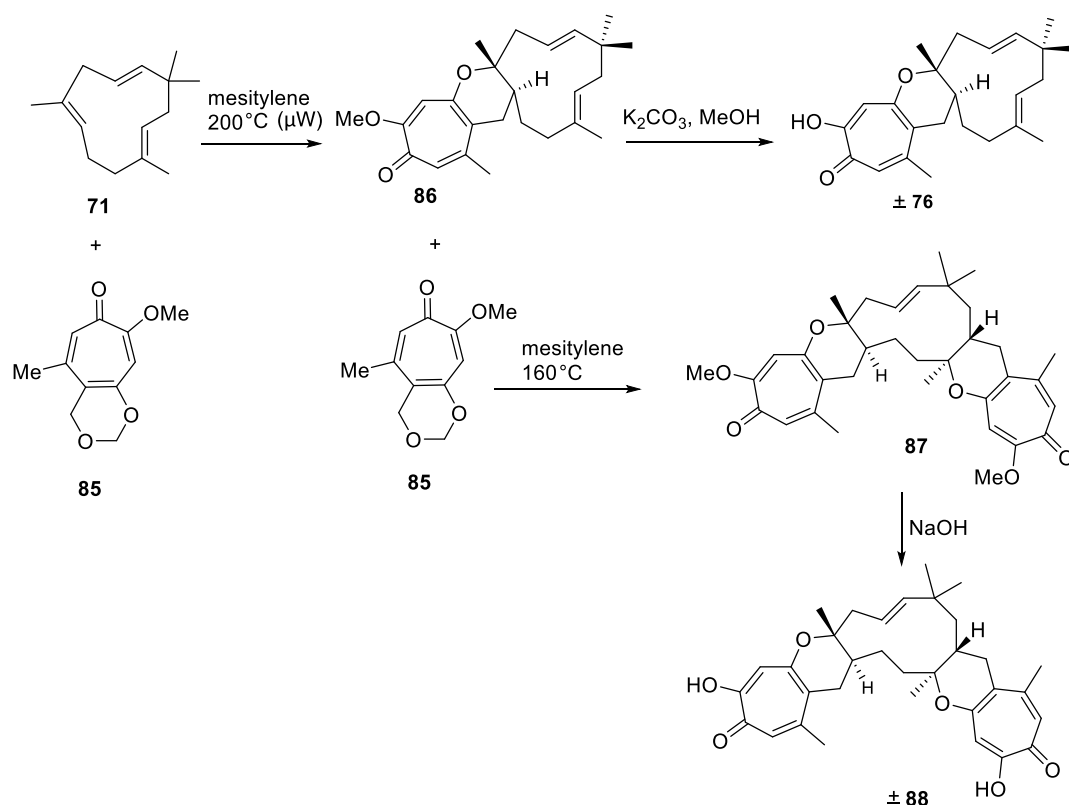


**Scheme 1.2.4.3** Biomimetic Synthesis of lucidene **80**. Reagents and conditions: (d) sealed tube, compound **77** (2.05 equivalents), 170°C; (e) sealed tube, compound **77** (6 equivalents), 170°C.



**Scheme 1.2.4.4** Biomimetic Synthesis of the core skeleton of xenovulene A **24**.

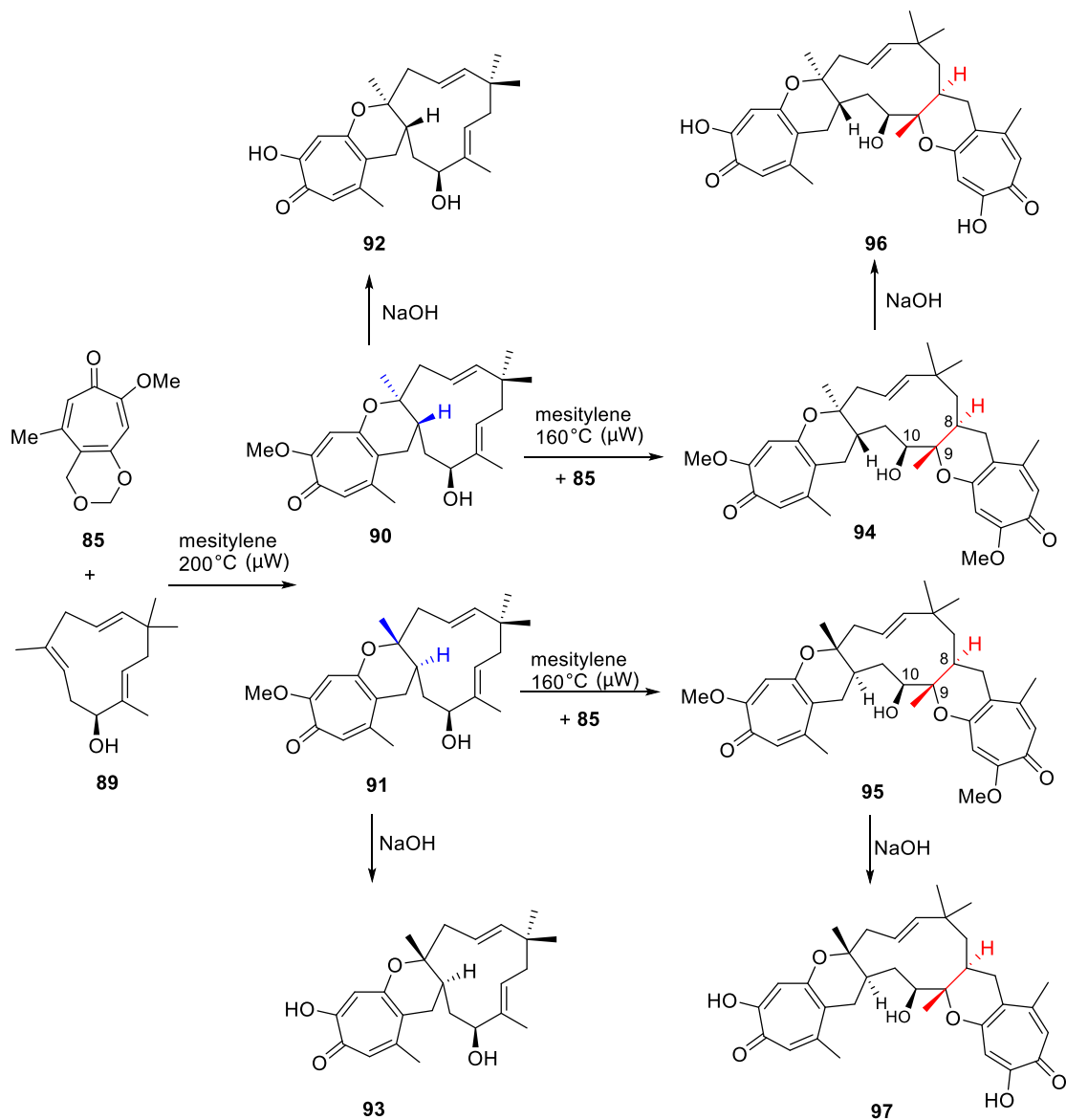
A biomimetic, regio- and stereoselective approach to the core skeleton of xenovulene A **24** was also described by Kirschning and co-workers.<sup>46</sup> This core structure is proposed to be constructed through an hDA cycloaddition between an  $\alpha,\beta$ -unsaturated ketone compound and *2E*-humulene **71**. In their study, heating vinyl ketone **82** with an excess of *2E*-humulene **71** for 5 days, followed by TBAF mediated TBS deprotection, afford two hDA cycloadducts compound **83** and **84** (2:1 mixture of diastereomers) in 51 % combined yield (Scheme 1.2.4.4).



**Scheme 1.2.4.5** Biomimetic Synthesis of dehydroxyepolone B **76** and  $(\pm)$ -dehydroxypycnidione **88**.

In 2021, a concise hDA reaction based-approach to synthesize a series of pycnidione type TS, including (-)-epolone B **93**, (+)-isoeolone B **92**,  $(\pm)$ -dehydroxypycnidione **88**, 8,9-*epi,epi*-(+)-pycnidione **97** and (-)-10-*epi*-pycnidione **96** was described by Sarlah and co-workers.<sup>47</sup> This study reveals that heating tropolone *o*-QM precursor **85** and *2E*-humulene **71** under microwave irradiation generates monoadduct TS **86** in 82 % yield as a single constitutional isomer. This is further deprotected under basic conditions to give racemic dehydroxyepolone B **76**. For the second hDA cycloaddition, compound **85** was added slowly as a solution over the course of several hours to

compound **86** under heating condition, which results in the formation of *bis*-adduct TS **87** (1:1 mixture of diastereomers) in 13 % yield. Further base-induced deprotection of compound **87** delivers ( $\pm$ )-dehydroxypycnidione **88** (Scheme 1.2.4.5).



**Scheme 1.2.4.6** Biomimetic synthesis of (-)-epolone B **93**, (+)-isoepolone B **92**, 8,9-*epi, epi*-(+)-pycnidione **97** and (-)-10-*epi*-pycnidione **96**.

On the other hand, the application of 2*E*-humulenol **89** with methylated tropolone *o*-QM precursor **85** also leads to the desired hDA reaction, furnishing two monoadduct TS **90** and **91** (1.4:1 mixture of diastereomers) in 89 % combined yield. After separation of the diastereomers, base-induced deprotection of **90** and **91** afford (+)-isoepolone B **92** and (-)-epolone B **93**. As for the second hDA

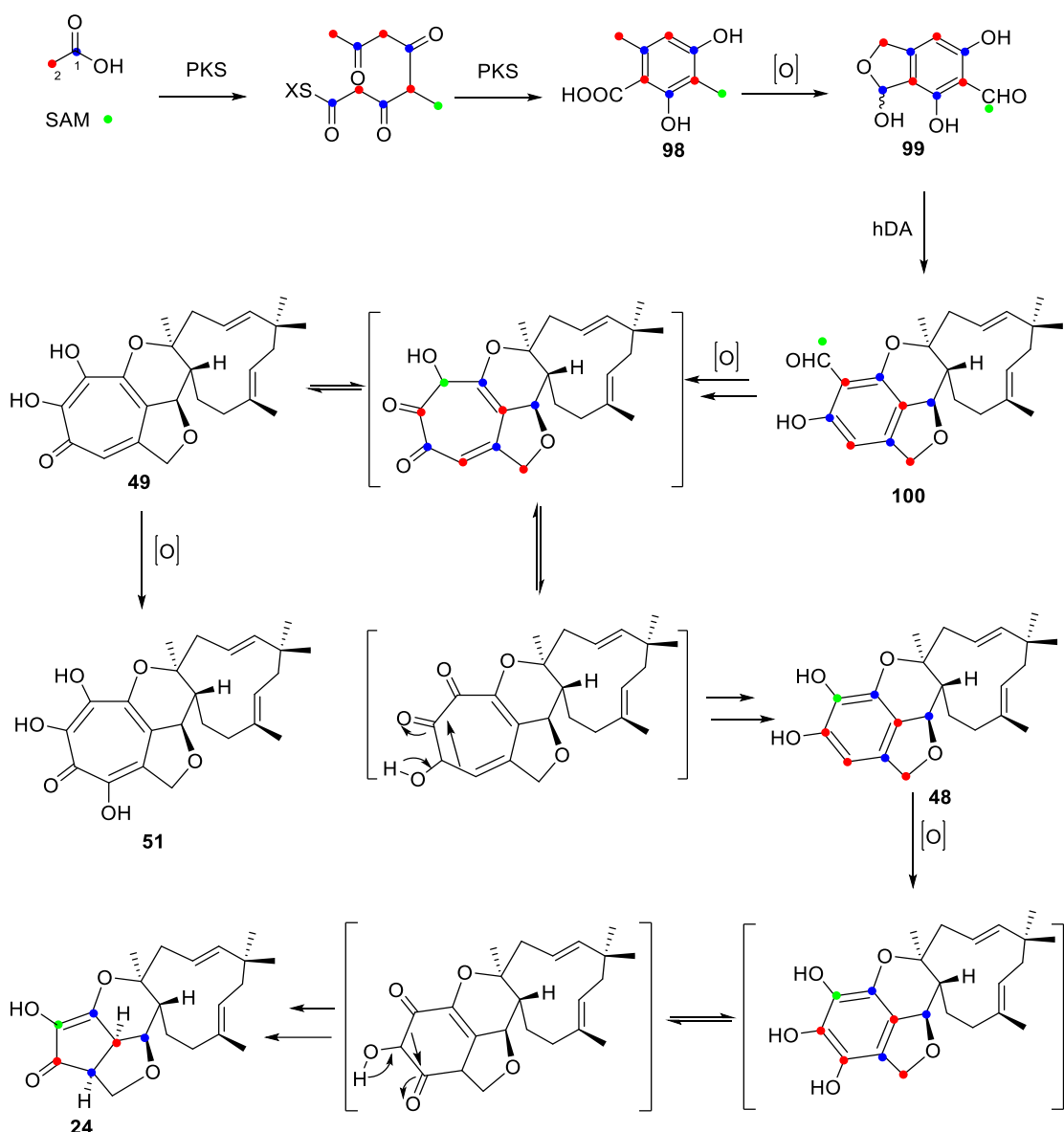
cycloaddition, compound **90** and **91** are separately reacted with compound **85**, they each produce only a single diastereomer corresponding to compound **94** and **95** in 24 % and 26 % yield. This observation highlights the importance of the hydroxy group in compound **89** on the pronounced selectivity of the second hDA reaction. Further deprotection of **94** and **95** in base condition led to two bis-adduct TS (-)-10-*epi*-pyncnidione **96** and 8,9-*epi,epi*-(+)-pyncnidione **97**, the diastereomers of pyncnidione **22** (Scheme 1.2.4.6).

In conclusion, synthesis of lucidene **80**, analogues of epolone B **44** and pyncnidione **22** *via* hetero-DA additions of benzene *o*-QM /tropolone *o*-QM and 2*E*-humulene **71**/2*E*-humulenol **89** provide further support for their biosynthetic pathway *via* hDA reactions. However, it is clear that in the native biosynthetic pathways of TS the hDA steps must be catalyzed effectively because non-catalysed reactions during chemical synthesis require very high temperatures.

### 1.2.5 Biosynthesis of Tropolone Sesquiterpenoids

Tropolone sesquiterpenoids (TS) are fungal meroterpenoid natural products that display a significant array of biological activities, including cytotoxic, antifungal, antimalarial, and anthelmintic effects.<sup>17,18,33,34,37,38,41,48-51</sup> Their unique structures and profound biological activities have attracted much attention from biochemists too.

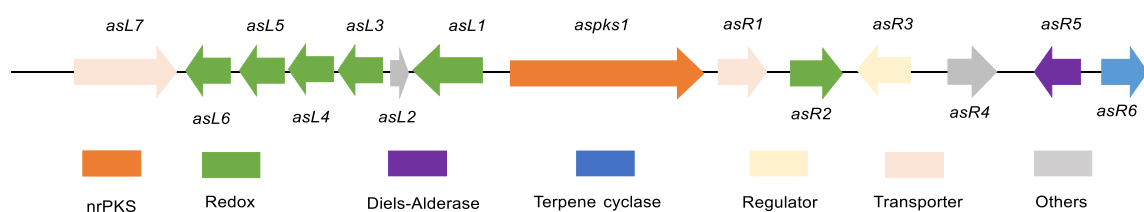
In 1997, Simpson and co-workers proposed a plausible pathway to xenovulene A **24** and its co-metabolites based on the incorporation studies with <sup>13</sup>C-labelled acetates and methionine in *Acromonium strictum*. In this proposed pathway the non-terpenoid portion is derived from 3-methylorsellinic acid **98** which would be converted to the lactol **99** by standard biosynthetic modifications. Compound **99** could be subsequently converted to a benzene *o*-QM and be attached to the 2*E*-humulene **71** *via* an hDA reaction to form the key tetracyclic intermediate **100**. The resultant meroterpenoid **100** could then undergo hydroxylation and ring expansion *via* an  $\alpha$ -ketol rearrangement to yield **49** and **51**. This reaction could be followed by oxidative decarboxylation of **49** to form the trioxygenated benzenoid meroterpenoid **48** in which the aromatic ring contains a methionine-derived carbon. Finally, compound **48** would be converted into the end product xenovulene A **24** through further oxidation,  $\alpha$ -ketol mediated ring contraction, and decarboxylation (Scheme 1.2.5.1).<sup>52</sup>



**Scheme 1.2.5.1** Simpson's proposed pathway to the xenovulene A **24** based on the results of  $^{13}\text{C}$ -labelled acetates and methionine feeding experiments in *Acremonium strictum*.

Subsequent molecular genetic studies on the isolation and heterologous expression of a polyketide synthase (PKS) gene isolated from xenovulene-producing cultures of *A. strictum* indicates that 3-methylorcinaldehyde **19** is the initial polyketide-derived intermediate of xenovulene A **24**. This is the direct product of the fungal nonreducing polyketide synthase (nr-PKS) encoded by gene *aspks1*, but further elucidation of the xenovulene A BGC was initially not possible because of the difficulty in obtaining targeted knockouts.<sup>19</sup>

11 years later, Cox and co-workers finally identified the biosynthetic gene cluster and the chemical steps in the biosynthetic pathway to xenovulene A **24** (Figure 1.2.5.1) after the genome of *A. strictum* was fully sequenced.<sup>20</sup> BLAST searching using the previously identified *aspsk1* gene rapidly identified a 1.6-Mb scaffold that contained a 49-kb BGC potentially involved in xenovulene A biosynthesis, in which homologues of all four key tropolone biosynthetic proteins were detected. In order to link the potential cluster to the biosynthesis of xenovulene A **24**, targeted knockout (KO) experiments were carried out. Only one true *aspsk1* knockout was found to lose the production of xenovulene A **24** and related co-metabolites. An additional transcriptomic comparison of expression levels of genes surrounding *aspsk1* under producing and non-producing conditions clearly revealed that genes from *asL7* to *asR7*, inclusive, are significantly upregulated under producing conditions, while gene expression levels outside this area are similar under the two different conditions.



**Figure 1.2.5.1** The BGC of xenovulene A **24**.

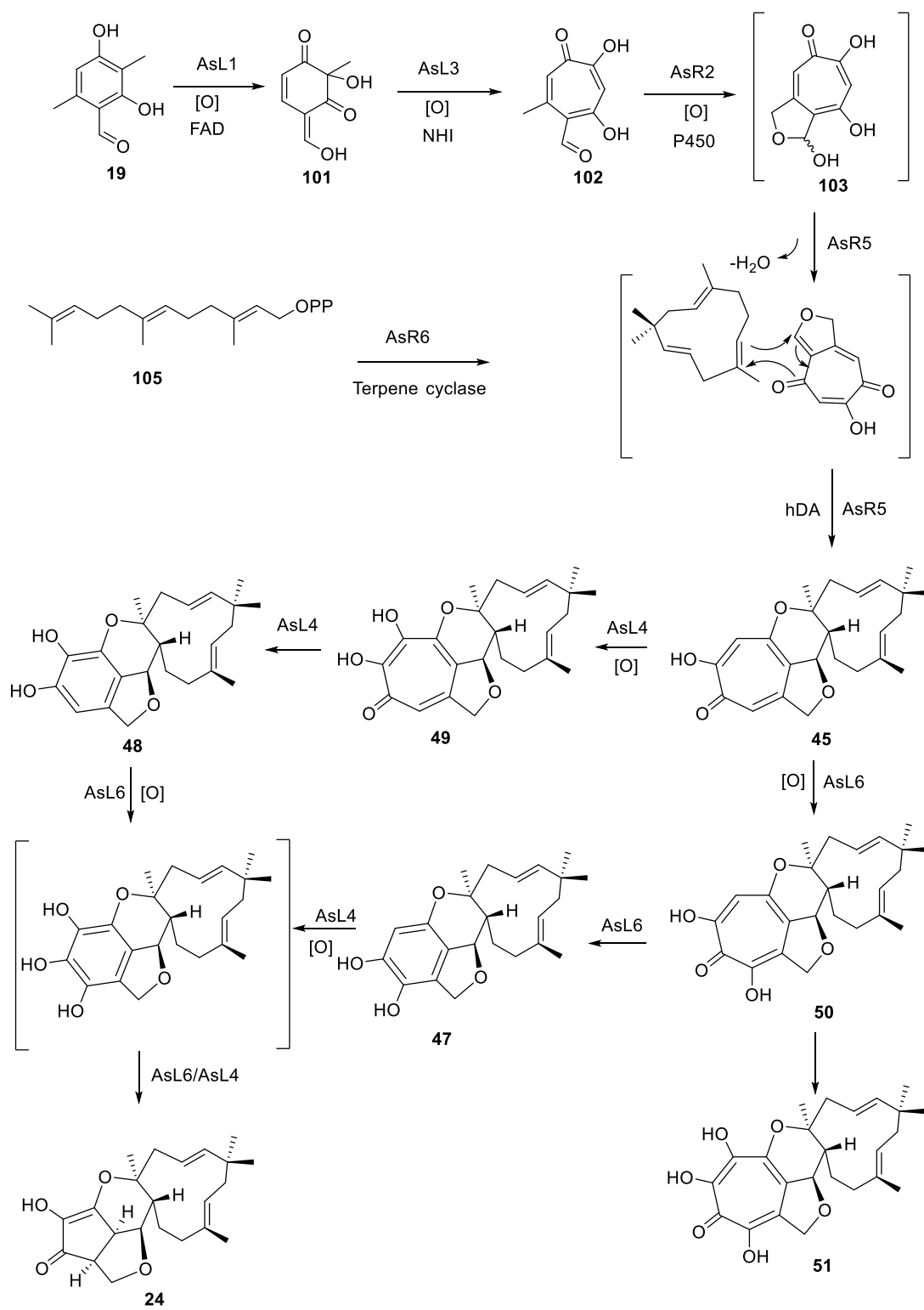
The function of the genes from the *aspsk* BGC and the chemical steps in the biosynthetic pathway were confirmed by heterologous expression experiments (Scheme 1.2.5.2). The biosynthesis of Xenovulene A **24** shares the same early-stage reactions with those of stipitatic acid **115**.<sup>53</sup> At the outset, the nr-PKS produces 3-methylorcinaldehyde **19** that subsequently undergoes oxidative dearomatisation catalyzed by the FAD dependent salicylate monooxygenase AsL1 to give enone **101**. The non-heme Fe II dependent dioxygenase AsL3 then uses compound **101** as the substrate to catalyze the methyl-hydroxylation and ring expansion to provide the first tropolone, stipitaldehyde **102**, which is then probably transformed into unstable lactol **103** by the cytochrome P450 AsR2. In the next biosynthetic process, with the aid of a putative hDAase AsR2, compound **103** reacts with the *2E*-humulene **71** synthesized by terpene cyclase AsR6 to form xenovulene B **45** as the first

important TS intermediate. Compound **47** could be derived from compound **45** by a ring-contraction reaction catalyzed by a putative NAD/FAD-dependent oxidoreductases AsL4. Interestingly, compound **45** can also be transformed into compound **48** by another putative NAD/FAD-dependent oxidoreductases AsL6. AsL6 and AsL4 appear to be able to partially complement each other's activities as xenovulene A **24** biosynthesis is severely reduced, but not abolished when one of them was knocked out. Both **47** and **48** could be further transformed to xenovulene A **24** to complete the biosynthesis in the presence of AsL4 or AsL6.

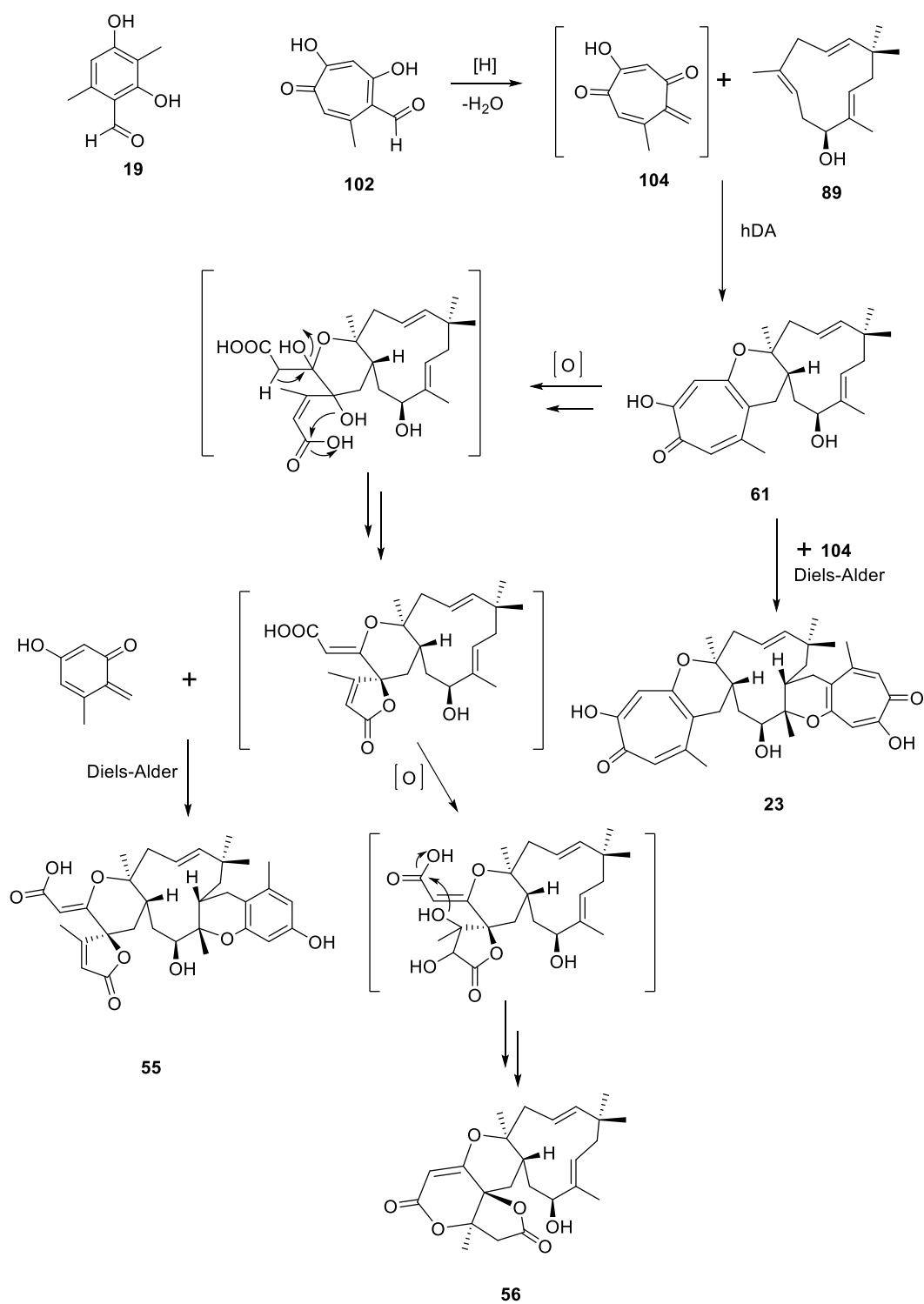
In short, the research indicates that at least 8 genes are necessary for the biosynthesis of xenovulene A **24**, including *aspsk1*, *asL1* encoding an FAD-dependent hydroxylase, *asL3* encoding a non-haem iron dioxygenase, *asR2* encoding a cytochrome P450 monooxygenase, *asR5* encoding an hDAase, *asR6* encoding a terpene cyclase and *asL6* and *asL4* encoding putative NAD/FAD-dependent oxidoreductases. The suggested pathway is consistent with not only the biomimetic synthesis of TS but also Simpson's earlier isotopic labelling studies. hDA addition between the tropolone product and 2*E*-humulene **71** generates the important meroterpenoid intermediate, which is further converted to the final product xenovulene A **24** by two consecutive ring-contraction reactions.

The time-course study for the production of epolone B **44** and pycnidione **22** during a fermentation process suggested that epolone B **44** could be the biosynthetic precursor of pycnidione **22**.<sup>33</sup> Che and co-workers proposed a hypothetical biosynthetic pathway, based on the known xenovulene A pathway, for eupenifeldin **23** and related meroterpenoids, which could be derived *via* hDA additions of the hypothetical precursor 2*E*-humulenol **89** and the tropolone *o*-QM **104** (Scheme 1.2.5.3). Compound **104** is the reactive form of tropolone stipitaldehyde **102** that is derived from 3-methylorcinaldehyde **19** *via* a series of oxidative reactions. Enzymatic-catalyzed hDA addition between **89** and **104** would generate neosetophomone B **61**, a diastereomer of epolone B **44**. The final products compounds **55**, **56**, and **1** could be generated *via* different reaction routes including the second hDA additions and oxidations, In addition, the hDA reaction between **89** and **104** gives the unfavoured *exo*-product state in the proposed biosynthetic pathways, indicating that the reaction could be catalyzed by a DAase.<sup>39</sup>





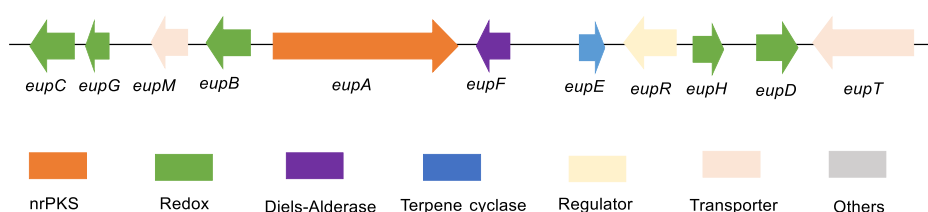
**Scheme 1.2.5.2** The proposed biosynthetic pathway for xenovulene A **24** based on *aspks* BGC.



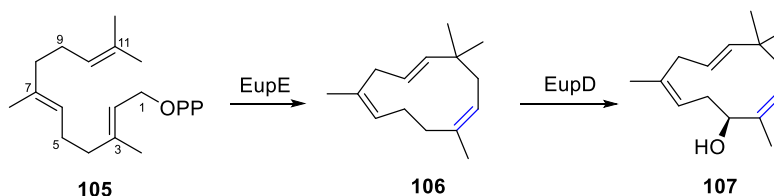
**Scheme 1.2.5.3** Proposed biosynthetic pathways for eupenifeldin **23** and related meroterpenoids by Che and co-workers.

To explore the biosynthesis of the tropolone sesquiterpenes, the genome of a eupenifeldin-producing fungi *Phoma sp.* was sequenced and analyzed.<sup>54</sup> The biosynthetic gene cluster of eupenifeldin (*eup*

BGC) was identified and the 5 functional enzymes from the BGC were validated by genomic analysis, gene disruption, and product analysis (Figure 1.2.5.2). A nr-PKS EupA, an FAD-dependent monooxygenase EupB, and a non-heme Fe II-dependent dioxygenase EupC, were identified as the enzymes responsible for tropolone precursor formation, while a terpene cyclase EupE, highly homologous to AsR6, was also identified to be likely to catalyze 2*Z*-humulene **106** formation *via* the cyclization of FPP **105** and a cytochrome P450 enzyme EupD is responsible for the production of 2*Z*-humulenol **107** by hydroxylation (Scheme 1.2.5.4). This study paved the way to further decipher the biosynthetic pathway of eupenifeldin TS.



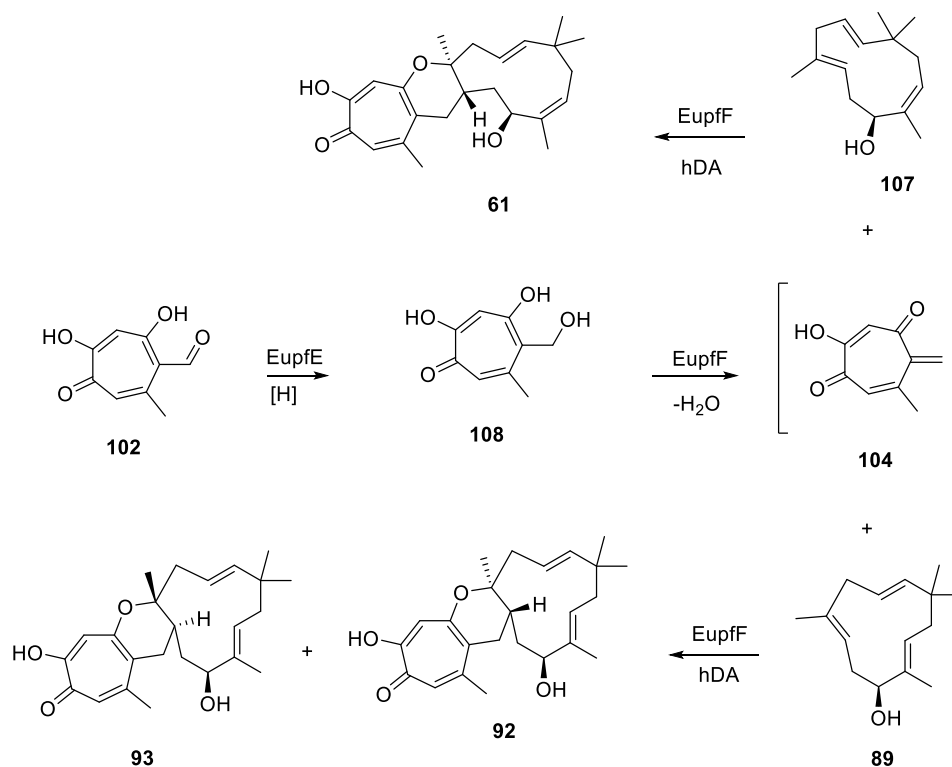
**Figure 1.2.5.2** The *eup* BGC of eupenifeldin **23** from *Phoma.sp.*



**Scheme 1.2.5.4** The chemical steps catalyzed by EupE and EupD during the biosynthesis of eupenifeldin **23**.

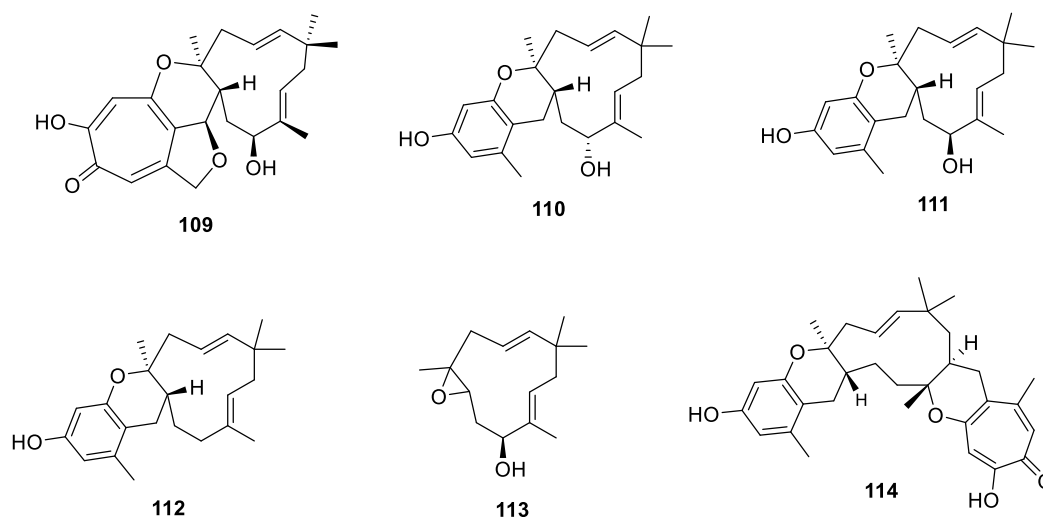
In the same year, an intermolecular hDA reaction involved during the biosynthesis of neosetophomone B **61** was reported.<sup>55</sup> The putative hDAase EupfF from the *eupf* BGC was identified to catalyze the hDA addition reaction of stipitol **108** and 2*Z*-humulenol **107** into enantiomerically pure neosetophomone B **61**. Interestingly, when the natural 2*Z*-humulenol **107** is switched to unnatural substrate 2*E*-humulenol **89**, two diastereomers compound **92** and **93** (1:1 ratio) are detected in the *in vitro* EupfF assay. In addition, dehydration of **108** catalyzed by EupfF leads to the reactive tropolone o-QM **104** as a diene precursor for the hDA reaction. The short-chain dehydrogenase EupfE is identified to reduce tropolone precursor stipitaldehyde **102** to stipitol **102**

(Scheme 1.2.5.5). Thus, EupfF is the first fully characterized as intermolecular hDAase that directly controls the stereo-selectivity of hDA reaction in the case of natural substrate 2Z-humulenol **107**.



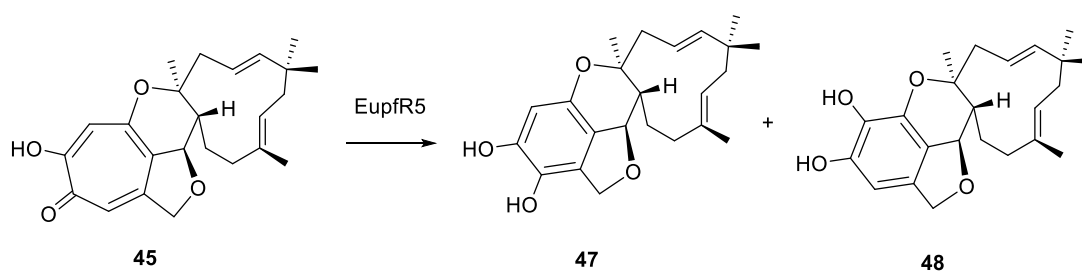
**Scheme 1.2.5.5** The chemical steps catalyzed by EupfE and EupfF in the biosynthetic pathway of eupenifeldin **23**.

In 2020, the combinatorial biosynthesis of a series of unprecedented analogs of pycnidione **22** and xenovulene A **24** were reported by Cox and co-workers.<sup>35</sup> In their study, the *eup2* BGC involved in eupenifeldin **23** biosynthesis was identified in *CF150626* while the *pyc* BGC from the *CF236968* is confirmed to be responsible for formation of pycnidione **22**. Rational extension and diversification of the xenovulene A **24** biosynthetic pathway by mixing and matching genes from the *eup2* BGC and the *pyc* BGC in the fungal host *A. oryzae* NSARI lead to production of six new unnatural tropolone sesquiterpenoids compounds **109-114** in good yields (Figure 1.2.5.3).

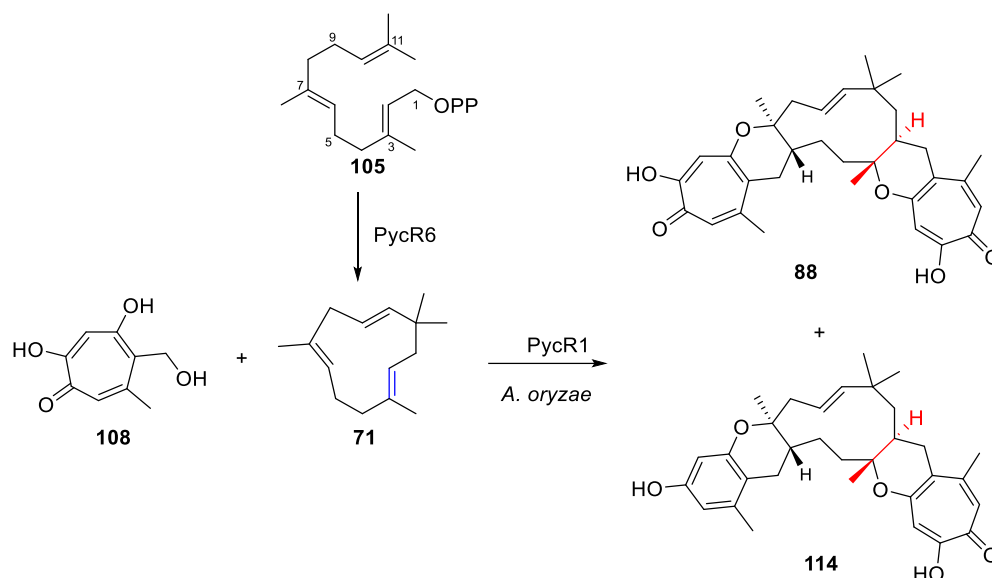


**Figure 1.2.5.3** New unnatural tropolone sesquiterpenoids from *A. oryzae* NSAR1 transformants.

Furthermore, formation of the benzopyranyl moieties in TS was identified to be derived *via* a ring-contraction of tropolone precursor moieties in TS based on the isotope-feeding studies with  $^{13}\text{C}$ -acetate. In addition, the FAD-binding monooxygenase Eup2R5 was identified to catalyze the regioselective oxidative ring contractions of xenovulene B **45** to yield a mixture of compound **47** and **48**, which are the previously reported products of AsL4 and AsL6 (Scheme 1.2.5.6). The *in vitro* assay of the putative humulene synthase PycR6 confirms that it converts farnesylpyrophosphate (FPP) **105** to 2*E*-humulene **71**, while co-expression of *pycR1* that encodes hDAase, *pycR6*, and the stipitaldehyde **102** producing genes from the BGC of xenovulene A **24** in *A. oryzae* lead to the production of two di-TS, dehydroxypycnidione **88** and compound **114** (Scheme 1.2.5.7).



**Scheme 1.2.5.6** The ring-contraction of xenovulene B **45** catalyzed by Eup2R5.

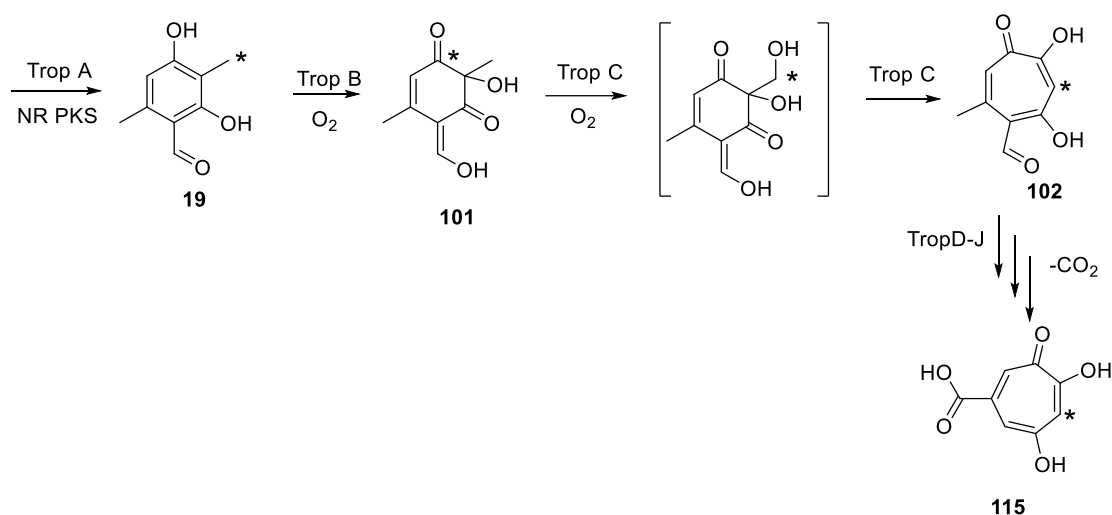


**Scheme 1.2.5.7** The chemical steps catalyzed by PycR6 and PycR1 in the biosynthetic pathway of pycnidione **22**.

### 1.2.6 Biosynthesis of Tropolone and Humulene

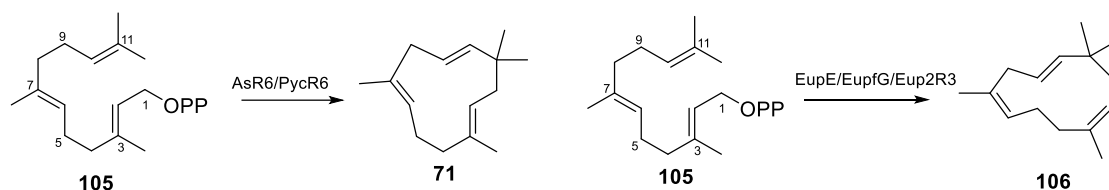
Most TS contain a tropolone core, a seven-membered ring aromatic system. The first discovered tropolone natural product was stipitatic acid **115** from *Penicillium stipitatum* in 1942.<sup>56</sup> Its biosynthesis was fully elucidated by Cox and co-workers in 2012.<sup>57</sup> The research shows that at least three proteins (a nr-PKS with a C-methyltransferase (Trop A), a FAD-dependent monooxygenase (Trop B), and a non-heme Fe (II)-dependent oxidase (Trop C)) are necessary for the construction of the tropolone ring. These proteins work together to form a polyketide and oxidise it twice to achieve the rearrangement required to produce the tropolone nucleus.<sup>57</sup> Undoubtedly, the discovery of BGC for tropolone ring lays a good foundation for the identification of BGC for TS.

The chemical steps and functional enzymes involved in the biosynthesis of the tropolone class of fungal metabolites such as stipitatic acid **115** have already been elucidated by gene knockout and heterologous expression experiments. Biosynthesis of stipitatic acid **115** starts with the production of 3-methylorcinolaldehyde **19**. It is then oxidized by Trop B to obtain the hydroxylated dearomatized product **101**. The next step of the process is the ring expansion of compound **101** to give the first tropolone stipitaldehyde **102** catalyzed by TropC (a non-heme iron  $\alpha$ -ketoglutarate dependent dioxygenase), which involves the hydroxylation of the methyl group at C-3. Further transformations by Trop D-J are required to convert the intermediate **102** into final product **115** (Scheme 1.2.6.1).<sup>58</sup>



**Scheme 1.2.6.1** Biosynthesis of stipitatic acid **115**.

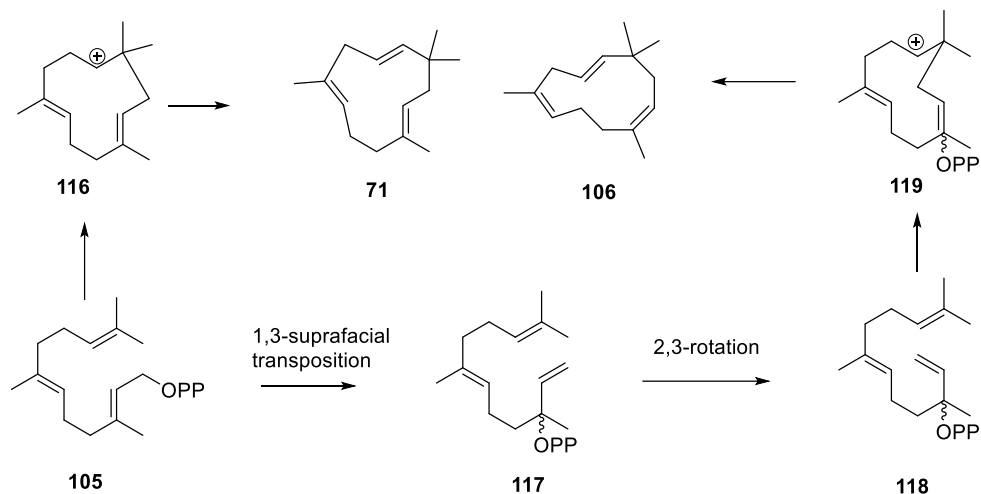
All TS have an 11-membered sesquiterpene ring that shows great resemblance to humulene. Humulene is a common intermediate in the biosynthesis of many fungal cyclic sesquiterpenoids, which has been isolated from fungi *Fusarium fujikorii*,<sup>59</sup> *Colletotrichum acutatum*<sup>60</sup> and *Stereum hirsutum*.<sup>61</sup>



**Scheme 1.2.6.2** Biosynthesis of 2E-humulene **71** and 2Z-humulene **106**.

Two cryptic terpene cyclases (AsR6, PycR6) have recently been identified as responsible for 2E-humulene **71** formation in pycnidione type TS, while the homologous enzymes (EupE; EupfG and Eup2R3) are responsible for the production of 2Z-humulene **106** in eupenifeldin type TS (Scheme 1.2.6.2).<sup>62</sup> They bear no significant sequence identity to any known terpene synthase such as Ffsc4 and CaTPS.<sup>59,60</sup> From a mechanistic viewpoint, 2E-humulene **71** formation involves a 1,11-cyclization of FPP **105**, which yields the 2E-humulyl cation **116** and deprotonation then affords **71** without further rearrangement. While the formation of 2Z-humulene **106** requires a formal isomerization of the 2E-alkene in FPP **105**, that is proposed to proceed *via* a 1,3-suprafacial transposition of the PPI in **105** and the rotation of the C-2/C-3

bond in compound **117** to give *cisoid* nerolidyl diphosphate **118** (NPP). Further 1-11-cyclization would yield the *Z,E*-humulyl cation **119** and deprotonation could then afford **106** (Scheme 1.2.6.3).<sup>63–65</sup>



**Scheme 1.2.6.3** Proposed mechanism of 1,11-cyclization reaction catalyzed by humulene synthase.

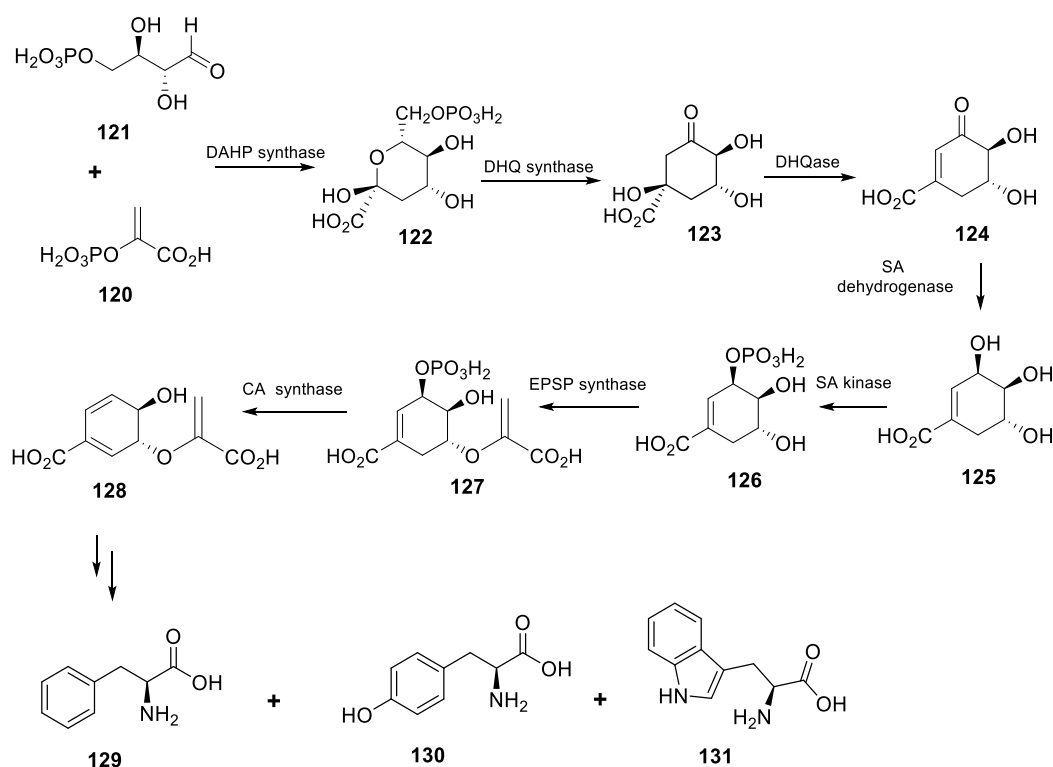
In 2021, the first crystal structure of non-canonical terpene cyclase AsR6 was reported by Cox and co-workers.<sup>62</sup> In this study, the *E. coli* codon-optimized genes *asR6* and *eup2R3* were expressed in *E. coli* BL21 and isolated as polyhistidine-tagged enzymes. Incubation of AsR6 with FPP **105** gives 2*E*-humulene **71** while identical *in vitro* assays with Eup2R3 affords 2*Z*-humulene **106**. Further structure-based site-directed mutagenesis of AsR6 and Eup2R3 identified a single residue, L285/M261, that controls the production of either 2*E*- or 2*Z*-humulene.

### 1.3 Natural products derived from the shikimate pathway

The shikimic acid pathway, ubiquitous in microorganisms and plants, provides precursors for the biosynthesis of primary metabolites such as the aromatic amino acids L-phenylalanine **129** (L-Phe), L-tyrosine **130** (L-Tyr), and L-tryptophan **131** (L-Trp), which are required for protein biosynthesis.<sup>66</sup> The shikimate pathway consists of seven reaction steps (Scheme 1.3.1), beginning with an aldol-type condensation of phosphoenolpyruvic acid **120** (PEP) and D-erythrose-4-phosphate **121** (E4P) to produce 3-deoxy-D-arabino-heptulosonic acid 7-phosphate **122** (DAHP), catalyzed by DAHP synthase. The second reaction of the shikimate pathway is an intramolecular aldol-type reaction



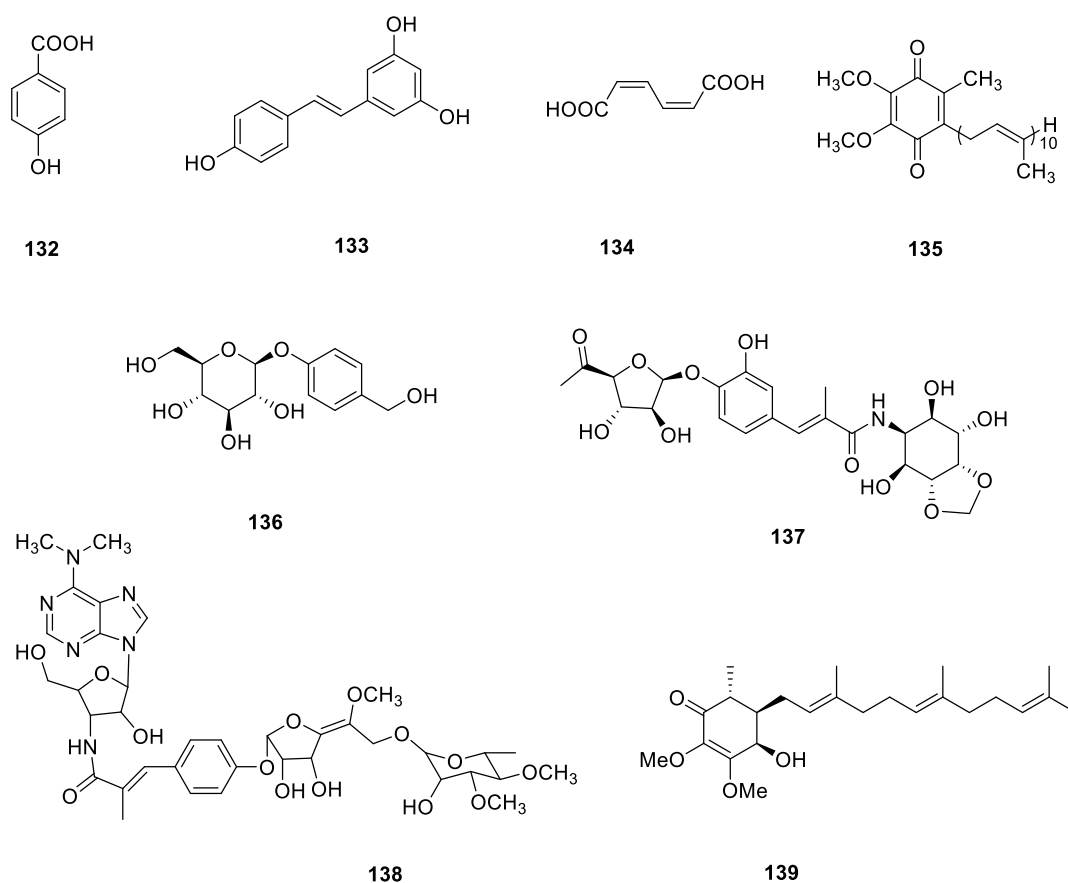
cyclization catalyzed by 3-dehydroquinate synthase (DHQ synthase), where the enol (C-6/C-7) of DAHP **122** nucleophilically attacks the carbonyl group (C-2), to produce the 3-dehydroquinic acid **123** (DHQ). In the third reaction step, the 3-dehydroquinase (DHQase) converts DHQ **123** into 3-dehydroshikimic acid **124** (DHS) by eliminating water and this reaction is reversible. The DHS **124** is converted to shikimic acid **125** (SA) by the catalytic action of shikimate dehydrogenase (SA dehydrogenase) with NADPH in the fourth reaction step. Shikimate kinase (SA kinase) then catalyzes the phosphorylation of the shikimic acid **125**, which is then converted by 5-enolpyruvylshikimate 3-phosphate synthase (EPSP synthase) to generate 5-enolpyruvylshikimic acid 3-phosphate **127** (EPSP) by the addition of PEP **120** to shikimic acid 3-phosphate **126** (SA 3-P). The last reaction step of the shikimate pathway is the production of chorismic acid **128** (CA) catalyzed by chorismate synthase that needs reduced flavin mononucleotide (FMNH<sub>2</sub>) as a cofactor.<sup>67</sup>



**Scheme 1.3.1** The shikimate pathway consists of seven chemical reaction steps.

At this metabolic node chorismate **128** either follows a biosynthetic pathway to the aromatic amino acids L-tyrosine **130** and L-phenylalanine **129** through the action of chorismate mutase, or it follows an alternative anabolic route *via* anthranilate synthase to form L-tryptophan **131** (Figure 1.3.1).<sup>68</sup> The

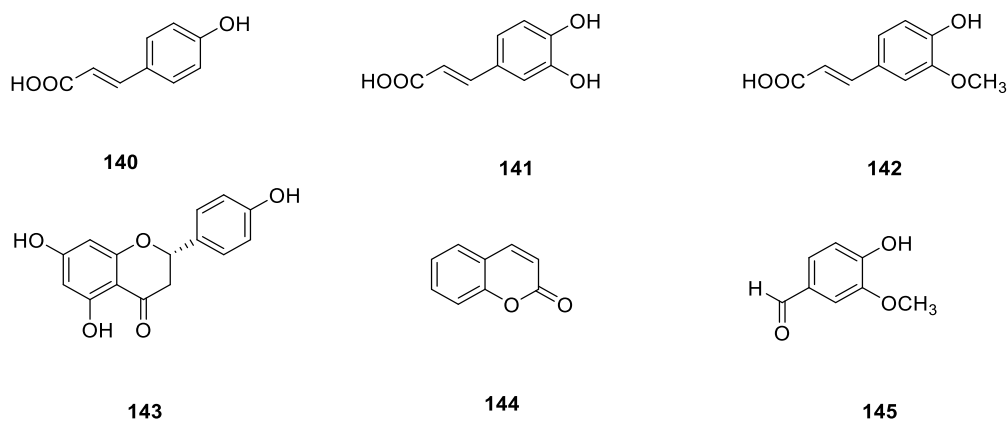
shikimate pathway plays an important role during the biosynthesis of primary metabolites such as the aromatic amino acids and folic acid. Several branch-points from the primary metabolic pathway also provide aromatic and, in some unusual cases, nonaromatic precursors for the biosynthesis of a considerable array of secondary metabolites.<sup>69</sup>



**Figure 1.3.1** Natural products derived from 4-Hydroxybenzoic acid (4-HBA) **132**. Stereochemistry of **138** is unknown.

4-Hydroxybenzoic acid **132** (4-HBA) is used as an important monomer for the production of polymers, fibers, and esters.<sup>70,71</sup> Biological production of 4-HBA is achieved by employing chorismate **128** as the precursor with glucose as the substrate in *E. coli*. Over-expression of the endogenous chorismate pyruvate lyase gene *ubiC* enhances the expression of the genes involved in the shikimate pathway, and yields 12 g/L of 4HBA in a fed-batch fermentation.<sup>72</sup> The 4-HBA-derived natural products are a large group of secondary metabolites which exhibit a wide variety of biological and pharmaceutical activities, including resveratrol **133**,<sup>73</sup> muconic acid **134**,<sup>74</sup> ubiquinone-10 **135**,<sup>75</sup> gastrodin **136**,<sup>76</sup> hygromycin A **137**,<sup>77</sup> A201A **138**,<sup>78</sup> and antroquinonol **139**<sup>79</sup> (Figure 1.3.1).

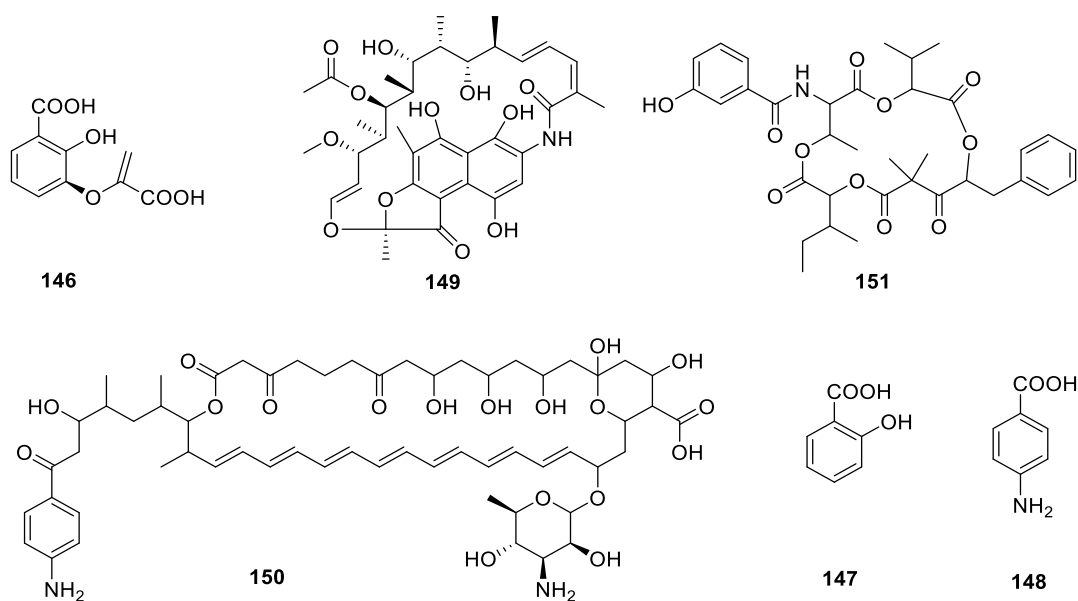
*p*-Coumaric acid **140** (*p*-CA) is a phenolic acid naturally produced by various seeds, fruits, and vegetables. It is a common precursor for the production of flavonoids, polyphenols, and polyketides.<sup>80</sup> Biological production of *p*-CA **140** is reported to be obtained in *S. cerevisiae* with phenylalanine **129** as precursor.<sup>81</sup> A more straightforward biosynthetic route is constructed in *E. coli*, which employs *L*-tyrosine **130** as the direct precursor.<sup>82</sup> *p*-CA **140** plays a central role in secondary metabolism because it can be subsequently used for the biosynthesis of naringenin **143**,<sup>83</sup> coumarin **144**<sup>84</sup> and phenolic acids (e.g. caffeic acid **141**,<sup>85</sup> ferulic acid **142**<sup>86</sup> and vanillin **145**<sup>87</sup>, Figure 1.3.2). Phenolic acids derivatives are commonly derived from the shikimate pathway. *Para* and ortho-hydroxylation patterns tend to occur among shikimate-derived aromatics, in contrast with the *meta*-hydroxylation patterns of phenolics arising from the polyketide pathway, which kind of facilitates in deducing the likely biosynthetic origin of an aromatic compound or unit.<sup>88</sup>



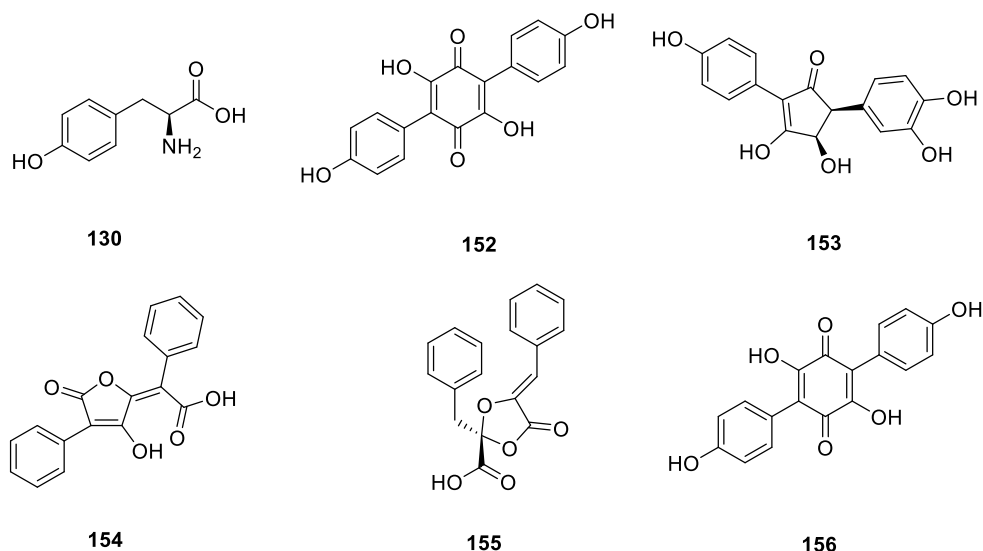
**Figure 1.3.2** Natural products derived from *p*-Coumaric acid **140**.

The shikimate pathway and variations of this pathway also play a critical role in providing aromatic precursors for the biosynthesis of a wide range of bioactive compounds (Figure 1.3.3).<sup>89</sup> For instance, shikimate-derived *p*-aminobenzoic acid **148** is used as a precursor for the biosynthesis of the polyene antibiotic candicidin **150**.<sup>90</sup> Chorismate **128** is a key precursor for biosynthesis of various value-added aromatic chemicals, such as anticancer agents unantimycin B **151** and immunosuppressants rafamycin **149**.<sup>91-92</sup> Salicylic acid **147** is an important phenolic acid that has been used to synthesize acetylsalicylic acid (aspirin), which is an effective drug for treatment of pain, fever and inflammation. Biosynthesis of salicylic acid **147** is achieved by recruiting chorismate **128** as the precursor, which is converted to isochorismate **146** catalyzed by isochorismate synthase.

Sequentially, isochorismate pyruvate lyase is responsible for converting isochorismate **146** into salicylic acid **147**.<sup>93</sup>



**Figure 1.3.3** Bioactive compounds derived from the shikimate pathway. Stereochemistry of **150** and **151** are unknown.



**Figure 1.3.4** Involutin **153** and its analogues.

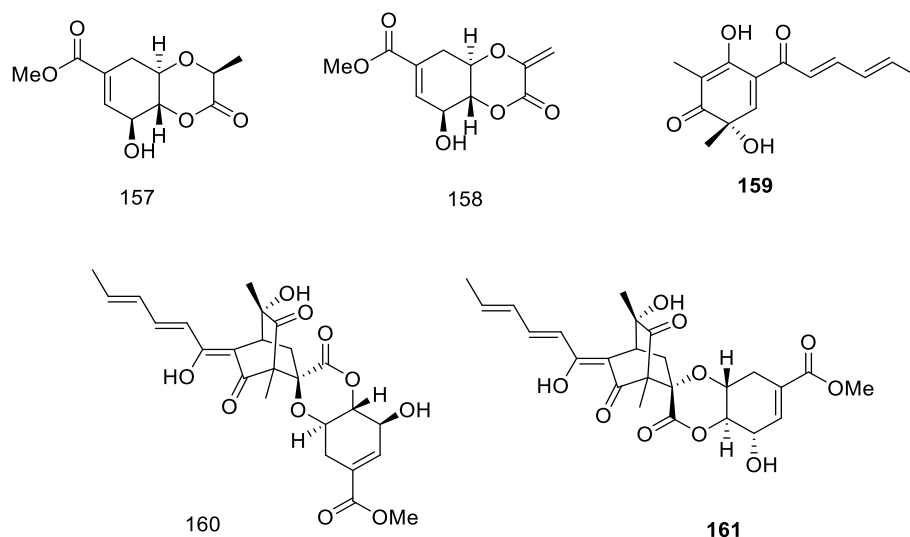
Involutin **153** is a diarylcyclopentenone secreted by the ectomycorrhizal fungus *Paxillus involutus*, which plays an important role as an  $\text{Fe}^{3+}$  reductant in Fenton-based decomposition of organic matter by mobilizing embedded nutrients, thereby increasing their accessibility to the host plant.<sup>94</sup> Stable-isotope labeling confirms *L*-tyr **130** as the metabolic origin of the terphenylquinone atromentin **152**,

which serves as the precursor of involutin **153** (Figure 1.3.4).<sup>95</sup> Structurally related shikimate-derived metabolites, including pulvinic acid **154**, polyporic acid **156**, and phenguignardic acid **155** are pigments or antioxidants and can show antimicrobial and cytotoxic activities.<sup>96-98</sup>

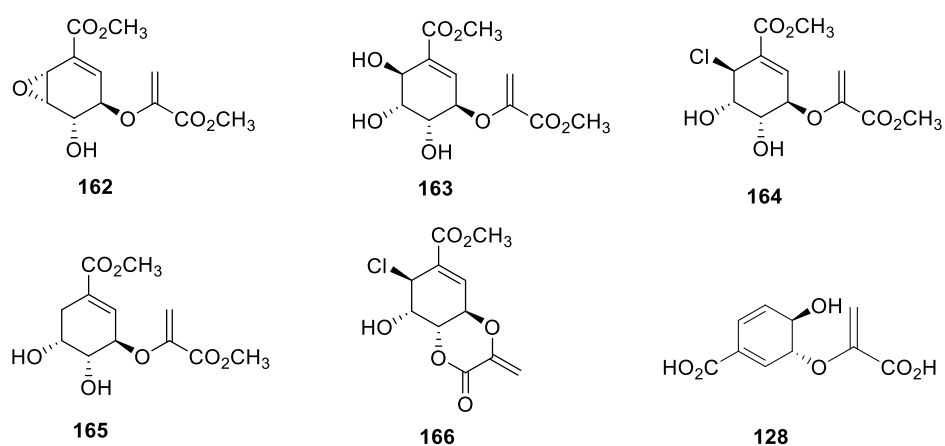
#### 1.4 Scytolide and Related Natural Products

Scytolide **158** was first reported in 1993, as a metabolite produced by *Scytalidium uredinicola* grown on malt extract. Scytolide inhibits germination of spores of *Endocronartium harknessii*, which causes a fungal disease of pine trees named western gall rust. Based on its structure, scytolide **158** was proposed to be derived biogenetically from the shikimate pathway.<sup>99</sup> After 10 years, scytolide was also isolated and identified from the fungus *Phyllosticta cirsii* together with the closely related metabolite phyllostin **157**. Their absolute configurations were assigned based on ORD, ECD and VCD analysis.<sup>100</sup> Both scytolide **158** and phyllostin **157** show promising bioactivity in suicidal germination, an alternative strategy for weed control of some *Orobanch*e species, which are parasitic weeds of important agrarian crops including tomato, cabbage, sunflower, and legumes.<sup>101</sup> In 2019, scytolide was also identified as the precursor of spirosorbicillinols A **160** and B **161** in *Trichoderma reesei* since knockout of *sorA* in *T. reesei* allows scytolide **158** to accumulate. The intermolecular Diels Alder (DA) reaction between sorbicillinol **159** and scytolide **158** catalyzed *in vivo* by the flavin-dependent monooxygenase SorD results in the formation of spirosorbicillinols A **160** and B **161**.<sup>102</sup> Even though the biosynthetic pathway of sorbicillinol **159** has been clarified,<sup>103</sup> the biosynthesis of scytolide **158** is still unknown.

In 1990, the first chlorine-containing shikimate-related metabolite compound **166** was reported from two fungi by two Japanese groups jointly. The structure and absolute stereochemistry were established by spectroscopic and X-ray crystallographic studies. In consideration of the structure of compound **166**, it is proposed to be derived from chorismic acid **128**.<sup>104</sup> Screening of the genus *Clitocybe cyathiformis* for new fungal-derived natural products led to isolation and identification of an unexpected series of alicyclic metabolites including cyathiformines A-D **162-165** and the reported chlorine-containing shikimate-related metabolite **166** (Figure 1.4.2).<sup>105</sup> The structures of compounds **162-165** suggest a biosynthetic origin from chorismate **128** by an *in vivo* epoxidation reaction. The chemical synthesis of compound **166**, cyathiformines A **162** and C **164** based on the biosynthetic hypothesis have been published.<sup>106,107</sup>



**Figure 1.4.1** Structures of scytolide **158**, phyllostin **157**, sorbicillinol **159**, spirosorbicillinols A **160** and B **161**.



**Figure 1.4.2** A few natural analogues of scytolide **158**.

## 1.5 Method for biosynthetic studies

### 1.5.1 Genome Sequencing

Whole-genome sequencing provides a deep insight into the DNA sequence of microbial genomes, which enables the identification of natural product biosynthetic gene clusters and comparison of multiple genomes. The rapid technological expansion in next generation sequencing represented by Illumina technology is making genomic information more accurate with exponential

advancement.<sup>108–110</sup> Illumina sequencing method is based on fluorescent dNTPs polymerized into a template strand. The incorporated dNTP is recognized by fluorophore excitation with millions of PCR-amplified gDNA fragments sequenced in parallel, which enables the identification of single bases.<sup>111</sup> Illumina sequencing operates in Paired-End sequencing that involves sequencing both ends of the DNA fragments in a library and aligning the forward and reverse as read pairs. The alignment algorithm uses the information to map the reads over repetitive regions precisely as the distance between each paired read is known. This provides highly accurate but short read alignment and facilitates detection of genomic rearrangements such as insertions, deletions, inversions, and gene fusions.<sup>112</sup>

Oxford nanopore sequencing is another common method for genome sequencing, in which a single molecule of DNA is passed through a nanopore and is sequenced without the need for PCR amplification or chemical labelling of the sample. Nanopore sequencing exploits electrophoresis to transport an unknown sample through an orifice of diameter  $10^{-9}$  meters.<sup>113</sup> Sequencing is possible because characteristic changes in the density of the electric current flowing through the nanopore depend on the composition of DNA that occupies the nanopore. Oxford nanopore sequencing gives very long reads, but with lower accuracy than Illumina sequencing.<sup>114</sup> At the moment, combining the results from Illumina sequencing and Oxford nanopore sequencing is ideal to provide the long and precise sequence data that is required for finding secondary metabolite BGC in microorganisms.<sup>115</sup>

### **1.5.2 Fungal Transformation and Gene Knockout**

Target gene knockout is an effective tool to elucidate the function of a single gene in the biosynthetic pathway of natural products. The loss of a functional enzyme interrupts the synthetic pathway, often resulting in the loss of target natural products or accumulation of intermediates. This shows *in vivo* potentially chemical steps involved in the biosynthesis of the target product. A commonly used strategy to disrupt a gene of interest from the biosynthetic gene cluster (BGC) is based on stable integration of an antibiotic resistance gene into the native genome locus by fungal transformation, which is usually achieved by a polyethyleneglycol (PEG)/  $\text{CaCl}_2$  method.<sup>116</sup>

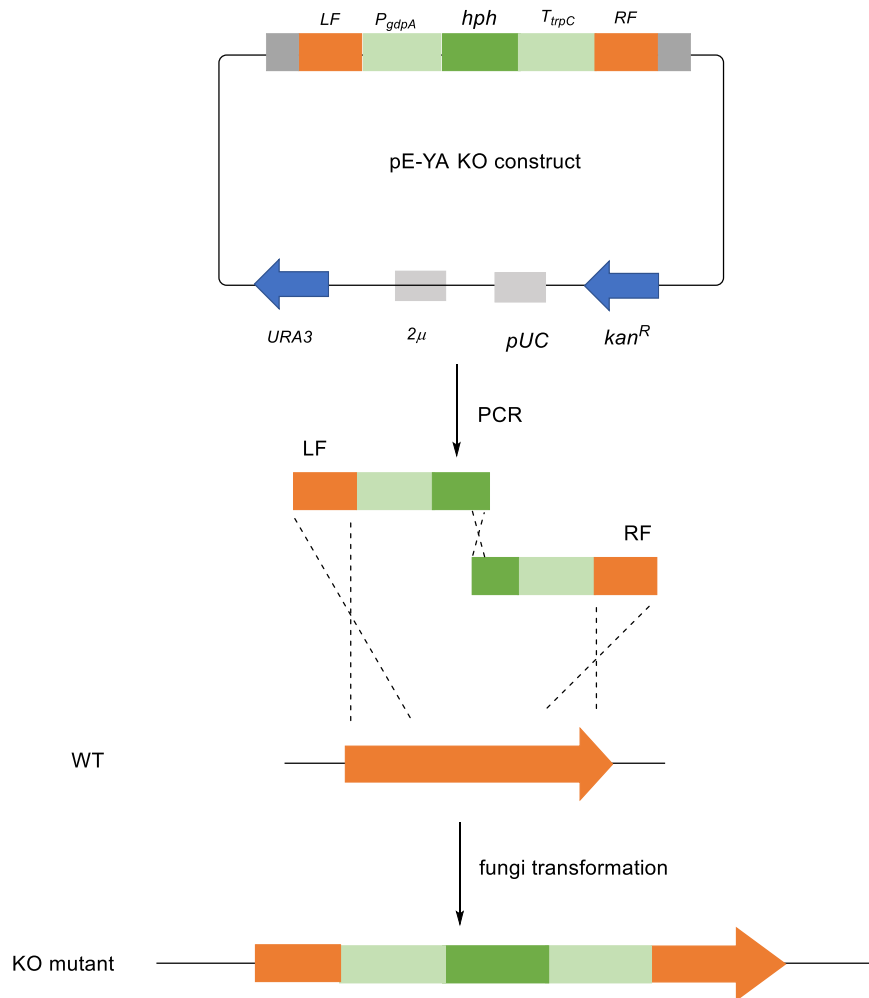
To transform fungal cells, the cell wall must be digested with hydrolytic enzymes to allow the penetration of exogenous DNA. The fungal cell wall consists of glycoproteins and polysaccharides;

mainly mannoproteins, 1,3-glucan and chitin, which can be digested by fungal cell wall hydrolases that show cellulase, chitinase or protease activities.<sup>117,118</sup> Concentrated mixtures of lysing enzymes extracted from *Trichoderma harzianum* and *basidiomycota* are known to be effective in the total digestion of fungal cell walls.<sup>119,120</sup> The resulting digested single cell, namely a protoplast, is easy to transform with exogenous DNA. However, protoplasts are particularly sensitive to osmotic pressure, and must be preserved with proper saline buffer.

To increase the opportunity of destroying the targeted gene, an efficient bipartite knockout method has been described by Nielsson and co-workers.<sup>121</sup> It is a precise site-directed tool that exploits the natural occurring homologous recombination present in filamentous fungi for gene deletions, promoter replacements, in-frame GFP fusions and specific point mutations. The method involves adding two overlapping, non-functional DNA fragments, amplified from the constructed KO cassettes by PCR, by protoplasts transformation. Three events of recombination are required for rebuilding the select marker gene and destroying the target gene (Figure 1.5.2.1).<sup>122</sup> Transformed protoplasts are then plated on agar with antibiotics and the growing colonies are transferred onto new plates with the same concentration of antibiotic for another two times to find real mutants. Transformant mycelia with selection marker (e.g. *hph*) are resistant to the antibiotic, whilst wild type fungi are incapable of surviving. However, the colonies growing on antibiotic selection plates could get the selection marker gene by either ectopic integration or homologous re-combinations, and transformants obtaining selection marker gene by ectopic integration are false positive transformants as the target gene is not destroyed in this way.

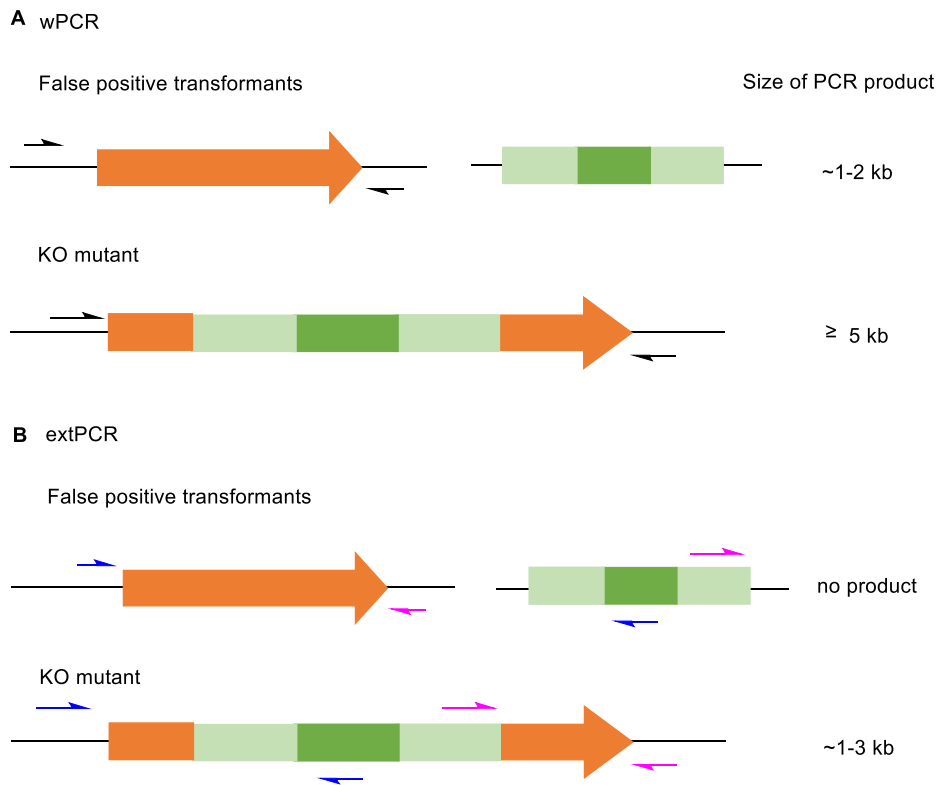
False positive transformants are ruled out by genomic analysis with PCR. Two kinds of diagnostic PCR are used: whole PCR (wPCR) and external PCR (extPCR, Figure 1.5.2.2). wPCR spans the entire target gene and reveals whether the WT gene is still in place. The PCR product from real mutant is much longer than that from false positive transformants, sometime no product is observable in a real mutant as the Taq polymerase is incapable to yield the large PCR product. extPCR on the other hand is designed to cover the sites of homologous recombination at each end of the target gene and indicates site-specific integration of the exogenous DNA fragments. In this case, the PCR product should only be observed from real mutants instead of false positive transformants. Further confirmation can be gained by sequencing the extPCR products.





**Figure 1.5.2.1** Bipartite targeted knockout strategy showing amplified parts (KO LF and KO RF) for fungal transformation.

Other gene KO methods worth mentioning include transformation mediated by *Agrobacterium tumefaciens*, CRISPR/Cas directed genome editing, and post-transcriptional gene silencing. *A. tumefaciens* is a bacterial plant pathogen that causes tumors in the host by introducing and integrating a segment of DNA (namely T-DNA).<sup>123</sup> CRISPR/Cas exploits the bacterial immune system against virus infections to edit targeted DNA sequences with the help of guided RNA.<sup>124</sup> Post-transcriptional gene silencing is a method that relies on degradation of targeted mRNA that transfers the genetic information between the transcriptional and translational machinery.<sup>125</sup>



**Figure 1.5.2.2** Two kinds of diagnostic Tag PCR for identifying true transformants: **A**, wPCR, from start to end of the gene. **B**, extPCR, from the edge of the HygR to a region outside the target gene.

### 1.5.3 Heterologous Expression

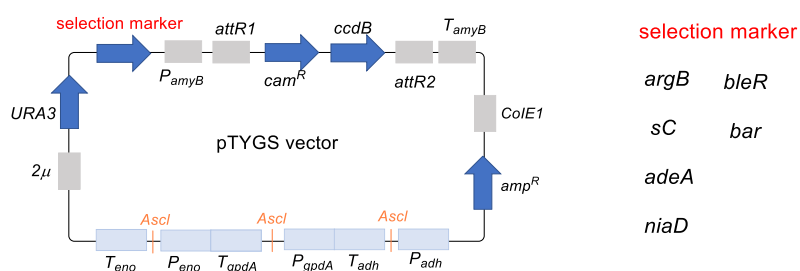
Heterologous expression allows the reconstruction of partial or entire biosynthetic pathways of natural products in a different host strain by transforming a set of exogenous genes under the control of constitutive or inducible promoters into the genome of the host strain.<sup>126</sup> *Aspergillus oryzae*, *Aspergillus nidulans* and *Saccharomyces cerevisiae* are three common fungal hosts for the reconstruction of the biosynthesis of fungal metabolites.<sup>127–129</sup> They are auxotrophic strains that could survive after being transformed with vectors carrying the genes to reintroduce autotrophy plus the genes of interest. The first successfully expressed complete multigene fungal cluster was the tenellin gene cluster in an arginine auxotrophic *A. oryzae* M-2-3 strain.<sup>130</sup> Later, more complete and partial biosynthetic pathways have been expressed in fungal host for the biosynthesis of secondary metabolites such as byssochlamic acid and squalestatin.<sup>131,132</sup>

Heterologous expression experiments in *A. oryzae* are useful tools to allow the verification and elucidation of biosynthetic pathway for fungi secondary metabolites.<sup>133</sup> As a heterologous host, *A. oryzae* has a number of advantages. Firstly, this species has been given the status of "Generally

Regarded As Safe" organism, accordingly it is safe to be used for gene manipulation. Secondly, it is usually stable and easily cultured in the lab. Furthermore, practical and efficient transformation protocol for genome modification in *A. oryzae* is already built up.<sup>134–136</sup>

Reconstitution of larger biosynthetic pathways have been achieved in the quadruply auxotrophic *A. oryzae* NSARI. This fungus is deficient in arginine (*argB*) biosynthesis, adenine (*adeA*) biosynthesis, sulfur metabolism (*sC*) and nitrate reduction (*niaD*). These metabolic pathways are necessary for the growth of *A. oryzae* NSARI. The four auxotrophies can be exploited for transformation of fungal expression vectors by complementation with the four different selection markers (*argB*, *sC*, *adeA*, *niaD*) or by supplementing with arginine, adenine, methionine or ammonium salts respectively.<sup>137</sup> In addition, the natural sensitivities of *A. oryzae* NSARI towards the antibiotics bleomycin and the herbicide glufosinate (basta) provide two more selection markers (*bleR* and *bar*).

The pTYGS vectors are a series of fungal expression vectors that are available with the above mentioned auxotrophic and antibiotic selection markers containing four sets of fungal promoters/terminators (*P/T<sub>amyB</sub>*, *P/T<sub>adh</sub>*, *P/T<sub>gpdA</sub>* and *P/T<sub>eno</sub>*). The first three are flanked by an *AscI* site that can be used to simultaneously add up to three tailoring genes *in vivo* by yeast homologous recombination. This allows simultaneous expression of up to 24 individual genes in *A. oryzae* NSARI. Furthermore, the expression vectors contain an origin of replication *colE1* for replication in *E. coli* and an ampicillin resistance gene *amp<sup>R</sup>* as well as a chloramphenicol gene (*cam<sup>R</sup>*). For selection and propagation in uracil deficient *S. cerevisiae* strains, the vector contains the *URA3* gene, encoding orotidine 5'-phosphosphate decarboxylase, and the *2μ* origin of replication. Six different versions of the vectors are available, each with one of the six above-mentioned selection markers (Figure 1.5.3).<sup>138</sup>



**Figure 1.5.3** pTYGS vector families used for heterologous expressions in *A. oryzae* NSARI

#### 1.5.4 Enzyme Assay

The bacterium *Escherichia coli* (*E. coli*) is most preferred system used for heterologous enzyme production. Many tailoring enzymes from fungi have been characterized *via* bacterial heterologous expression, such as methyltransferases, cytochrome P450 oxygenases and FAD-dependant monooxygenase. For the investigation of fungal enzymes, they can be expressed in active form if the introns from its corresponding gene are removed, as *E. coli* is incapable of splicing out introns of eukaryotic genes. Yet, there are some main challenges in using bacterial expression for identifying the function of enzymes from fungi.<sup>139</sup> For example, in *E. coli*, many eukaryotic enzymes do not fold correctly resulting in the formation of insoluble and inactive proteins. Despite the appropriate expression system, some proteins express at very low level or not at all. In order to improve the amount of soluble protein production, a number of vectors, their compatible hosts and culture conditions can be tested for optimization. Also, vectors with the fusion tags/chaperons facilitate protein expression in soluble fraction and assist in proper protein folding.<sup>140</sup> A further requirement for successful *in vitro* assay of fungal biosynthetic systems is the availability of suitable substrates. Sometimes these may be commercially available, but often they must be first isolated from WT or engineered fungal strains and are therefore only available in low amounts. In addition, suitable analytical chemistry assays must be available or developed in order to observe enzyme activity. Thus *in vitro* enzyme assays of fungal biosynthetic may often require significant work.

#### 1.5.5 Isotopic Feeding Experiments

Isotopic labelling studies have been used for many years to investigate biosynthetic pathways of natural products. The method reveals information about the starter unit, the biosynthetic origin and important insights into biosynthetic pathways of natural products.<sup>141</sup> Results of labelling studies can also give information about biosynthetic mechanisms in some cases. Early labelling studies were conducted using radioactive labels, such as <sup>14</sup>C and <sup>3</sup>H. The disadvantages of this technique are not only the harm of radioactivity but that the determination of the exact site of isotope incorporation in the molecule involves complex chemical degradation and further radio-detection. With the advent of NMR, feeding building blocks labeled with stable isotopes such as <sup>13</sup>C, <sup>2</sup>H, <sup>15</sup>N and <sup>19</sup>F give deeper structural information with much less effort. For example, incorporation of a singly <sup>13</sup>C-labelled precursor enhances the content of <sup>13</sup>C at particular carbons which result in an increase in intensity of

the carbon signal in the  $^{13}\text{C}$  NMR spectrum compared to natural abundance of  $^{13}\text{C}$  (1.1 %).<sup>142</sup>  $^{18}\text{O}_2$  incorporates into keto- and hydroxy groups and causes a slight chemical shift of the tethered  $^{13}\text{C}$ , by the effect of the greater mass of the oxygen isotope.<sup>143</sup> Multiple labels such as  $[1,2\text{-}^{13}\text{C}_2]$ -acetate result in the splitting of the signals in a pair of doublet with identical coupling constant  $J$ , rendering easy to understand atom connectivity.<sup>144</sup>

## 2 Biosynthesis of Tropolone Sesquiterpenoids

Tropolone sesquiterpenoids (TS) are an intriguing family of fungal meroterpenoids that are proposed to derive from a unique intermolecular hDA reaction between humulene and tropolones. These natural products have been recognized for their impressive biological activities, such as antimalarial, antimicrobial and anticancer (chapter 1.2.3). A few unnatural TS have been synthesized *via* the hDA addition reaction between tropolone *o*-QM and *2E*-humulene by different groups in recent years (chapter 1.2.4). The biosynthesis of both tropolones and xenovulene A have been elucidated by the Cox group, including the identification of the BGCs and the chemical reaction catalyzed by functional proteins in the biosynthetic pathway (chapter 1.2.5-6). Undoubtedly, these researches shed light on further investigating the biosynthesis of more complex and biologically active TS family members including pycnidione **22**, eupenifeldin **23** and their co-metabolites.

### 2.1 Project Aims

Understanding and engineering the biosynthesis of tropolone sesquiterpenoids such as pycnidione **22** and eupenifeldin **23** is the key focus of this project. For this purpose, the secondary metabolites from two producing fungi (*CF236968*, *CF253087*) will be analyzed in detail. Previously reported pycnidione **22** and eupenifeldin **23** and related novel compounds will be isolated and purified. All their structures will be elucidated based on subsequent HR-MS and full NMR analysis.

No genomic data of *CF236968* and *CF150626* are available in the literature, and no biosynthetic cluster of pycnidione **22** and eupenifeldin **23** had been reported in any other producer strain at the start of this project, therefore the composition of pycnidione and eupenifeldin BGCs are yet to be determined. Sequencing of the whole genome for three producing fungi will be conducted to enable a full bioinformatics analysis and the prediction of the possible BGC for pycnidione **22** and eupenifeldin **23**. In order to more clearly link the clusters with the biosynthesis of eupenifeldin **23** and delineate the boundaries of the co-transcribed genes, the mRNAs of *CF150626* under producing and non-producing conditions will be prepared and used for reverse transcription polymerase chain reaction (RT-PCR). At the same time, the most promising putative BGC for eupenifeldin biosynthesis, and the included functional genes, will be confirmed and analyzed through gene knockout in fungi *CF150626*.

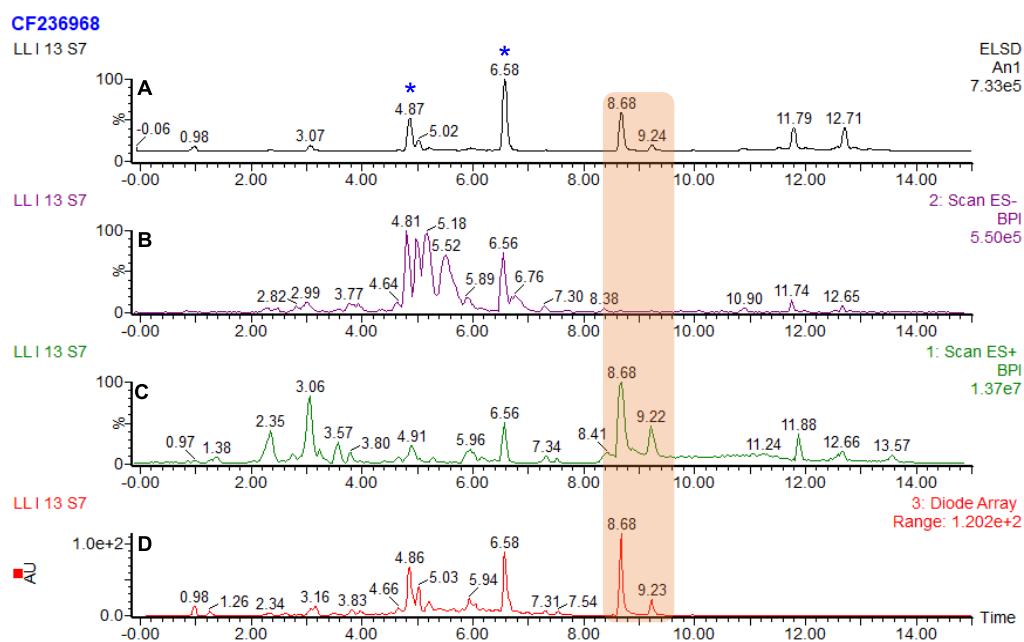
Early tropolone biosynthesis in *CF150626* is proposed to be similar to Xenovulene A **24** biosynthesis in *Acremonium strictum*. Considering the structural difference between Xenovulene A **24** and eupenifeldin **23**, it will be of interest to investigate the function of the different genes between the BGC of xenovulene A **24** and the BGC of eupenifeldin **23** by heterologous expression in *A. oryzae* or *in vitro* assay. Since both pycnidione **22** and eupenifeldin **23** are ditropolone sesquiterpenoids, so the identification of the enzymes responsible for the hDA addition reaction that connecting two tropolone precursors to one humulenol precursor are of special interest.

## 2.2 Results-Analysis of Tropolone sesquiterpenoids Produced by Fungi

In order to find the fungi that produce the tropolone sesquiterpenoids (pycnidione **22** and eupenifeldin **23**), we investigated the metabolites of fungi *CF236968*, *CF253087* and *Pleosporales sp 1987* in different media with LCMS. *CF236968*, *CF253087* are two TS producing fungi kindly provided by the Singh's group from Fundación MEDINA (Granada, Spain). They reported the discovery of noreupenifeldin and eupenifeldin from *CF150626* in 2008.<sup>37</sup> According to reported literature, pycnidione **22** and eupenifeldin **23** are diastereomers and possess the same molecular mass ( $m/z = 548$ ).

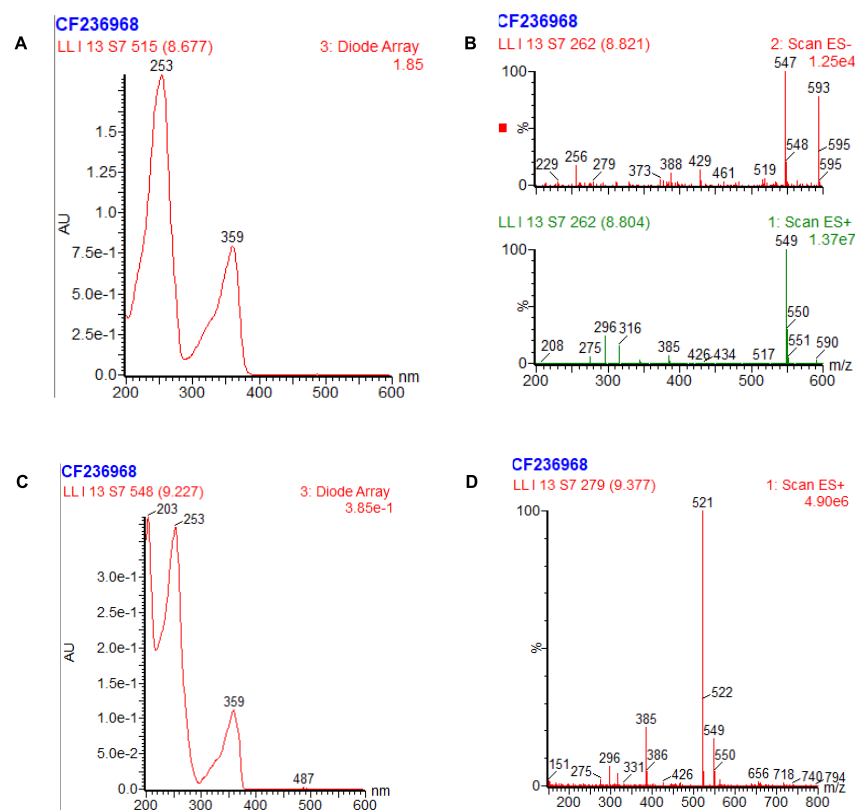
### 2.2.1 Analysis of Secondary Metabolites from *CF236968*

*Leptobacillium sp. CF236968* was typically cultured in production media (PD agar) at 22 °C in the dark. After 14 days' cultivation, two PD plates were cut into small pieces and transferred to a clean beaker. 200 ml ethyl acetate (EtOAc) was added to extract the metabolites with 4 hour's stirring. Solvents were removed by rotary evaporation and dried extracts was dissolved in dichloromethane:methanol (1:1) to prepare 10 mg/mL sample for analytical high performance liquid chromatography (HPLC, Figure 2.2.1.1). The HPLC chromatogram of wild type *CF236968* showed a single dominant peak at  $t_R = 8.7$  min with UV absorption  $\lambda_{max}$  253 and 359 nm. Extracted ion chromatograms using 549 Da ( $ES^+$  spectrum) led to the identification of a peak corresponding to pycnidione **22** (Figure 2.2.1.2). Additionally, the minor component eluting at  $t_R = 9.2$  min with UV absorption  $\lambda_{max}$  255 and 361 nm was proposed to be a new tropolone sesquiterpenoids. The mass spectra of the new peak showed the molecular weight is 520 based on the 521 Da ion signal in the  $ES^+$  spectrum (Figure 2.2.1.2).



**Figure 2.2.1.1** The LCMS chromatogram of *Leptobacillum* sp. CF-236968 culture extracts after 14 days' growth on PD agar, tropolone sesquiterpenoids are highlighted with orange color: **A**, ELSD chromatogram of the culture extracts; **B**, ES<sup>-</sup> TIC chromatogram of the culture extracts; **C**, ES<sup>+</sup> TIC chromatogram of the culture extracts; **D**, UV DAD chromatogram of the culture extracts. TS unrelated compounds are highlighted with \*.

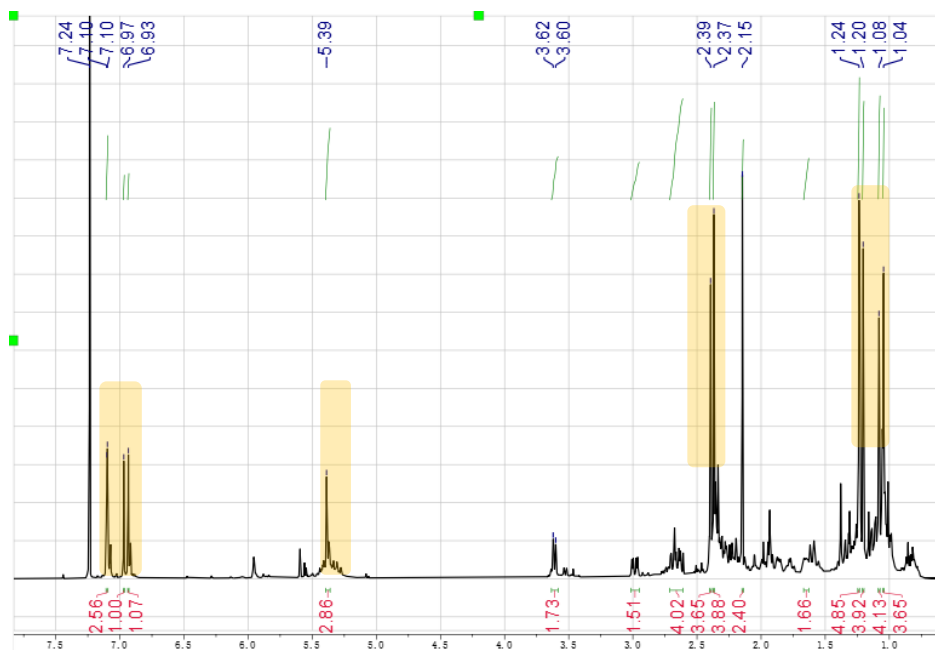




**Figure 2.2.1.2** Characterisation of pycnidione **22** and the new peak eluting at 9.2 min: **A**, UV spectrum of pycnidione **22**; **B**, MS spectrum of pycnidione **22**; **C**, UV spectrum of the new peak; **D**, MS spectrum of the new peak.

## 2.2.2 Structural Identification of Pycnidione **22** and Epolone **C 42**

To confirm the assignment of the compound eluting at 8.7 min as pycnidione **22** and identify the structure of the new peak eluting at 9.2 min, the two peaks were targeted for preparative-LCMS purification and NMR analysis. *CF236968* was grown on 20 PD agar plates for 14 d at 22 °C. The mycelia together with agar were cut into small pieces and then extracted with 2 L ethyl acetate by soaking and stirring for 4 hours. Solids were removed by filtration. The organic extract was filtered and evaporated to afford 1 g crude metabolites. In total, 10 mg pycnidione **22** and 5 mg new compound were purified by mass-directed preparative HPLC for full NMR analysis. The obtained  $^1\text{H}$  and  $^{13}\text{C}$  NMR data of major compound with  $t_{\text{R}} = 8.7$  min were identical to the literature data of pycnidione **22** (Table 2.2.2.1).<sup>33</sup>

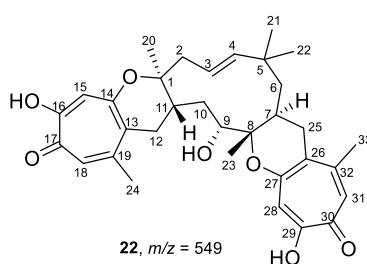


**Figure 2.2.2.1**  $^1\text{H}$  NMR spectrum of pycnidione **22**; The most identifiable signal was highlighted in color.

The new peak eluted at  $t_{\text{R}} = 9.2\text{min}$  had the molecular formula  $\text{C}_{32}\text{H}_{40}\text{O}_6$  with 13 degrees of unsaturation based on HRMS analysis, ( $[\text{M}] \text{H}^+$  calculated  $\text{C}_{32}\text{H}_{41}\text{O}_6$  521.2903, found 521.2905). The  $^1\text{H}$  NMR spectrum (Figure 2.2.2.2) of **42** revealed the presence of six singlet methyl [ $\delta_{\text{H}}$  2.42 (s, H<sub>3</sub>-27), 2.18 (s, H<sub>3</sub>-32), 1.22 (s, H<sub>3</sub>-31), 1.19 (s, H<sub>3</sub>-28), 1.10 (s, H<sub>3</sub>-29), and 1.05 (s, H<sub>3</sub>-30)], an oxygenated methine [ $\delta_{\text{H}}$  3.61 (d,  $J = 9.7$  Hz, H-11)] and six aromatic/olefinic proton [ $\delta_{\text{H}}$  7.12 (s, H-7), 6.97 (s, H-3), 6.27 (d,  $J = 2.5$  Hz, H-25), 6.21 (s,  $J = 2.5$  Hz, H-23), 5.44 (m), 5.42 (m)]. The  $^{13}\text{C}$  NMR spectra (Table 2.2.2.2) of **42** showed thirty-two carbon resonances ascribed to six methyls, five methylenes, nine methines (six aromatic/olefinic carbons, one oxygenated carbon and two aliphatic carbons), twelve quaternary carbons (one ketone group, eight aromatic/olefinic carbons, two oxygenated carbon and one aliphatic carbons). These data suggested that **42** was a tropolone sesquiterpenoids, which has similar structure as epolone A **41** possessing a tropolone ring and a phenol ring attached to a humulene ring.<sup>33</sup>

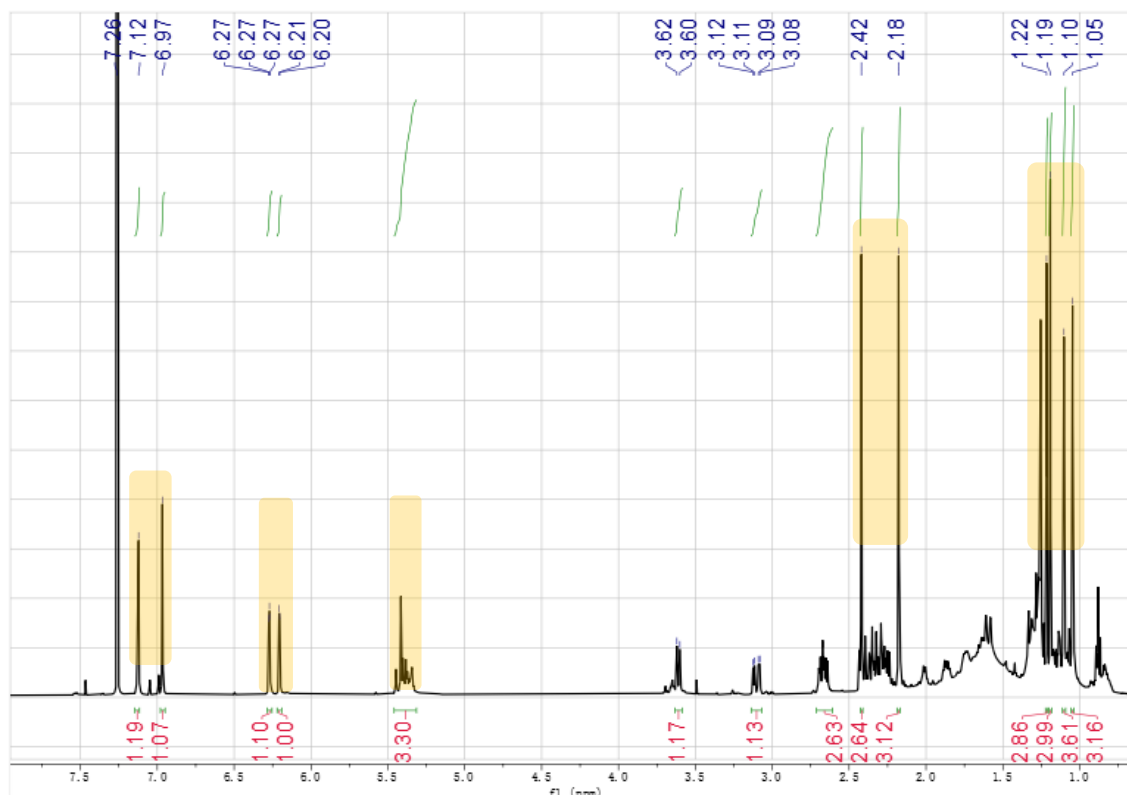
Characteristic proton signals of the methyl at  $\delta_{\text{H}}$  2.42 ppm and two aromatic methines at  $\delta_{\text{H}}$  7.12, 6.97 ppm indicated the presence of a tropolone ring, which was confirmed by the HMBC correlations from: H-15 to C-17, C-16, C-14, and C-13; H-18 to C-13, C-16, C-17, C-19 and C-24;

and from H-24 to C-18 and C-19. The  $^1\text{H}$ - $^1\text{H}$  COSY correlation of H-2/H-3, H-6/H-7, and H-9/H-10, together with HMBC correlations from H-20 to C-1, C-2, and C-11, H-21 and H-22 to C-4, C-5, and C-6, and H-23 to C-7, C-8, and C-9 established the humulene unit. Meanwhile, the phenol ring was constructed by the HMBC correlations of: H-28/C-26, C-27, and C-30; of H-30/C-26, C-28, C-29, and C-32; of H-32/C-26, C-30 and C-31 (Figure 2.2.2.3).

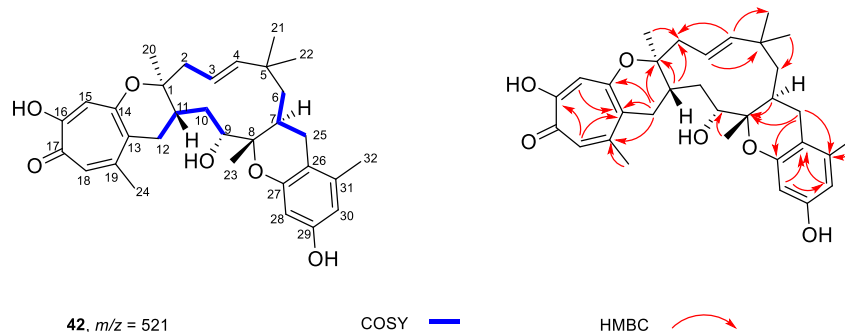


Position	$\delta_{\text{C}}/\text{ppm}$ <b>22</b>	$\delta_{\text{H}}/\text{ppm}$ <b>22</b>	$\delta_{\text{C}}/\text{ppm}$ <b>22</b> literature <sup>33</sup>	$\delta_{\text{H}}/\text{ppm}$ <b>22</b> literature <sup>33</sup>
1	81.5		81.6	
2	48.6	2.69, m; 2.30, m	48.6	2.69, m; 2.34, m
3	121.9	5.38, m	122.0	5.38, m
4	143.7	5.36, m	143.7	5.40, m
5	36.7		36.8	
6	46.1	1.60, m; 1.14, m	46.1	1.60, m; 1.12, m
7	31.4	1.65, m	31.4	1.65, m
8	83.6		83.7	
9	77.0	3.60, bd, 10	76.8	3.61, d, 10
10	33.6	1.60, bdd, (15, 10); 1.14, m	33.6	2.22 m; 1.12 m
11	39.6	1.85, m	39.6	1.85, m
12	33.9	2.98, m; 2.26, m	33.9	2.99, dd (5, 17); 2.26, m
13	121.8		121.9	
14	160.4		160.4	
15	113.5	6.93, s	113.6	6.93, s
16	163.9		163.9	
17	171.7		171.8	
18	124.0	7.09, s	124.1	7.10, s
19	149.4		149.4	
20	18.6	1.20, s	18.6	1.20, s
21	21.9	1.08, s	21.9	1.08, s
22	31.2	1.05, s	31.4	1.05, s
23	15.9	1.24, s	15.9	1.24, s
24	27.3	2.39, s	27.3	2.39, s
25	34.0	2.63, m; 2.31, m	34.0	2.67, m; 2.31, m
26	119.9		120.0	
27	159.7		159.7	
28	112.4	6.97	112.6	6.97, s
29	163.3		163.3	
30	172.6		172.5	
31	124.6	7.10	124.7	7.10, s
32	149.6		149.7	
33	27.1	2.36, s	27.2	2.36, s

**Table 2.2.2.1** NMR data for pycnidione **22** in  $\text{CDCl}_3$  (500 MHz). Literature data<sup>33</sup> was also measured in  $\text{CDCl}_3$  (500 MHz).



**Figure 2.2.2.2**  $^1\text{H}$  NMR spectrum of epolone C **42**; The most identifiable signal is highlighted in color.



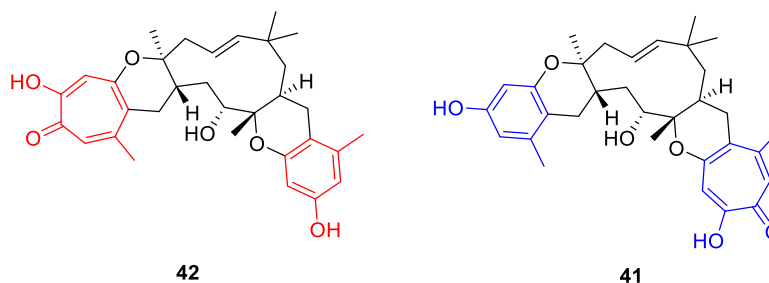
**Figure 2.2.2.3** Key COSY and HMBC correlations of epolone C **42**.

Thus, apart from 11 degrees of unsaturation occupied by tropolone ring, humulene ring and phenol ring, the remaining two degrees of unsaturation suggested that two additional rings existed in **42**. Correlations from H-12 to C-1, C-11, C-13, C-14 and C-19, plus the chemical shift values for C-1 ( $\delta_C$  81.9) and C-14 ( $\delta_C$  160.7), led to the postulation that tropolone ring and humulene ring were simultaneously connected by C-12 and an oxygen atom, forming a dihydro-2H-pyran ring. In turn, correlations from H-25 to C-7, C-8, C-27 and C-31, plus the chemical shift values for C-8 ( $\delta_C$  81.9)

and C-27 ( $\delta_C$  153.4), implied that C-25 was located between C-7 and C-26, and C-8 and C-27 were both connected to the same oxygen atom to form the second dihydropyran ring. Collectively, the planar structure of **42** was definitely assigned. The relative stereochemistry at positions C-1, C-7, C-8, C-9 and C-11 were assumed to be identical to pycnidione **22**, but not further determined. So, epolone C **42** was proved to be a diastereomer of epolone A **41** (figure 2.2.2.4)

Pos	$\delta_C$ /ppm <b>42</b> referenced to 77.0 ppm	$\delta_H$ /ppm <b>42</b> referenced to 7.26 ppm	COSY	HMBC
1	81.9			
2	48.6	2.70, m; 2.36, m		
3	121.6	5.42, m	2	4, 5, 2
4	144.3	5.44, m		22, 21, 2, 3, 5
5	36.7			
6	46.9	1.62, m; 1.15, m		
7	32.0	1.77, m	25, 6	8, 6
8	81.9			
9	77.5	3.61, d, 9.7	10	10, 11
10	33.7	2.30, m; 1.17, m	9, 11	
11	40.0	1.89, m	10, 12	10, 1, 2, 20
12	33.5	3.14, dd, 17, 5 2.37, m		11, 1, 14, 13, 19
13	122.6			
14	160.7			
15	113.7	6.97, s		13, 14, 16, 17
16	164.1			
17	171.5			
18	123.9	7.12, s		16, 13, 15, 24, 19, 17
19	149.0			
20	18.9	1.22, s		11, 1, 2
21	31.4	1.05, s		4, 5, 6, 22
22	22.1	1.10, s		4, 5, 6
23	15.2	1.19, s		7, 8, 9
24	27.2	2.42, s		19, 18
25	30.5	2.69, m; 2.27, m		7, 8, 31, 27
26	112.4			
27	153.4			
28	101.0	6.21, d 2.4		30, 26, 27
29	154.6			
30	109.1	6.27, d 2.5		32, 28, 26, 29
31	138.0			
32	19.2	2.18, s		26, 30, 31

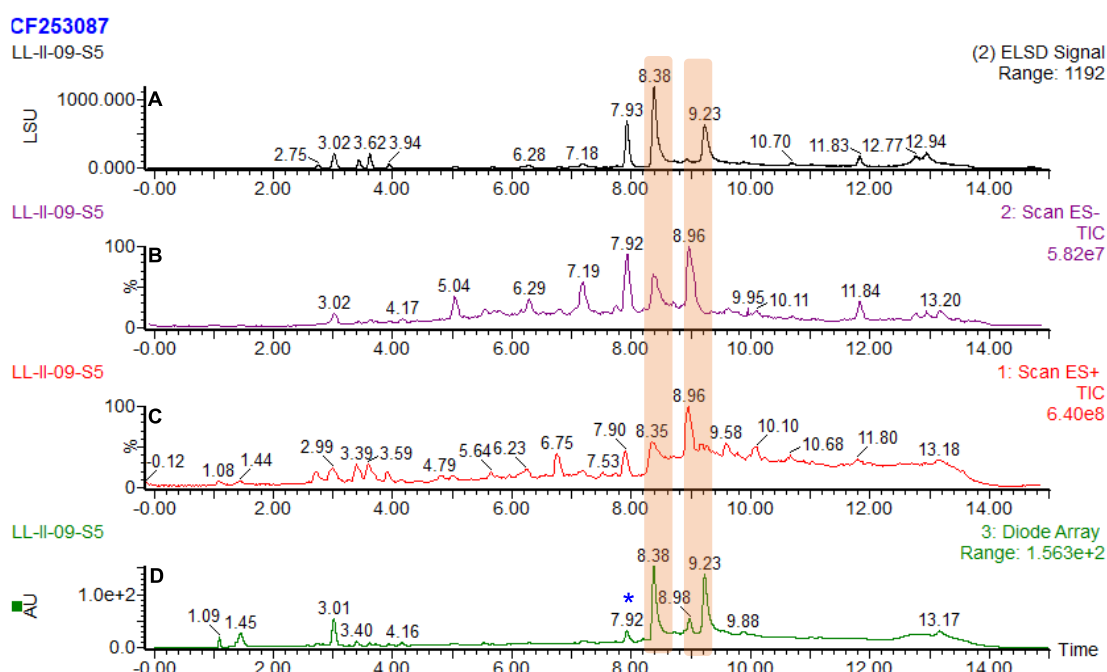
**Table 2.2.2.2** NMR data for epolone C **42** in CDCl<sub>3</sub> (500 MHz).



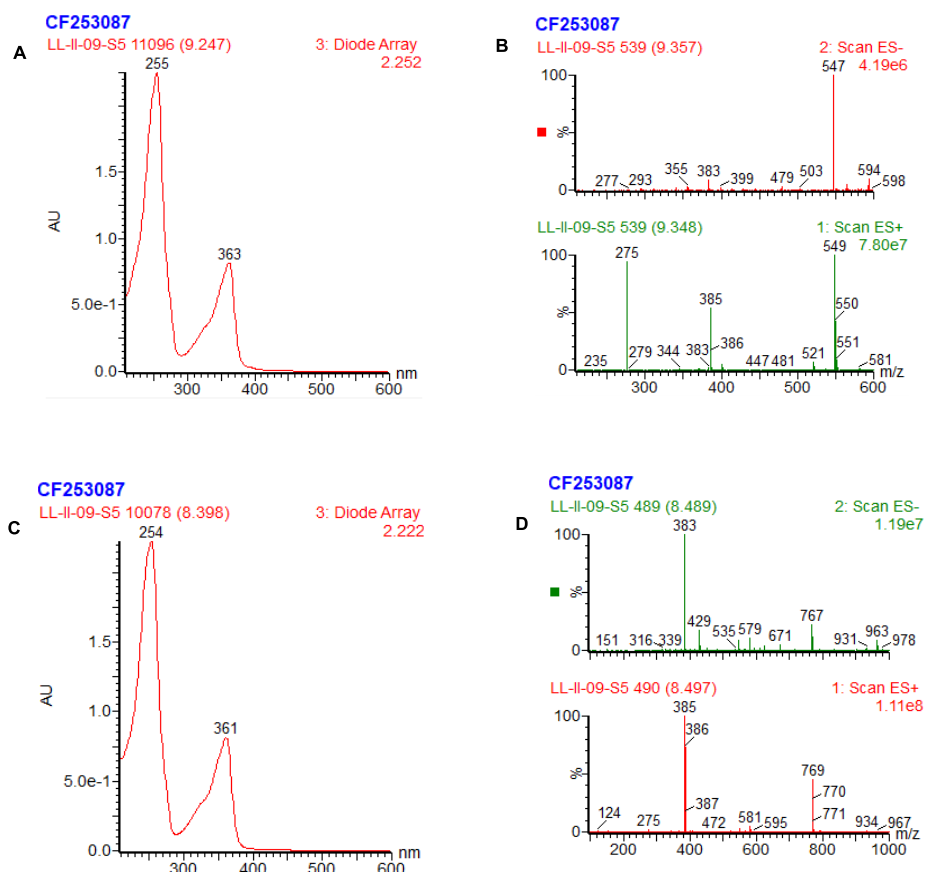
**Figure 2.2.2.4** Structural difference between epolone C **42** and epolone A **41**.

### 2.2.3 Analysis of Secondary Metabolites from *CF253087*

Fungus *CF253087* was also cultured on production media (PD agar) for 19 d at 22 °C. Two PD plates were cut into small pieces and transferred to a clean beaker. 200 ml EtOAc was added to extract the metabolites with 4 hour's stirring. Solids were removed by filtration. Solvents were removed by rotary evaporation and the dried extracts were dissolved in dichloromethane:methanol (1:1) to prepare 10 mg/mL sample for analytical HPLC. The HPLC chromatogram of wild type *CF253087* (Figure 2.2.3.1) showed a minor peak at  $t_R = 9.2$  min with UV absorption  $\lambda_{max}$  255 and 363 nm (Figure 2.2.3.2A). Extracted ion chromatograms using 549 Da ( $ES^+$  spectrum) led to the identification of eupenifeldin **23**. Additionally, the dominant peak eluting at  $t_R = 8.4$  min with UV absorption  $\lambda_{max}$  254 and 361 nm was proposed to be a related tropolone sesquiterpenoid (Figure 2.2.3.2C). The mass spectra of the new peak showed the molecular weight is 384 based on the 385 Da ion signal in the  $ES^+$  spectrum and 383 Da ion signal in the  $ES^-$  spectrum (Figure 2.2.3.2D).



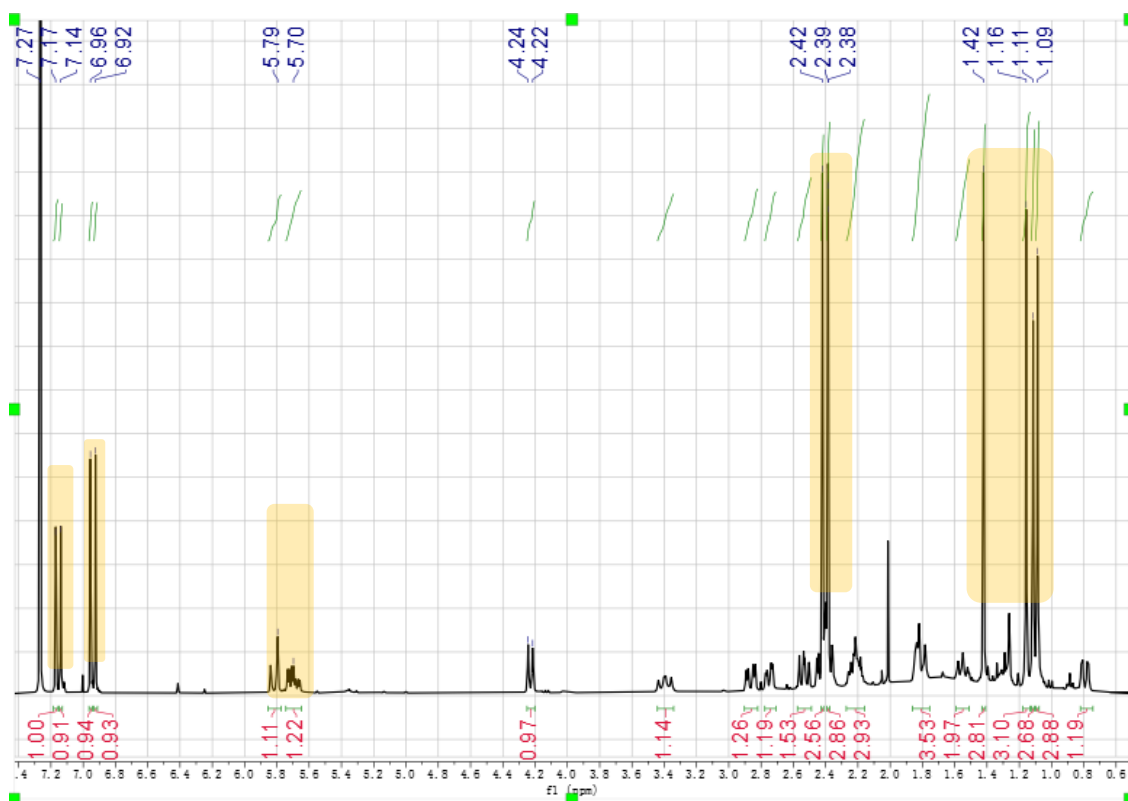
**Figure 2.2.3.1** The LCMS chromatogram of *CF 253087* culture extracts after 19 days' growth on PD agar, tropolone sesquiterpenoids are highlighted with orange color: **A**, ELSD chromatogram of the culture extracts; **B**,  $ES^-$  TIC chromatogram of the culture extracts; **C**,  $ES^+$  TIC chromatogram of the culture extracts; **D**, UV chromatogram of the culture extracts. TS unrelated compounds are highlighted with \*.



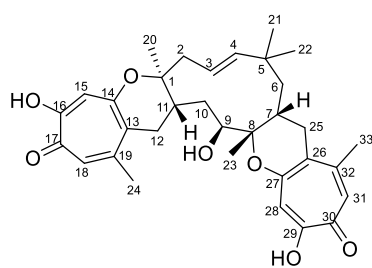
**Figure 2.2.3.2** Characterisation of eupenifeldin **23** and the new peak eluting at 8.4 min: **A**, UV spectrum of eupenifeldin **23**; **B**, MS spectrum of eupenifeldin **23**; **C**, UV spectrum of the new peak; **D**, MS spectrum of the new peak.

## 2.2.4 Structural Identification of Eupenifeldin **23** and Neosetophomone **B 61**

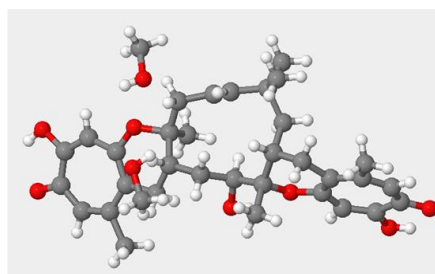
To confirm the assignment of the compound eluting at 9.1 min as eupenifeldin **23** and identify the structure of the new peak eluting at 8.4 min, the two peaks were targeted for preparative-LCMS purification and NMR analysis. *CF253087* were grown on 20 PD agar plates for 14 d at 22 °C. The mycelia together with agar were cut into small pieces and then extracted with 2 L ethyl acetate by soaking and stirring for 4 hours. After this, solids were removed by filtration and the organic extract was evaporated to afford 1g crude metabolites. In total, 10 mg eupenifeldin **23** and 3 mg new compound were purified by mass-directed preparative HPLC for full NMR analysis. The obtained  $^1\text{H}$  and  $^{13}\text{C}$  NMR data of major compound with  $t_{\text{R}} = 9.1$  min were identical to the literature data of eupenifeldin **23** (Table 2.2.4.1).<sup>39</sup> Besides, the relative configuration of eupenifeldin **23** was further confirmed with X-ray single crystal diffraction (Figure 2.2.4.2).



**Figure 2.2.4.1**  $^1\text{H}$  NMR spectrum of eupenifeldin **23**; The most identifiable signal was highlighted in color.



**23**,  $m/z = 549$



**Figure 2.2.4.2** Structure of eupenifeldin **23** and crystal structure of eupenifeldin **23** in methanol.

The purified compound eluting at  $t_R = 8.4$  min was observed as minor product in EtOAc extract of *CF253087*. This compound showed a nominal mass of 384, determined by low resolution mass spectrometry, indicating a molecular formula of  $\text{C}_{24}\text{H}_{32}\text{O}_4$  with nine degrees of unsaturation. It has characteristic UV absorption at 254 nm and 360 nm similar to eupenifeldin **23**. The  $^1\text{H}$  NMR spectrum of **61** revealed two aromatic proton singlets [ $\delta_H$  7.13 (s, H-18) and 7.00 (s, H-15)], and



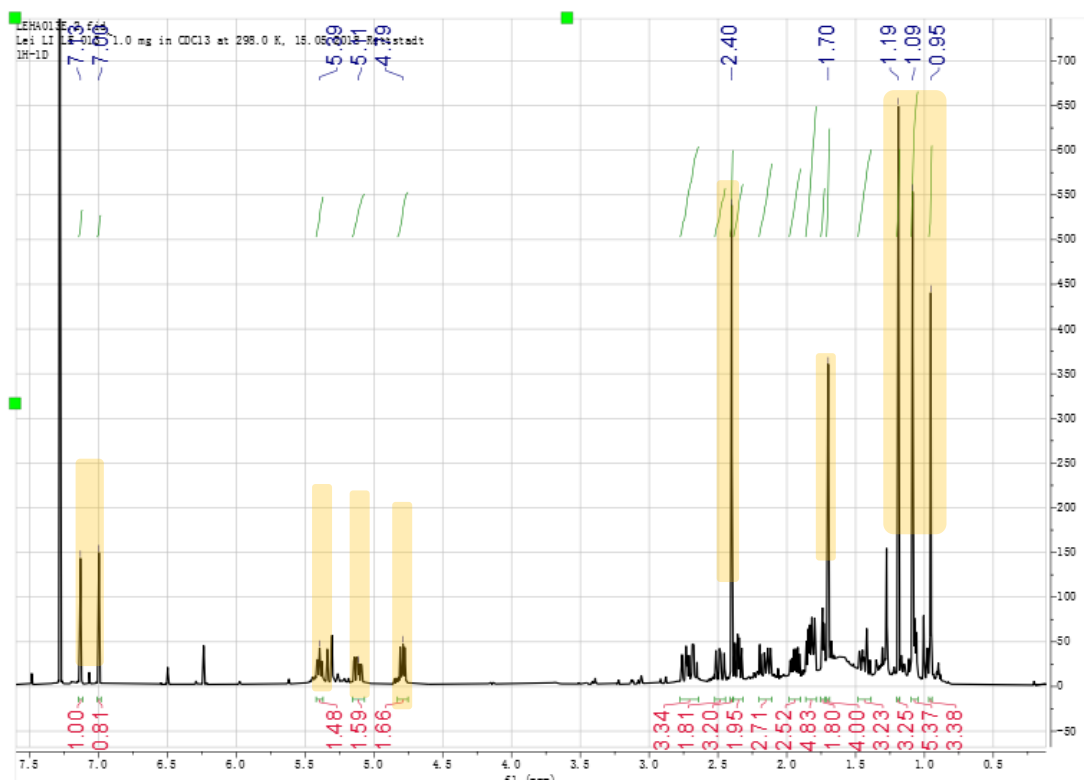
three olefinic proton [ $\delta_H$  5.12 (ddd,  $J = 15.8, 10.0, 2.9$  Hz, H-3), 5.33 (dd,  $J = 15.8, 1.9$  Hz, H-4) and 5.40 (ddd,  $J = 9.5, 7.6, 1.6$  Hz)], and five singlet methyl [ $\delta_H$  1.19 (s, H-20), 1.09 (s, H<sub>3</sub>-21), 0.95 (s, H<sub>3</sub>-22), 1.70 (s, H<sub>3</sub>-23), and 2.40 (s, H-24)] suggesting the presence of one tropolone ring and one humulene skeleton (Figure 2.2.4.3).

Position	$\delta_C$ /ppm <b>23</b>	$\delta_H$ /ppm <b>23</b>	$\delta_C$ /ppm <b>23</b> literature <sup>39</sup>	$\delta_H$ /ppm <b>23</b> literature <sup>39</sup>
1	80.7		80.4	
2	46.3	2.74, dd (13.4, 4.2) 2.53, dd (13.4, 11.1)	46.2	2.75, m 2.53, dd, 13.4, 10.5
3	125.8	5.69, ddd (16.0, 11.1, 4.2)	125.7	5.70, m
4	144.1	5.81, d (16.0)	143.9	5.81, dd, 16.1, 1.7
5	35.0		34.9	
6	46.6	1.79, d (14.7); 0.78, dd (14.7, 4.3)	46.5	1.80, m; 0.79, dd, 14.3, 4.1
7	32.0	1.83, dd (5.3, 4.3)	31.9	1.83, m
8	82.0		81.9	
9	70.8	4.22, d (11.2)	70.7	4.23, d, 11.2
10	30.2	2.21, dd (12.5, 5.9); 1.53, dd (12.5, 11.2)	30.1	2.24, m; 1.56, m
11	41.6	2.22, ddd (14.0, 5.9, 4.4)	41.4	2.21, m
12	33.1	3.40, dd (16.9, 14.0); 2.42, d (16.9, 4.4)	32.9	3.40, dd, 18.1, 13.5; 2.42, m
13	122.9		122.3	
14	160.6		160.1	
15	113.8	6.99, s	113.3	6.95, s
16	163.1		163.0	
17	172.2		172.5	
18	124.8	7.17, s	124.6	7.15, s
19	150.7		150.4	
20	19.9	1.41, s	19.3	1.42, s
21	29.7	1.08, s	29.7	1.09, s
22	27.2	1.11, s	27.0	1.11, s
23	16.1	1.15, s	16.0	1.15, s
24	27.5	2.42, s	27.3	2.42, s
25	34.4	2.86, dd (17.2, 5.3); 2.37, d (17.2)	34.2	2.84, dd, 17.3, 5.4; 2.38, m
26	118.7		118.5	
27	159.5		159.2	
28	112.8	6.93, s	112.6	6.92, s
29	162.9		162.8	
30	173.3		173.3	
31	125.6	7.17, s	125.5	7.17, s
32	151.5		151.3	
33	27.5	2.38, s	27.4	2.38, s

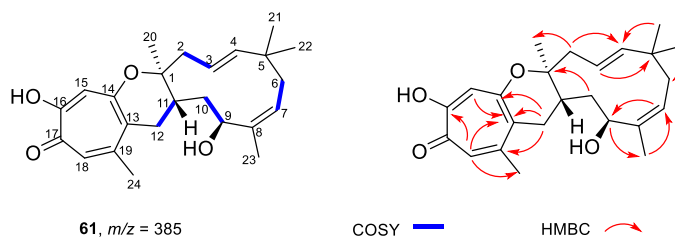
**Table 2.2.4.1** NMR data for **23** in CDCl<sub>3</sub> (500 MHz). Literature data<sup>39</sup> was measured in CDCl<sub>3</sub> (600 MHz).

The <sup>13</sup>C NMR spectrum of **61** displayed a total of 24 carbon signals, including five methyls, four methylenes, seven methines (five aromatic/olefinic carbons, one oxygenated carbon and one aliphatic carbons), eight quaternary carbons (one ketone group, five aromatic/olefinic carbons, one oxygenated carbon and one aliphatic carbon (Table 2.2.4.2). Upon comparison of the <sup>13</sup>C NMR data of **61** with that of eupenifeldin **23**, we found that **61** lacked one methylene carbon, one aliphatic methine carbon, and one oxygenated quaternary carbon of a pyran ring; this supported the proposal

that **61** had only one tropolone/pyran moiety. However, compound **61** had an additional olefinic proton indicated that one more double bond was located at the humulene ring. The additional double bond was likely to be located at C-7 and C-8 since humulene has a double bond at the same position. This was confirmed by the  $^1\text{H}$ - $^1\text{H}$  COSY correlation of H-6/H-7 and HMBC correlation from H-7 to C-6, C-9, and C-23 (Figure 2.2.4.4). Collectively, the planar structure of **61** was definitely assigned.



**Figure 2.2.4.3**  $^1\text{H}$  NMR spectrum of neosetophomone B **61**; The most identifiable signal was highlighted in color.

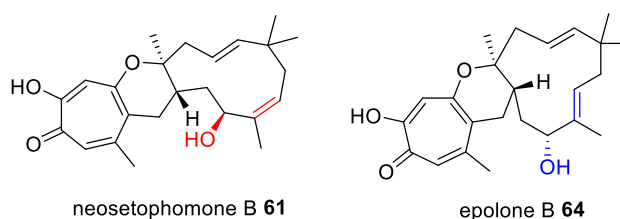


**Figure 2.2.4.4** Key COSY and HMBC correlations of neosetophomone B **61**.

Position	$\delta_c$ /ppm <b>61</b>	$\delta_H$ /ppm <b>61</b>	$\delta_c$ /ppm <b>61</b> literature	$\delta_H$ /ppm <b>61</b> literature
<b>1</b>	81.2		81.3	
<b>2</b>	43.3	2.75, m	44.5	2.70, d, $J = 15$ Hz
<b>3</b>	119.8	2.36, dd, $J = 14.9, 10.1$ Hz 5.12, ddd, $J = 15.8, 10.0, 2.9$ Hz	120.0	2.30, dd, $J = 14.9, 10.1$ Hz 5.07, ddd, $J = 15.6, 10.2, 2.4$ Hz
<b>4</b>	144.6	5.33, dd, $J = 15.8, 1.9$ Hz	144.8	5.27, d, $J = 16.2$ Hz
<b>5</b>	38.5		38.7	
<b>6</b>	40.6	2.14, dd, $J = 13.1, 9.8$ Hz 1.84, m	40.8	2.10, dd, $J = 12.0, 10.2$ Hz 1.78 overlap
<b>7</b>	127.0	5.40, ddd, $J = 9.5, 7.6, 1.6$ Hz	127.2	5.35, t, $J = 8.4$ Hz
<b>8</b>	137.9		138.1	
<b>9</b>	67.0	4.80, dd, $J = 10.1, 6.2$ Hz	67.2	4.75, dd, $J = 9.6, 6.0$ Hz
<b>10</b>	36.6	1.95, m 1.45, m	36.8	1.89, m 1.40, dd, $J = 13.8, 10.8$ Hz
<b>11</b>	33.0	1.83, m	33.2	1.77, overlap
<b>12</b>	34.4	2.70, m 2.49, dd, $J = 17.5, 12.6$ Hz	34.6	2.44, dd, $J = 16.8, 13.2$ Hz 2.65, dd, $J = 17.4, 5.4$ Hz
<b>13</b>	121.1		121.3	
<b>14</b>	160.9		162.3	
<b>15</b>	113.3	7.00, s	113.7	6.96, s
<b>16</b>	163.3		163.8	
<b>17</b>	172.5		172.8	
<b>18</b>	124.5	7.13, s	124.6	7.09, s
<b>19</b>	149.6		149.8	
<b>20</b>	19.4	1.19, s	19.6	1.14, s
<b>21</b>	21.6	0.95, s	21.8	0.91, s
<b>22</b>	30.1	1.09, s	30.3	1.04, s
<b>23</b>	18.1	1.70, s	18.4	1.66, s
<b>24</b>	27.1	2.40, s	27.2	2.35, s

**Table 2.2.4.2** NMR data for neosetophomone B **61** in CDCl<sub>3</sub> (500 MHz); Literature data was measured in CDCl<sub>3</sub> (600 MHz).

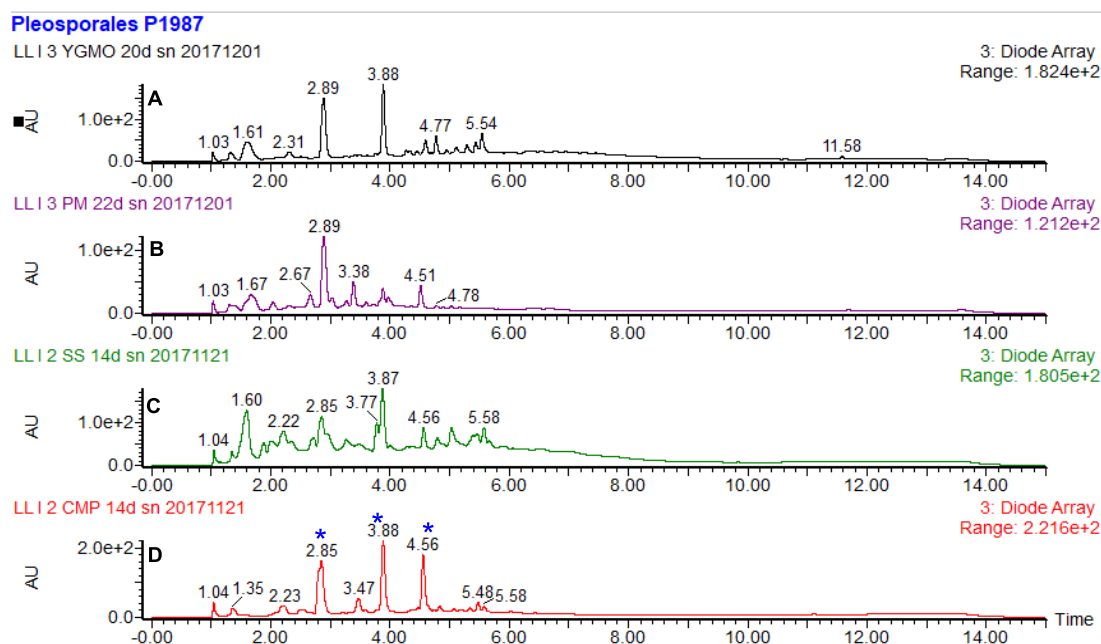
The relative stereochemistry at positions C-1, C-11 and C-9 were assumed to be identical to eupenifeldin **23** due to similarities of NMR data, but not further determined. However, the same NMR data of the compound was reported in 2019 by Oberlies group was named neosetophomone B **61**. The absolute configuration of neosetophomone B **61** was determined by modified Mosher's ester method and confirmed *via* single-crystal X-ray diffraction in the literature. So, neosetophomone B **61** was proved to be a diastereomer of epolone B **44**,<sup>33</sup> the differences are indicated in the figure 2.3.4.5.



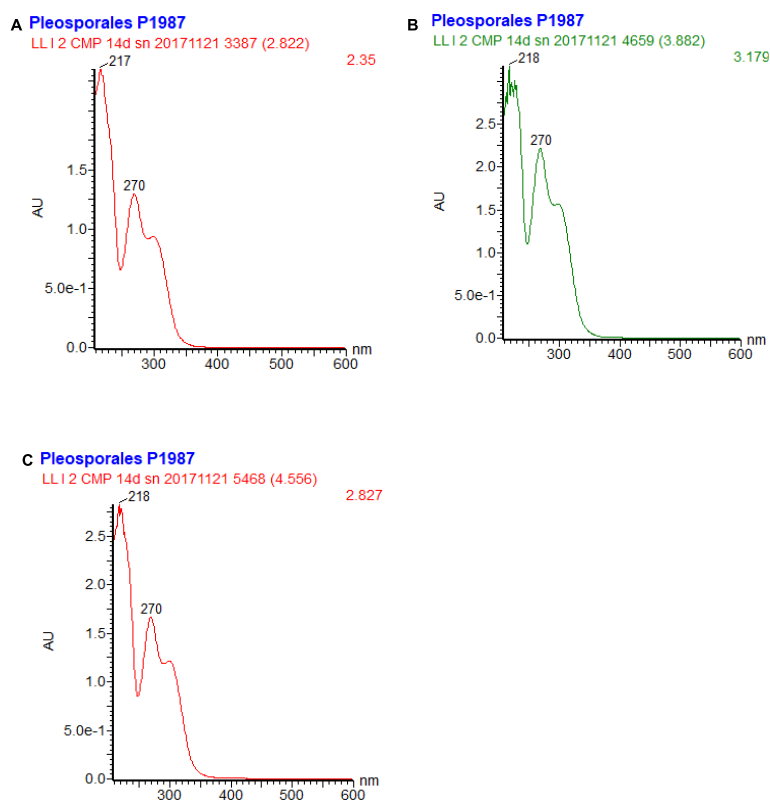
**Figure 2.2.4.5** Structural differences between neosetophomone B **61** and epolone B **44**.

### 2.2.5 Analysis of Secondary Metabolites from *Pleosporales sp 1987*

*Pleosporales sp 1987* was provided by Dr. Jose G. Maciá-Vicente, Institute of Ecology, Evolution and Diversity, Goethe University Frankfurt. Different liquid media including CMP, SS, PM and YGMO were used to test the producing conditions of tropolone sesquiterpenoids in *Pleosporales sp 1987*. All of the cultures were incubated at 28 °C and 110 rpm for 14 days. Then the mycelia were removed by filtration and the supernatant was extracted with ethyl acetate two times. Organic extracts were dried and dissolved in methanol to a concentration of 10 mg/mL for HPLC analysis. The HPLC chromatograms showed three main peaks eluting at 2.9 min, 3.9 and 4.6 min (Figure 2.2.5.1). Those compounds all showed UV absorption maxima  $\lambda_{\text{max}}$  218 nm and 270 nm (Figure 2.2.5.2) instead of characteristic UV absorption of tropolone ( $\lambda_{\text{max}}$  254 and 361 nm, Figure 2.2.3.2), which suggested that they were not TS. To find pycnidione **22** or eupenifeldin **23**, extracted ion chromatograms (EIC) of cultures were analyzed at 549 Da (ES<sup>+</sup> mode) corresponding to the expected [M + H]<sup>+</sup> molecular ion. However, no peaks were observed, this indicated that *Pleosporales sp 1987* was not a tropolone sesquiterpenoids producing strain under the tested conditions. Lack of presence of tropolone sesquiterpenoids or related compounds meant that further studies of this organism were not followed.



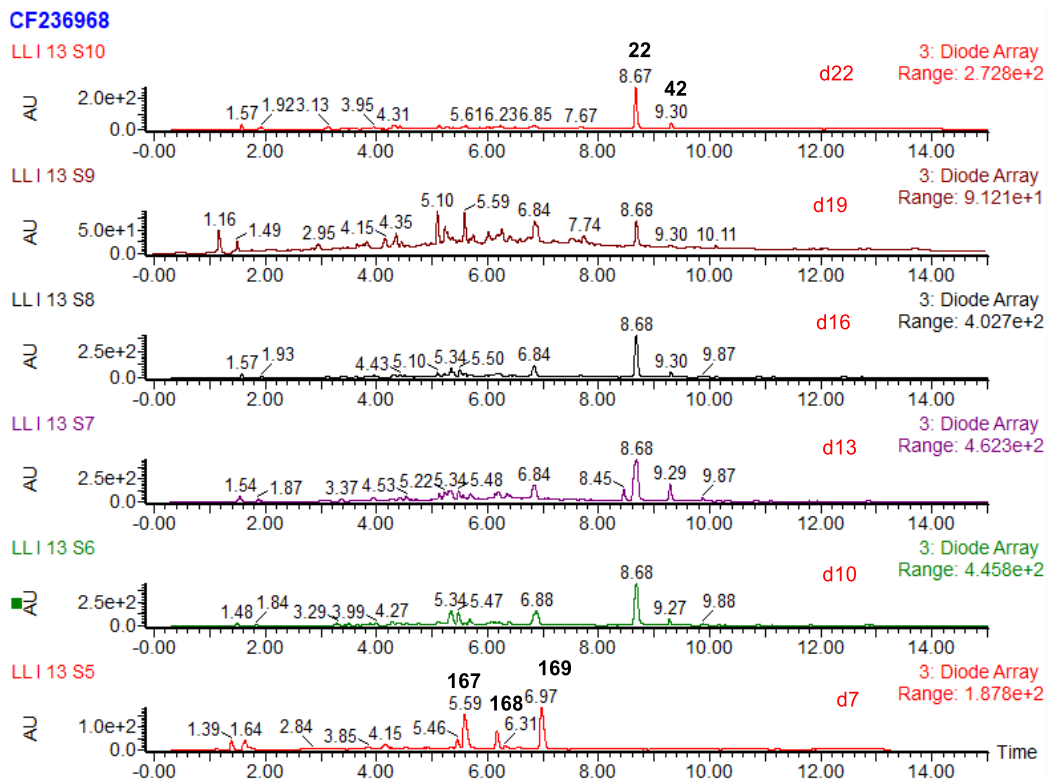
**Figure 2.2.5.1** The LCMS chromatogram of *Pleosporales sp.* 1987 culture extracts in different media: **A**, YGMO medium; **B**, PM medium; **C**, SS medium; **D**, CMP medium. TS unrelated compounds are highlighted with \*.



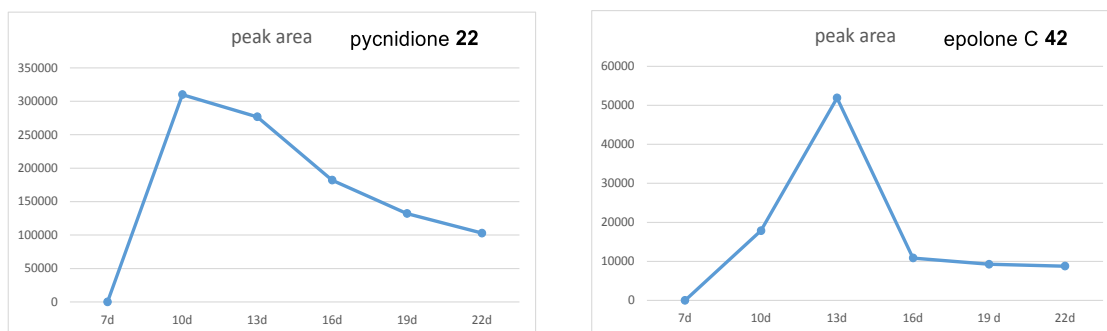
**Figure 2.2.5.2** UV spectrum of three main peaks eluting at 2.9 min, 3.9 and 4.6 min in the LCMS chromatogram of *Pleosporales sp.* 1987.

### 2.2.6 Timecourse of Metabolite Production from *CF236968*

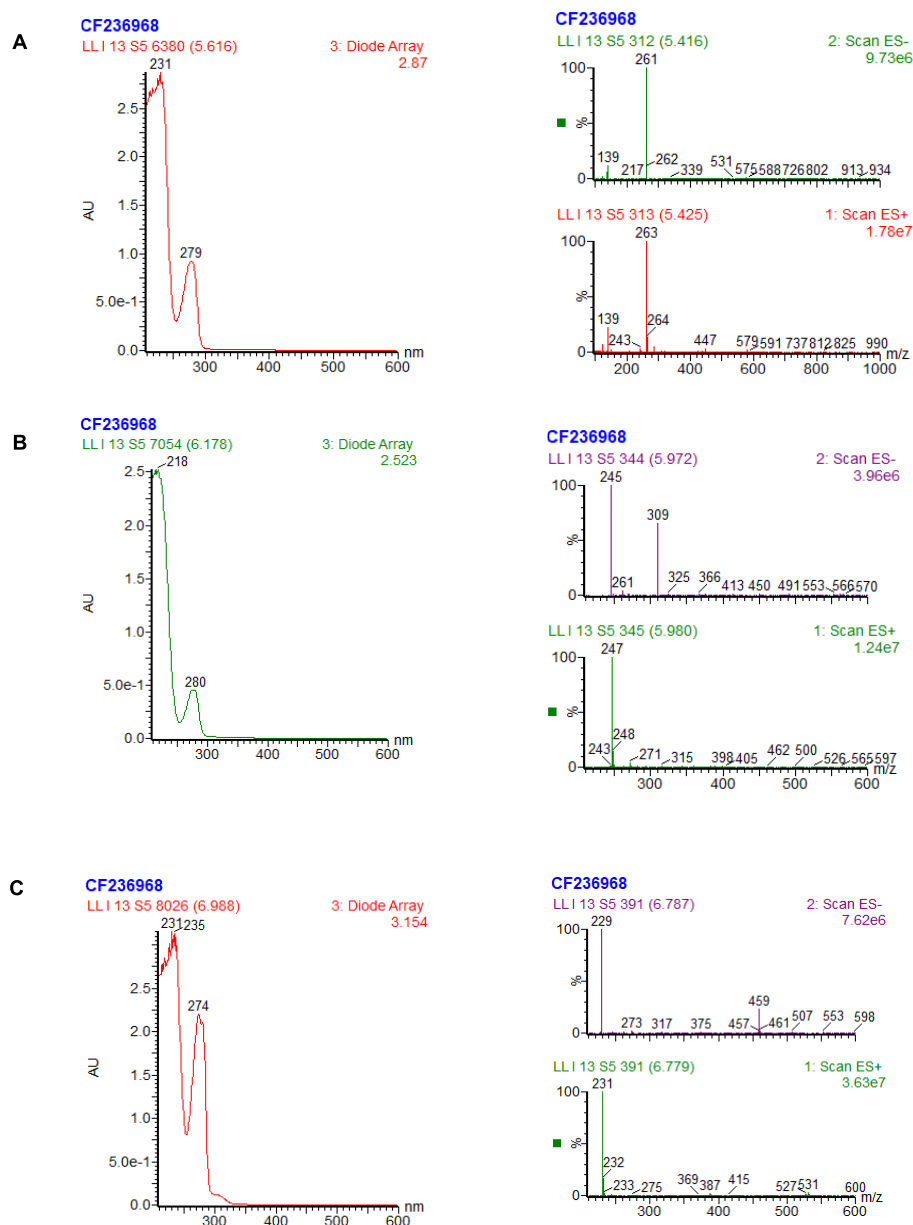
For the purpose of tropolone meroterpenoid isolation, time-course production experiments were performed to determine the time of maximum production of tropolone sesquiterpenoids and monitor the change of metabolite profile. For such experiments, 10 PD agar plates (20 ml in one plate) were inoculated using a spore-suspension of *CF236968* and grown at 22 °C. Analysis of crude metabolites with LCMS started from day 7 until day 22, and two PD plates were extracted with 200 ml ethyl acetate every 3 days (Figure 2.2.6.1). The relative concentrations of pycnidione **22** and epolone C **42** were indicated using DAD peak areas, which was only used as an internal comparison of the accumulation of targeted compound (Figure 2.2.6.2). The preliminary investigation showed that *CF236968* gave the maximum production of pycnidione **22** at 10 d and the maximum production of epolone C **42** at 13 d. It is worth noting that 3 new peaks were observed at day 7, which were all decreasing intensely with the production of tropolone sesquiterpenoids. It was proposed that the decline in the level of these three peaks in the early days might be associated with the increase in the production of tropolone sesquiterpenoids. Those compounds all showed UV absorption  $\lambda_{max}$  280 nm and had ES<sup>+</sup> molecular ions of  $m/z$  263 ( $t_R = 5.6$  min),  $m/z$  247 ( $t_R = 6.2$  min) and  $m/z$  231 ( $t_R = 7.0$  min) respectively (Figure 2.2.6.3).



**Figure 2.2.6.1** LCMS chromatogram of a time-course secondary metabolites production from *CF236968*.



**Figure 2.2.6.2** Production curve of pycnidione 22 and epolone C 42 production in *CF236968*.

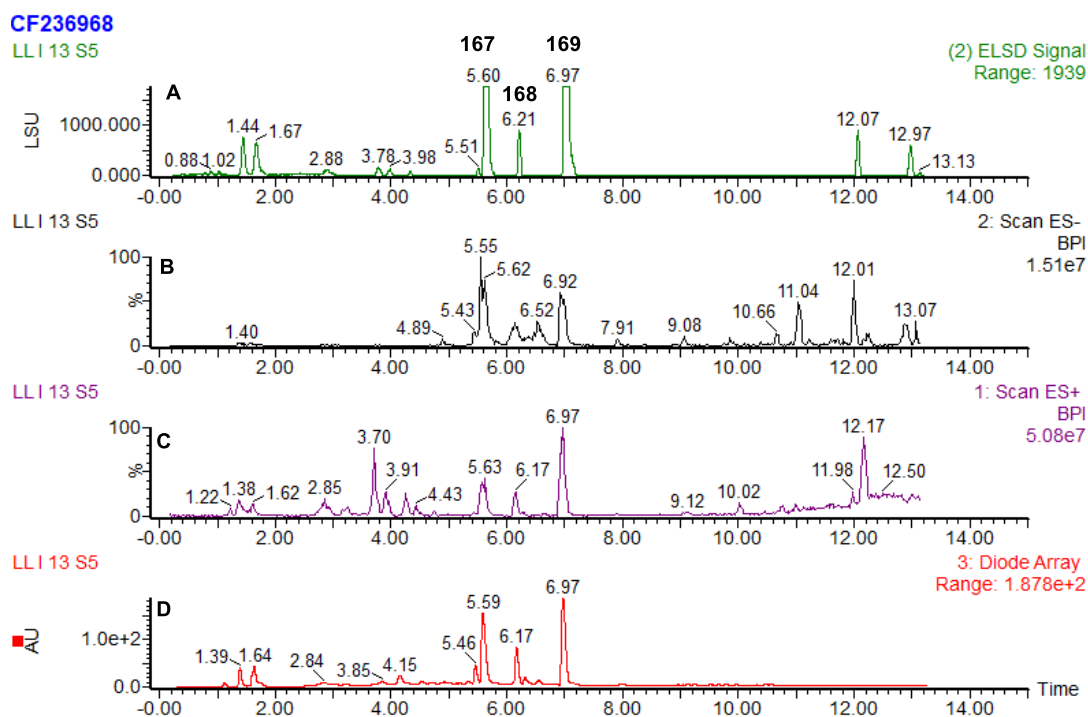


**Figure 2.2.6.3** UV and MS spectrum of three main peaks appeared in *CF236968* on day 7: **A**, the peak eluting at 5.6 min; **B**, the peak eluting at 6.2 min; **C**, the peak eluting at 6.7 min.

## 2.2.7 Structural Identification of violaceol I 167, cordyol C 168 and diorcinol 169

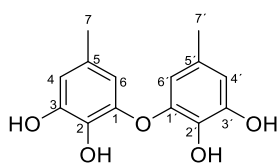
To confirm the assignment of the compounds eluting at 5.6 min, 6.2 min and 7.0 min, the three peaks were targeted for preparative LCMS purification and NMR analysis (Figure 2.2.7). *CF236968* were grown on 20 PD agar plates for 7 d at 22 °C. The mycelia together with agar were cut into small pieces and then extracted with 2 L ethyl acetate by soaking and stirring for 4 hours. The organic extract was filtered and evaporated to afford 1g crude metabolites. In final, 4 mg compound eluting at 5.6 min, 2.2 mg compound eluting at 6.2 min and 4.8 mg compound eluting at 7.0 min were then purified by mass-directed preparative HPLC.



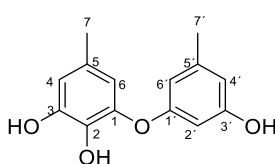


**Figure 2.2.7** The LCMS chromatogram of *Leptobacillum sp.* CF-236968 culture extracts after 7 days' growth on PD agar: **A**, ELSD chromatogram of the culture extracts; **B**, ES<sup>-</sup> TIC chromatogram of the culture extracts; **C**, ES<sup>+</sup> TIC chromatogram of the culture extracts; **D**, UV DAD chromatogram of the culture extracts.

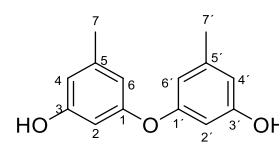
The <sup>1</sup>H- and <sup>13</sup>C-NMR spectra of the compound eluting at 5.6 min, together with its structural formula C<sub>14</sub>H<sub>14</sub>O<sub>5</sub>, suggested a symmetrical diphenyl ether structure for **167**. Compound **167** was identified as violaceol I **167**, which was a known metabolite that has been isolated from several different fungi (Table 2.2.7.1).<sup>145</sup> The compound eluting at 7.0 min has a molecular formula C<sub>14</sub>H<sub>14</sub>O<sub>3</sub> according to the MS spectrum. The <sup>1</sup>H- and <sup>13</sup>C-NMR spectral features suggested another symmetrical diphenyl ether, which was identified as diorcinol **169** (Table 2.2.7.2).<sup>146</sup> The <sup>1</sup>H-, <sup>13</sup>C- and 2D-NMR data (including COSY, HSQC, and HMBC) indicated that compound eluting at 6.2 min consisted of 3,4,5-trihydroxytoluene unit connected to 3,5-dihydroxytoluene residue, a known fungi metabolites named cordyol C **168** (Table 2.2.7.3).<sup>147</sup> Based on the structure analysis, these compounds are not the precursors of TS compounds as we expected.



**167**,  $m/z = 263$



**168**,  $m/z = 247$



**169**,  $m/z = 231$

**Figure 2.2.7** The structures of compounds **167**, **168** and **169**.

Position	$\delta_C$ /ppm <b>167</b>	$\delta_H$ /ppm <b>167</b>	$\delta_C$ /ppm <b>167</b> literature <sup>145</sup>	$\delta_H$ /ppm <b>167</b> literature <sup>145</sup>
1/1'	146.5		146.5	
2/2'	134.9		135.0	
3/3'	147.5		147.6	
4/4'	112.4	6.39 dd ( $J = 2.0, 0.8$ Hz)	112.6	6.39 m
5/5'	130.0		130.1	
6/6'	111.6	6.14 dd ( $J = 2.0, 0.8$ Hz)	111.8	6.15 m
7/7'	21.0	2.11 s	21.2	2.11 s

**Table 2.2.7.1** NMR data for violaceol I **167** in CD<sub>3</sub>OD (500 MHz). Literature data was also measured in CD<sub>3</sub>OD (600 MHz).

Position	$\delta_C$ /ppm <b>169</b>	$\delta_H$ /ppm <b>169</b>	$\delta_C$ /ppm <b>169</b> literature <sup>146</sup>	$\delta_H$ /ppm <b>169</b> literature <sup>146</sup>
1/1'	159.6		157.6	
2/2'	104.2	6.19, m	102.9	6.45, m
3/3'	159.5		158.4	
4/4'	112.0	6.35, m	111.1	6.33, m
5/5'	141.6		140.0	
6/6'	111.8	6.26, m	110.0	6.24, m
7/7'	21.5	2.21, s	21.1	2.18, s

**Table 2.2.7.2** NMR data for diorcinol **169** in CD<sub>3</sub>OD (500 MHz). Literature data was also measured in DMSO-*d*<sub>6</sub> (600 MHz).

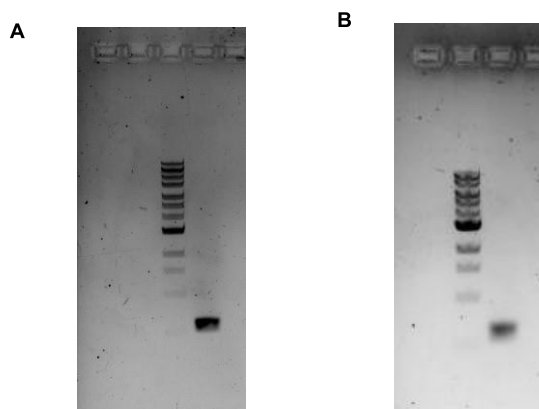
Position	$\delta_C$ /ppm <b>168</b>	$\delta_H$ /ppm <b>168</b>	$\delta_C$ /ppm <b>168</b> literature <sup>147</sup>	$\delta_H$ /ppm <b>168</b> literature <sup>147</sup>
1	145.3		143.4	
2	136.0		134.9	
3	147.7		146.5	
4	113.0	6.44 dd ( $J = 2.0, 0.7$ )	112.2	6.53 d ( $J = 1.6$ Hz)
5	130.0		128.7	
6	113.5	6.20 dd ( $J = 2.0, 0.7$ )	112.6	6.28 d ( $J = 1.8$ Hz)
7	20.9	2.14 d ( $J = 0.7$ )	20.0	2.16 s
1'	160.5		159.3	
2'	102.7	6.16 m	101.3	6.15 dd $J = 2.2, 2.1$ Hz
3'	159.4		158.4	
4'	111.2	6.29 m	110.1	6.35 br s
5'	141.3		140.0	
6'	110.3	6.24 m	108.8	6.22 br s
7'	21.6	2.20 s	20.6	2.19 s

**Table 2.2.7.3** NMR data for cordyol C **168** in CD<sub>3</sub>OD (500 MHz). Literature data was also measured in DMSO-*d*<sub>6</sub> (600 MHz).

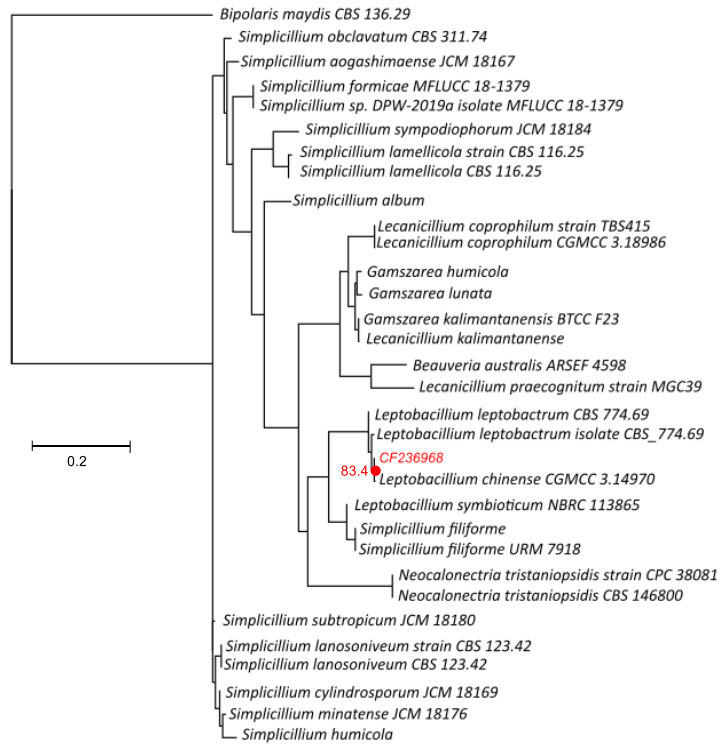
### 2.3 Results-ITS Characterization of *CF 236968* and *CF253087*

The internal transcribed spacer (ITS) is a genomic region between 18S and 28S ribosomal RNA (rRNA) including 5.8S rRNA. ITS is widely used in the classification of fungal species because it has highly conserved sequences in the same species but show a high degree of variation even between closely related species.<sup>148-149</sup> So, ITS PCR was performed in this part to determine the species relationships of *CF236968* and *CF253087* (Figure 2.3.1).

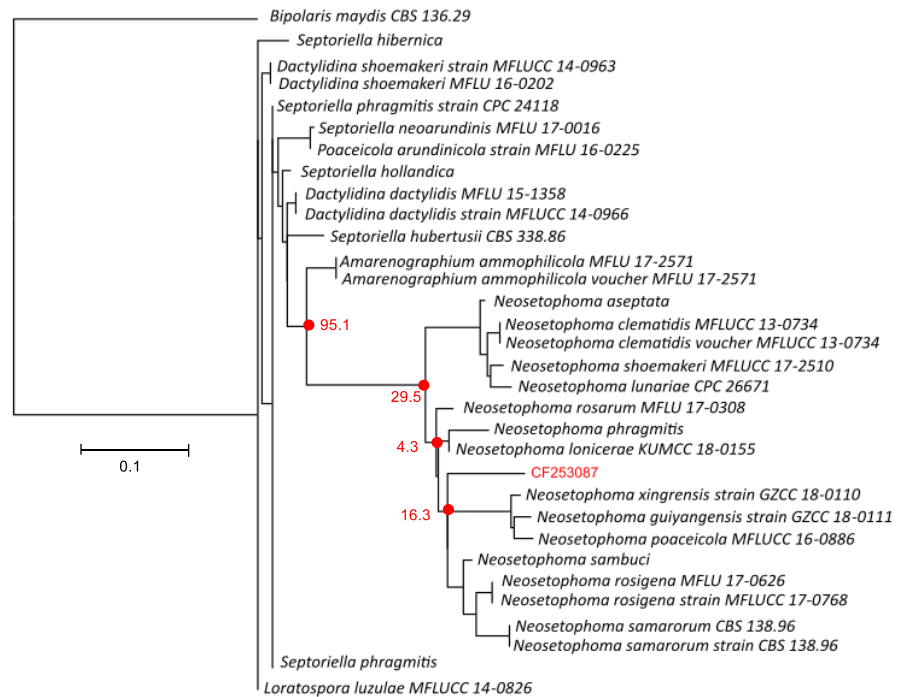
ITS sequences were amplified from gDNA of *CF236968* and *CF253087* by PCR using standard ITS1 and ITS4 primers.<sup>148</sup> The resulting PCR products were sequenced, and the results were then blasted into NCBI database using nBLAST (nucleotide Basic Local Alignment Search Tool). The top 30 known ITS sequences from different fungi together with the ITS sequences of *CF236968* were used for constructing a phylogenetic tree in PhyML 3.0.<sup>150</sup> The ITS sequence from the model organism *Bipolaris maydis* CBS 136.29 was used as an outgroup. This places *CF236968* in the center of the *Leptobacillium* clade, confirming its identity as *Leptobacillium* sp. *CF236968* (Figure 2.3.2). The phylogenetic tree of *CF253087* was constructed in the same way. Due to low support value between *CF253087* and *Neosetophoma* species, *CF253087* was then placed in the *Phaeosphaeriaceae* sp. family (Figure 2.3.3).



**Figure 2.3.1** 1 % agarose gel electrophoresis analysis of ITS PCR product from two fungi: **A**, *CF236968*; **B**, *CF253087*.



**Figure 2.3.2** A phylogenetic tree of *CF236968* constructed in PhyML 3.0 using ITS sequences.



**Figure 2.3.3** A phylogenetic tree of *CF253087* constructed in PhyML 3.0 using ITS sequences.

## 2.4 Results-*In silico* Analysis of Putative BGC for Tropolone Sesquiterpenoids

### 2.4.1 Whole Genome Sequencing of CF236968

The fresh and dry fungal mycelia of *CF236968* was prepared from 7 days old PDB liquid culture, and extracted by GenElute Plant Genomic DNA miniprep kit. Then, sodium acetate (40  $\mu$ l 3 M) and ethanol (1 ml) was added to precipitate the DNA which was subsequently washed 3 times with 70 % ethanol.

Assembled base (bp)	28,623,532
# of scaffolds	170
N50 scaffold (bp)	466,766
GC (%)	53.53

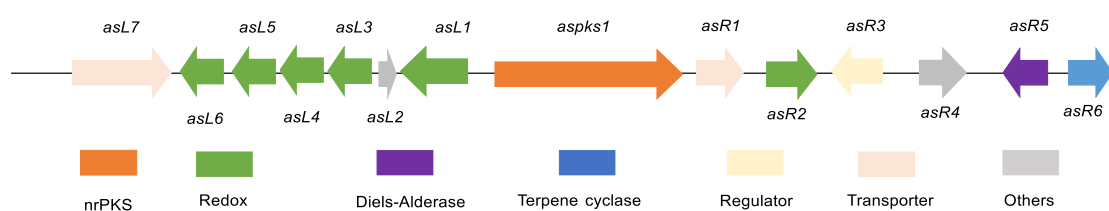
**Table 2.4.1** Key statistics of draft genome assembly in *CF236968*.

The 5  $\mu$ g sample was dissolved in water and sent to the Center for Biotechnology (CeBiTec, University of Bielefeld) for Illumina MiSeq paired-end sequencing. The raw data were processed by Dr Daniel Wibberg using an in-house software platform based on CASAVA 1.8.2 (Illumina).<sup>151</sup> A draft genome of 29 Mb consisting of 170 scaffolds was obtained with an N50 scaffold length of 0.5 Mb. Gene prediction and annotation was performed with AUGUSTUS 3.0.3 and the GenDBE platform.<sup>152</sup> A total number of 10507 genes were predicted (Table 2.4.1).

### 2.4.2 *In silico* Analysis of Putative BGC for Pycnidione

In the previous work of Dr. Raissa Schor, a BGC of xenovulene A **24** (*aspks* BGC) was characterized from *A. strictum*. Heterologous expression of the BGC in *A. oryzae* indicated that at least 8 genes are necessary for the biosynthesis of xenovulene A **24**, including *aspks1*, *asL1* encoding an FAD-dependent hydroxylase, *asL3* encoding a non-haem iron dioxygenase, *asR2* encoding a cytochrome P450 monooxygenase, *asR5* encoding an hDAase, *asR6* encoding a terpene cyclase and *asL6* and *asL4* encoding putative NAD/FAD-dependent oxidoreductases (Table 2.4.2.1).<sup>20</sup> The biosynthetic pathway of xenovulene A **24** is proposed according to the heterologous expression results of the *aspks* BGC in *A. oryzae* (Scheme 1.2.5.2).

Based on the structural analysis, pycnidione **22** might be derived biogenetically *via* hDA reaction between one humulenol and two tropolones. Early steps in the biosynthesis of pycnidione **22** were predicted to involve oxidative formation of the tropolone nucleus, similar with that in the biosynthesis of xenovulene A **24**. Thus, gene clusters controlling tropolone biosynthesis in both *A. strictum* and *CF236968* were expected to show considerable sequence homology. To find the possible biosynthetic gene cluster of pycnidione **22**, the translated amino acid (AA) sequences of three necessary tropolone biosynthetic genes (*aspks1*, *asL1* and *asL3*) from *A. strictum* were used as templates in a BLAST search of the similar proteins in *CF236968* genome.

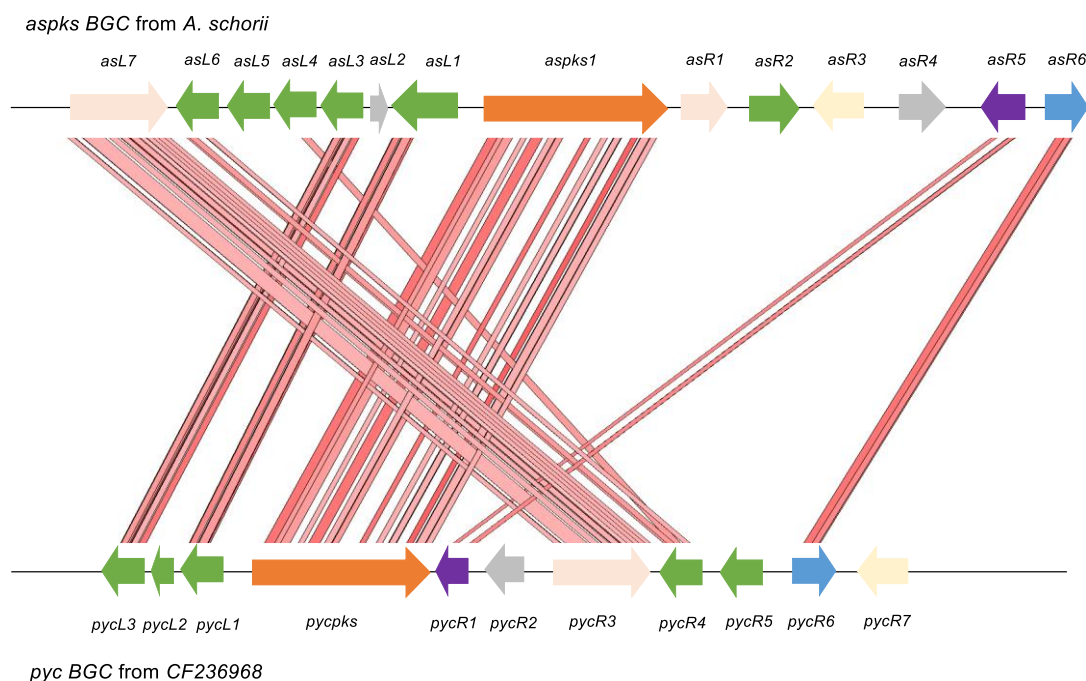


Name and Locus_tag	Annotation	Name and Locus_tag	Annotation
<i>asPKS1</i> <i>Asg3673</i>	nrPKS	<i>asL7</i> <i>Asg3680</i>	ABC transporter
<i>asL1</i> <i>Asg3674</i>	FAD-dependent salicylate hydroxylase	<i>asR1</i> <i>Asg3672</i>	MFS transporter
<i>asL2</i> <i>Asg3675</i>	Ferredoxin like	<i>asR2</i> <i>Asg3671</i>	Cytochrome P450
<i>asL3</i> <i>Asg3676</i>	Non-haem iron II oxygenase	<i>asR3</i> <i>Asg3670</i>	Regulator
<i>asL4</i> <i>Asg3677</i>	FAD-dependent oxidoreductase	<i>asR4</i> <i>Asg3669</i>	FAD-binding monooxygenase
<i>asL5</i> <i>Asg3678</i>	Short-chain dehydrogenase	<i>asR5</i> <i>Asg3668</i>	Hetero-Diels Alderase
<i>asL6</i> <i>Asg3679</i>	FAD-dependent oxidoreductase	<i>asR6</i> <i>Asg3667</i>	Terpene cyclase

**Table 2.4.2.1** Annotations of *aspks* BGC in *A. strictum*.

Three clustered genes in *CF236968* were found with strong sequence identity (> 49 %) to putative homologues in *A. strictum*. From this analysis, a promising BGC including 12 open reading frames was assumed to be responsible for pycnidione production, which was then named the *pyc* BGC. Genes from the *pyc* BGC were assigned generic names according to their position relative to the central PKS in its coding orientation; thus, the gene immediately left of the PKS was dubbed L1, and

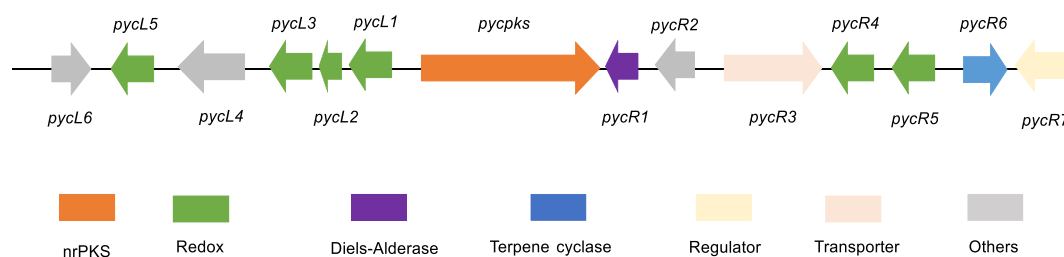
immediately to the right became R1. This BGC was further compared to the whole *aspks* BGC using the Artemis Comparison Tool.<sup>153</sup> The results indicated that 6 genes from the *pyc* BGC showed strong sequence identity (> 49 %) to putative homologues in *aspks* BGC (Figure 2.4.2).



**Figure 2.4.2** Artemis analysis of the *aspks* BGC and the *pyc* BGC.

Putative annotations and functions of all the genes from the *pyc* BGC were predicted by running protein Blast in NCBI and compared to the *aspks* BGC (Table 2.4.2.2). In the *pyc* BGC, nr PKS encoded by *pycPKS* is assigned as 3-methylorcinolaldehyde synthase, which catalyzes the production of 3-methylorcinolaldehyde **19**. The FAD-binding hydroxylase encoded by *pycL1* and non-haem iron II oxygenase encoded by *pycL3* are proposed to convert 3-methylorcinolaldehyde **19** to tropolone precursor **102**. *pycR6* is proposed to encode a terpene cyclase, which is responsible for the production of 2*E*-humulene **71**, while the enzyme encoded by *pycR1* is predicted to be an hDAase that catalyzes the required hDA reaction between tropolone and the terpene precursor. *PycR4* encoding a FAD-binding monooxygenase showed low sequence identity with both *aspL4* (38.6 %) and *aspL6* (35.6 %), might catalyze a ring-contraction reaction. However, *pycR5* encoding a cytochrome P450 oxygenase, showed low sequence identity with *aspR2* (31.8 %) and may be responsible for C-10 hydroxylation of the terpene moiety rather than oxygenation of the polyketide that is not required during pycnidione biosynthesis. In addition, the annotated ORFs also includes a

short-chain dehydrogenase (*pycL2*), an aldehyde dehydrogenase (*pycL5*), an ABC transporter (*pycR3*), a putative regulatory protein (*pycR2*) and an alkaline phosphatase D-related protein (*pycL4*) that have no homologs in the xenovulene A BGC.



Name and Locus_tag	Annotation	Name and Locus_tag	Annotation
<i>pycPKS</i> 4791_g	nrPKS	<i>pycR1</i> 4792_g	Hetero-Diels Alderase
<i>pycL1</i> 4790_g	FAD-binding hydroxylase	<i>pycR2</i> 4793_g	Regulator
<i>pycL2</i> 4789_g	Short-chain dehydrogenase	<i>pycR3</i> 4794_g	ABC transporter
<i>pycL3</i> 4788_g	Non-haem iron II oxygenase	<i>pycR4</i> 4795_g	FAD-binding monooxygenase
<i>pycL4</i> 4787_g	Alkaline phosphatase D-related protein	<i>pycR5</i> 4796_g	Cytochrome P450 monooxygenase
<i>pycL5</i> 4786_g	Aldehyde dehydrogenase	<i>pycR6</i> 4797_g	Terpene cyclase
<i>pycL6</i> 4785_g	Unknown	<i>pycR7</i> 4798_g	Regulator

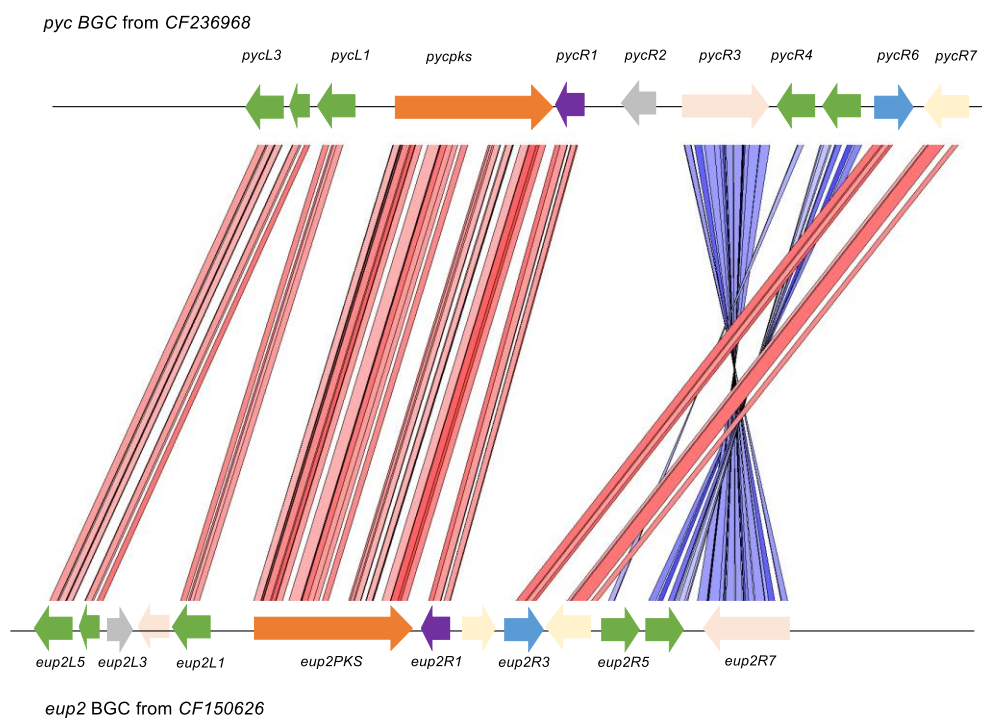
**Table 2.4.2.2** Annotations of the *pyc* BGC in *CF236968*

### 2.4.3 *In silico* Analysis of Putative BGC for Eupenifeldin

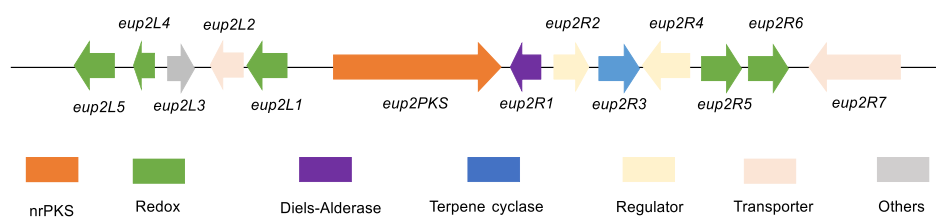
Eupenifeldin **23** and pycnidione **22** are diastereoisomers, therefore it is probable that their producing organisms should contain highly homologous gene clusters. The putative BGC of eupenifeldin **23** was predicted and named the *eup2* BGC via the bioinformatic analysis of the *CF150626* genome by Dr Carsten Schotte. The Artemis Comparison Tool analysis of the *pyc* BGC and the *eup2* BGC showed 10 homologous genes with similar predicted functions (Figure 2.4.3). All genes in the *eup2* cluster were analyzed by BLASTp (Table 2.4.3), and part of the gene functions were assigned as follows: Proteins encoded by *eup2PKS*, *eup2L1* and *eup2L5* are predicted to construct the tropolone



unit; A terpene cyclase (*eup2R3*) is proposed to produce terpene precursor; hetero DAase (*eup2R1*) will catalyze the hDA reaction between tropolone and terpene precursor. And cytochrome P450 (*eup2R6*) is responsible for C-10 hydroxylation of the terpene moiety.



**Figure 2.4.3** Artemis analysis of the *pyc* BGC and the *eup2* BGC.



Name and Locus_tag	Annotation	Name and Locus_tag	Annotation
<i>eup2PKS</i>	nr-PKS	<i>eup2R1</i>	hetero Diels-Alderase
<i>CF150626g2131</i>		<i>CF150626g2132</i>	
<i>eup2L1</i>	FAD-dependent monooxygenase	<i>eup2R2</i>	regulator
<i>CF150626g2130</i>		<i>CF150626g2133</i>	
<i>eup2L2</i>	Secondary transporter	<i>eup2R3</i>	Humulene-synthase
<i>CF150626g2129</i>		<i>CF150626g2134</i>	
<i>eup2L3</i>	Ferredoxin-like	<i>eup2R4</i>	regulator
<i>CF150626g2128</i>		<i>CF150626g2135</i>	
<i>eup2L4</i>	short-chain dehydrogenase (SDR)	<i>eup2R5</i>	FAD-dependent monooxygenase
<i>CF150626g2127</i>		<i>CF150626g2136</i>	
<i>eup2L5</i>	Non-haem iron II oxygenase	<i>eup2R6</i>	Cytochrome P450
<i>CF150626g2126</i>		<i>CF150626g2137</i>	monooxygenase (P450)
		<i>eup2R7</i>	ABC-transporter
		<i>CF150626g2138</i>	

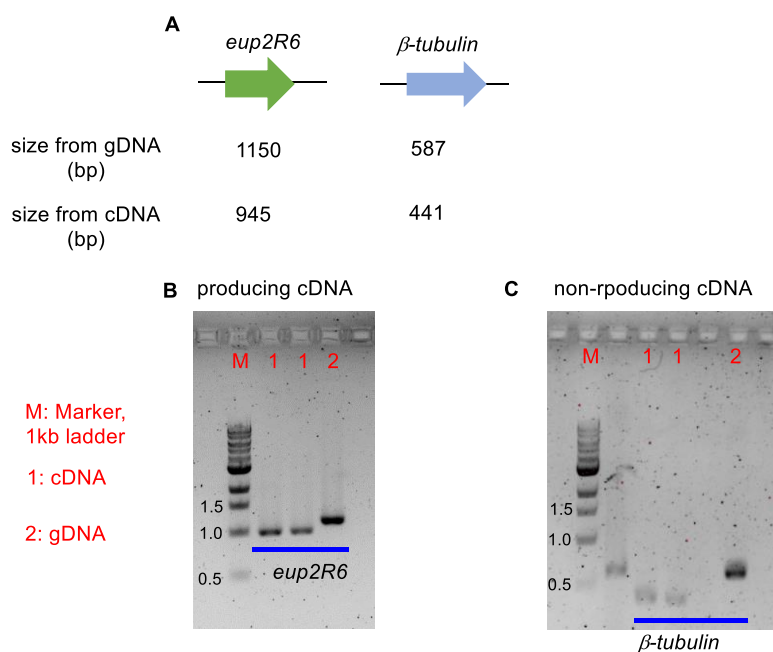
**Table 2.4.3** Annotations of the *eup2* BGC in *CF150626*

## 2.5 Results-RT-PCR Analysis for Eupenifeldin BGC in Fungus *CF150626*

In order to more clearly link the *eup2* BGC found by bioinformatic analysis with the biosynthesis of eupenifeldin **23**, the expression levels of these genes under producing and non-producing condition were investigated by reverse transcription polymerase chain reaction (RT-PCR).

*Phaeosphaeriaceae* sp. *CF150626* was obtained from the Fundación MEDINA (Granada, Spain). The producing condition of *CF150626* was first reported in 2008, production of eupenifeldin **23** together with other tropolone sesquiterpenoids were observed by growing fungi in PM medium.<sup>37</sup> In parallel, ½ TOM medium was found to be non-producing medium for comparative transcriptome analysis. Liquid cultures for both conditions (PM medium, ½ TOM medium) were inoculated with *CF150626* and cultured for 4 days. The messenger RNA (mRNA) of *CF150626* mycelium in PM medium was extracted using the trizol reagents and converted into the corresponding complementary

DNA (cDNA) with High Capacity RNA-to-cDNA Kit. The mRNA of *CF150626* in ½ TOM medium was extracted with ZR Fungal Bacterial RNA MiniPrep kit (Zymo Research) and immediately converted to cDNA in the same way.



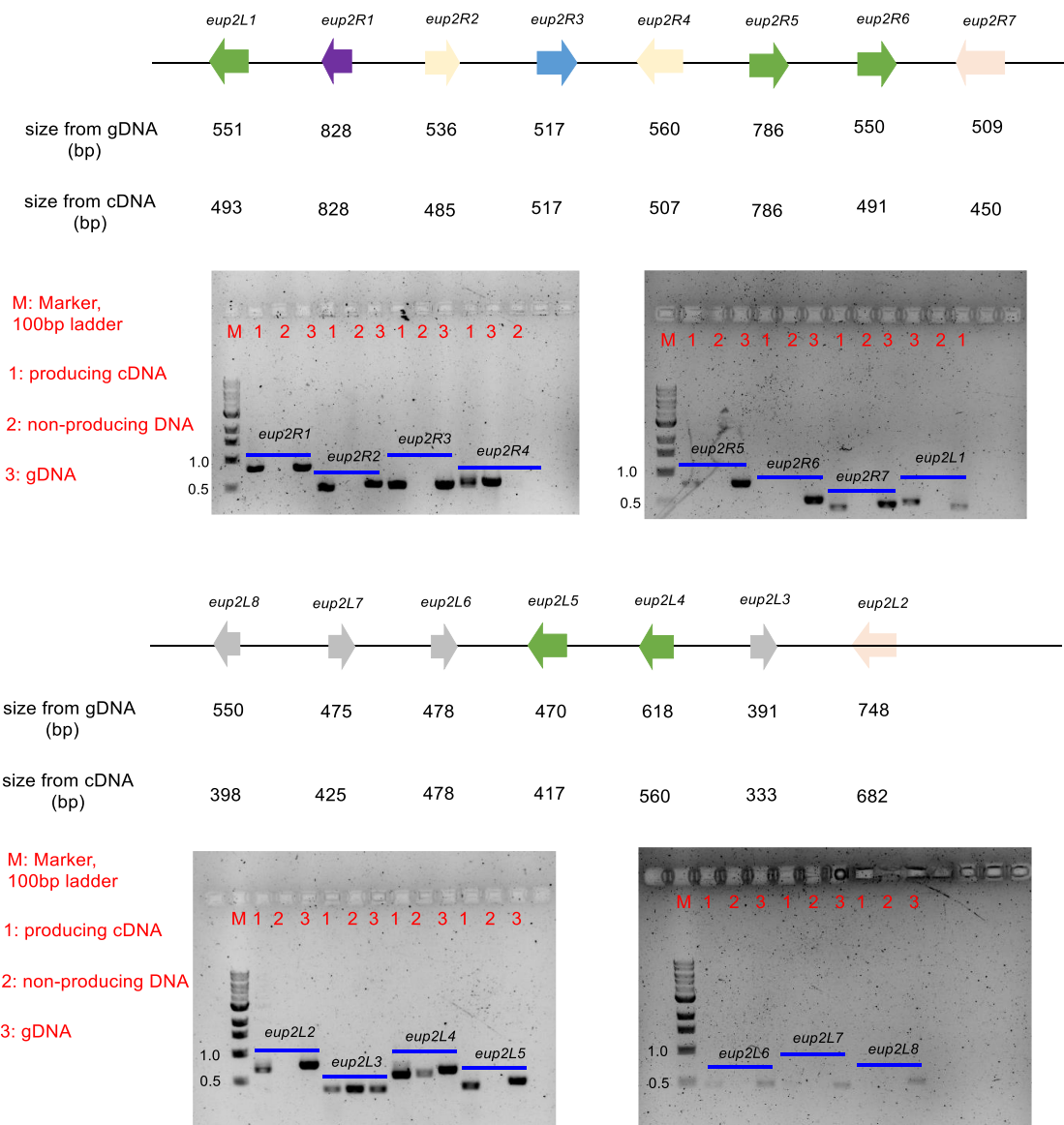
**Figure 2.5.1** PCR analysis of cDNA extracted from *CF150626* grown under producing and non-producing conditions.

We proposed that the genes from the *eup2* BGC would show strong expression in producing condition while the housekeeping gene (such as *β tubulin*) that are necessary for the survival of fungi would express in non-producing condition. Therefore, for the producing cDNA, PCR experiments were carried out with the oligonucleotides targeting *eup2R6* from the *eup2* BGC; For the non-producing cDNA, PCR experiments were performed the oligonucleotides to amplify *β tubulin*. gDNA of *CF150626* was also analyzed with the PCR experiment as control group. Since the targeted gene fragments contained introns, the PCR product from gDNA is longer than the product from cDNA. The shorter PCR product from *eup2R6* and *β tubulin* confirmed that both producing and non-producing cDNA were obtained successfully (Figure 2.5.1).

To gain more insight into the *eup2* BGC, RT-PCR experiments were started to check the expression of 8 consecutive genes on the left side of *eup2PKS* and 7 consecutive genes on the right side of

*eup2PKS* under both producing and non-producing conditions. Expression of *eup2R1-R5*, *eup2R7*, *eup2L1*, *L2*, *L5* were only observed under the producing condition and not under the non-producing condition (Figure 2.5.2). Except *eup2R2* and *eup2L2*, all the other genes have the corresponding homologous genes in the *pyc* BGC (chapter 2.5.3). Enzymes encoded by *eup2L1* and *eup2L5* are predicted to be involved in the formation of the tropolone unit; the terpene cyclase encoded by *eup2R3*, the FAD-dependent monooxygenase encoded by *eup2R5* and hetero DAase encoded by *eup2R1* are likely to be necessary for the biosynthesis of eupenifeldin **23**. However, it is hard to predict the chemical reaction catalyzed by *eup2R5*. These results validated that the *eup2* BGC is strongly correlated to the production of eupenifeldin **23**.

In addition, expression of *eup2L3* and *eup2L4* were observed under both producing and non-producing conditions, while the expression of *eup2R6*, *eup2L6- eup2L8* were not observed neither in producing condition nor in the non-producing condition. Among them, *eup2L4* (SDR) and *eup2R6* (P450) have homologous genes in the *pyc* BGC.

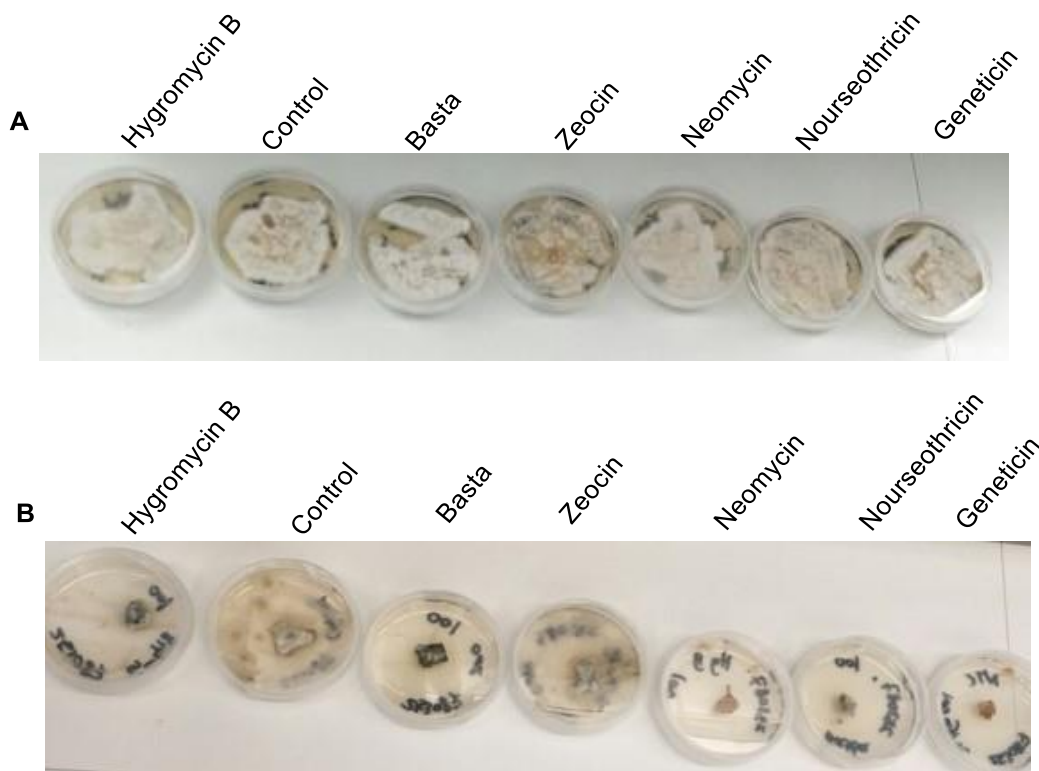


**Figure 2.5.2** Reverse Transcriptase (RT) analysis of genes from the *eup2* BGC.

## 2.6 Results-Antibiotic Sensitivity of Fungi *CF236968* and *CF253087*

In fungi, gene knockouts are normally achieved by inserting a suitable antibiotic resistance cassette into the target gene. Antibiotic testing is the first step of the process. To analyze whether *CF236968* and *CF253087* show sensitivity to antibiotics, they were sub-cultured on CD agar containing 100 µg/mL of six different antibiotics, including hygromycin B, basta, zeocin, neomycin, nourseothricin, and geneticin. Observation of the colonies at 22 °C for 14 days showed that *CF236968* is able to grow on all the plates containing antibiotics (Figure 2.6.A). Under the same cultivation conditions,

the growth of *CF253087* was effectively inhibited by hygromycin B, basta, neomycin, nourseothricin and geneticin after 14 days (Figure 2.6.B). These observations suggested that *CF253087* exhibited multiple antibiotic sensitivity and could be used for further gene manipulation.

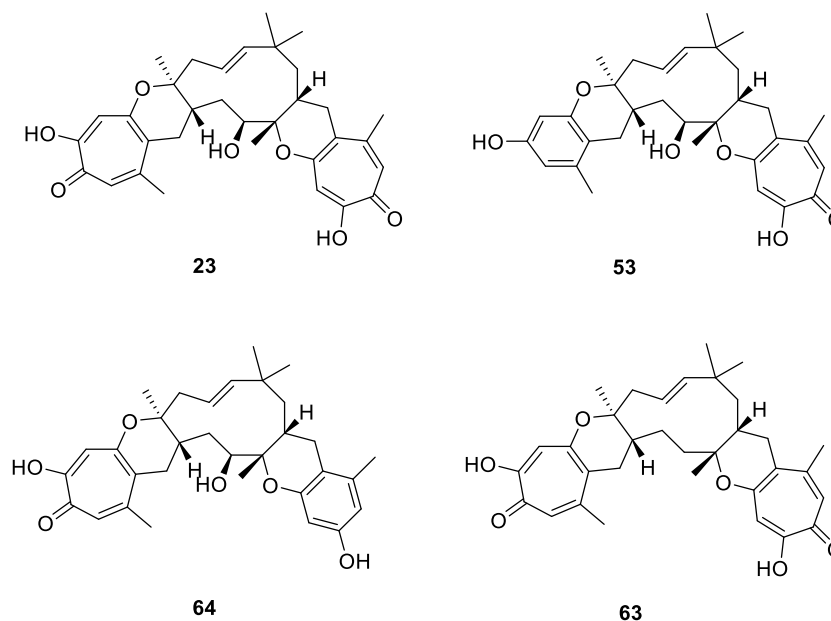


**Figure 2.6** A, *CF236968* on CD+S agar plates with different antibiotics at 16d; B, *CF253087* on CD+S agar plates with different antibiotics at 16d.

## 2.7 Results-Developing a Transformation Protocol for *CF150626*

Based on the bioinformatic analysis and RT-PCR findings, knock out experiments targeting the *eup2* BGC from *CF150626* should be carried out to clarify the biosynthetic pathway of eupenifeldin **23**. *CF150626* was reported to produce eupenifeldin **23** and other tropolone sesquiterpenoids (Figure 2.7.1),<sup>35,37</sup> but has not been transformed before. Transformation is a process to introduce and incorporate exogenous DNA into the cell which is very important in the exploration of secondary metabolites and biosynthetic pathways by targeted gene knockouts in fungi. Some fungi might not be transformable, so it is necessary to find a feasible transformation protocol for fungi *CF150626*

before the investigation of genes involved in the eupenifeldin biosynthetic pathway by molecular engineering. The antibiotic sensitivity of *CF150626* was tested by Dr Carsten Schotte. The fungus was obviously inhibited by 100 µg/mL hygromycin B. Thus, a pTH-GS-eGFP plasmid containing a *hph* gene and a *eGFP* gene was chosen as exogenous DNA for the transformation. This would help us assess the efficiency of the tested transformation protocols and if the  $P_{gpdA}$  and  $P_{amyB}$  promoters used for the *hph* and *eGFP* respectively works well in *CF150626*.

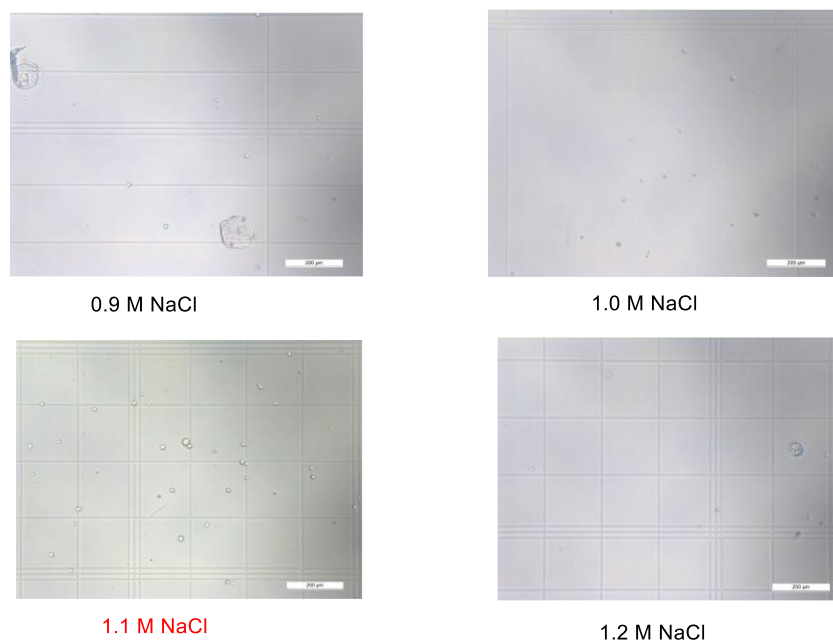


**Figure 2.7.1** Known tropolone sesquiterpenoids from *CF150626*.

First, A PEG-mediated transformation protocol based on the *A. oryzae* transformation protocol<sup>127</sup> was tested for the fungus *CF150626*. To transform fungal cells, the cell wall must be digested with hydrolytic enzymes, which will produce single cell without cell wall, namely a protoplast. The protoplasts are rather easy to transform with exogenous DNA and needed to be preserved in the saline buffer whose concentration ranges accordingly to the specific fungal species.

In order to obtain healthy and numerous protoplasts, different NaCl solutions were tested for maintaining protoplast homeostasis. The mycelia were collected from 2-day-old PDB culture (22 °C, still) by filtration using sterile miracloth as a filter. They are digested in different concentrations of 10 ml NaCl solutions (0.9, 1.0, 1.1 and 1.2 M ) with *Trichoderma* lysing enzymes (10 mg/ml) and

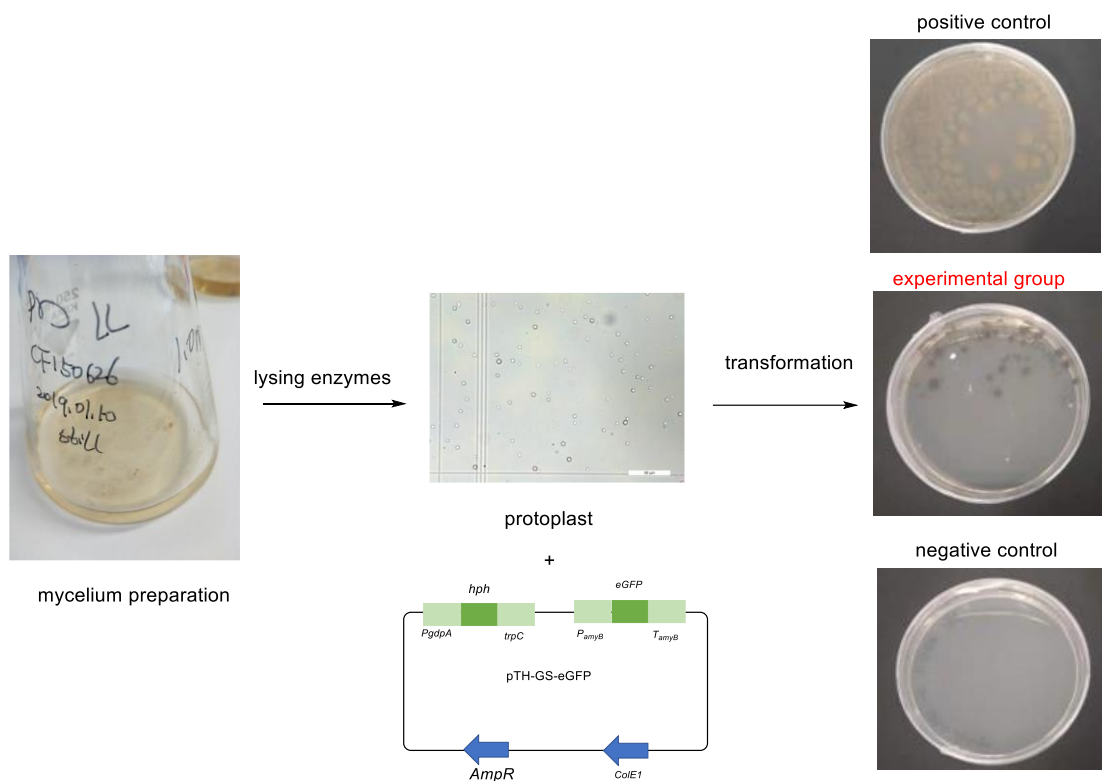
driselase from *Basidiomycetes* (5 mg/ml) and incubated at 30 °C, 110 rpm for 3h. After that, the NaCl solutions with protoplasts was filtered with sterile miracloth and centrifuged to get protoplast pellets, which were then suspended in 100 µl different concentrations of NaCl solutions. The concentration of protoplasts was checked under the microscope. The results indicated that a large number of protoplasts were only observed in 1.1 M NaCl solution, the best condition for *CF150626* protoplast preparations. (Figure 2.7.2)



**Figure 2.7.2** protoplast preparations in different NaCl solution.

The protoplasts were transformed with pTH-GS-eGFP<sup>154</sup> and cultivated in CD agar, which were covered with 10 mL of CD soft agar with 100 µg/mL hygromycin B after overnight incubation. Unfortunately, no transformant was obtained in this protocol after 8 transformation experiments in total, which indicated *CF150626* could not be transformed with *A. oryzae* transformation protocol. Then a new transformation protocol of *Phoma sp.* was tested on the protoplasts of *CF150626* with pTH-GS-eGFP plasmid as exogenous DNA.<sup>132</sup> After incubating the transformed protoplast at 22 °C for 10 days, some colonies were observed on the PD agar with 100 µg/mL hygromycin B (Scheme 2.7). These transformants were then selected for the second and third round of hygB selection to obtain stable transformants. However, no further analysis of these transformants with PCR or green fluorescence checking was carried out as the *eup* BGC was confirmed to be responsible for eupenifeldin biosynthesis by KO experiments in a *Phoma sp.* fungus by Hu and co-workers.<sup>54</sup>

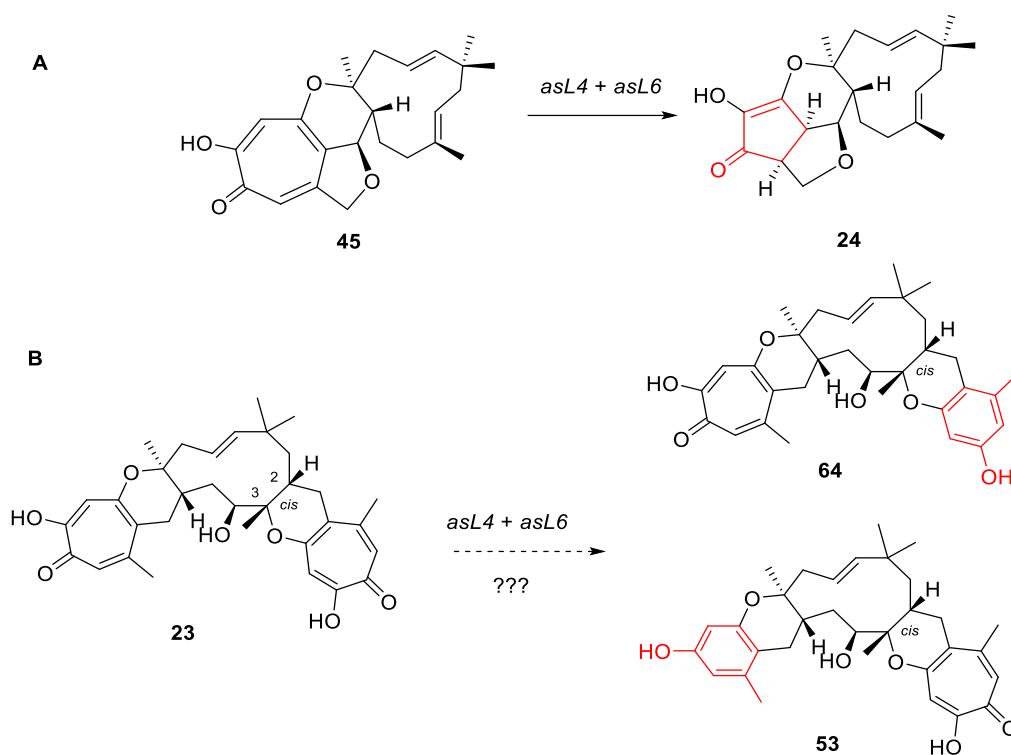




**Scheme 2.7** Transformation of *CF150626* with *Phoma sp.* transformation protocol.

## 2.8 Results-Heterologous Expression of *asL4* and *asL6* in *A. oryzae*

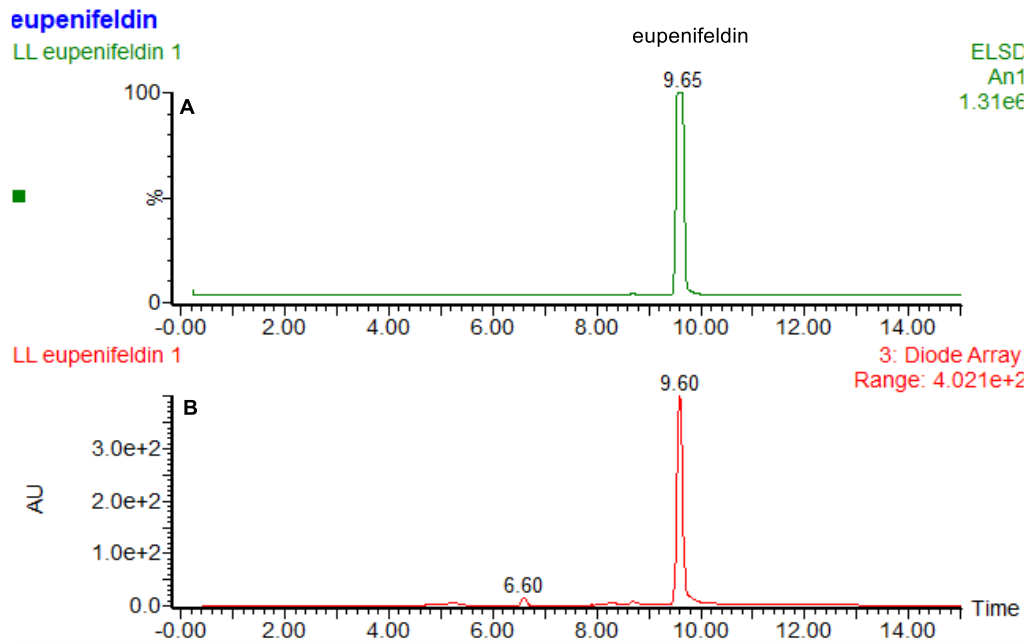
The genes *asL4* and *asL6*, encoding putative NAD/FAD-dependent oxidoreductases, were identified to catalyze two consecutive ring-contraction reactions that convert the compound **45**, the first tropolone sesquiterpenoids intermediate in the biosynthetic pathway, to xenovulene A **24** (Scheme 1.2.5.2).<sup>20</sup> Noreupenifeldin **53** and noreupenifeldin B **64** are reported to be the co-metabolites of eupenifeldin **23**.<sup>35,37</sup> It is proposed that compounds **53** and **64** are derived from a ring-contraction of eupenifeldin **23** similar to the biosynthesis of xenovulene A **24** (Scheme 2.8). Confirmation of this proposal will be investigated by transferring *asL4* and *asL6* into *A. oryzae* and supplementing the culture with eupenifeldin **23** as a substance.



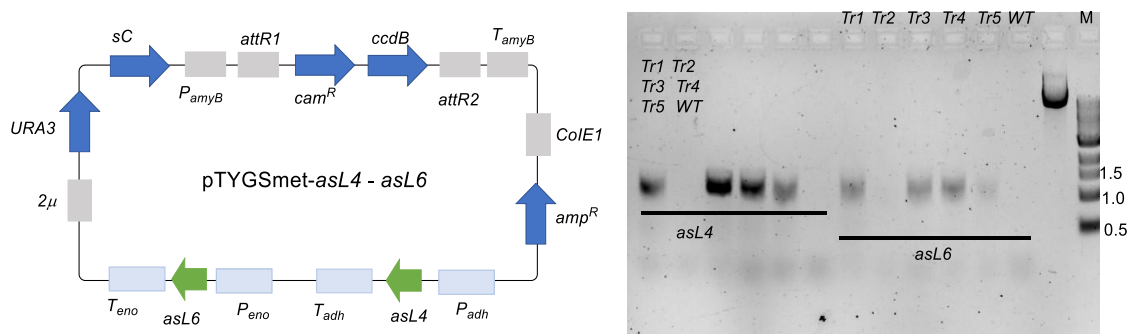
**Scheme 2.8** Ring-contraction reactions catalyzed by AsL4 and AsL6: **A**, natural substance of the enzymes; **B**, Proposed substance of the enzymes.

In order to get enough eupenifeldin **23** for the feeding experiments, *CF150626* was fermented according to a literature known fermentation protocol.<sup>37</sup> After 13 days' cultivation, 2L fungi cultures were mixed with equal volumes of acetone and stirred at RT for 1h. After filtration, the crude extract solution was concentrated to aqueous under reduced pressure. This aqueous phase was then extracted 2 times with equal amounts of EtOAc. The organic extracts were dried with anhydrous magnesium sulfate and evaporate to afford 2.6 g crude metabolites. In total, 40 mg pure eupenifeldin **23** was isolated by mass-directed preparative HPLC (Figure 2.8.1).

pTYGSmet-*asL4-asL6* was already constructed by Dr. Raissa Schor and was integrated into the genome of *A. oryzae* through transformation. Screening of 5 transformants were conducted by fermentation using CZD/S containing argenine but no methionine. To confirm the presence of *asL4* and *asL6*, the genomes of 5 transformants were used for PCR analysis (Figure 2.8.2). These data indicated that 4 transformants contained *asL4* and *asL6* genes as expected.

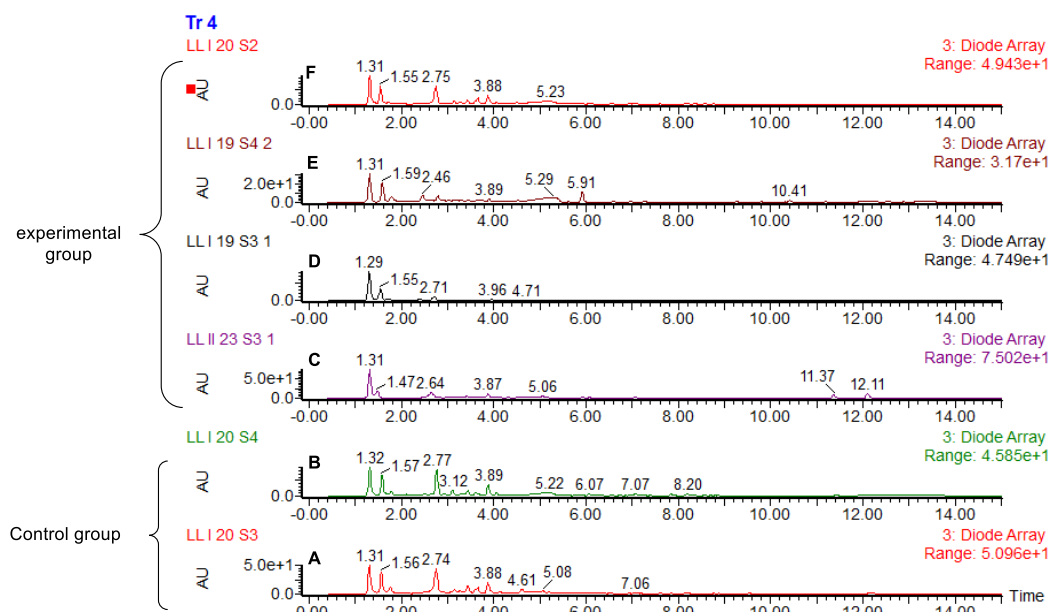


**Figure 2.8.1** LCMS analysis of eupenifeldin **23** used in the feeding experiments. **A**, ELSD chromatogram; **B**, UV chromatogram.



**Figure 2.8.2** PCR analysis from gDNA of 5 *A. oryzae* *asL4+asL6* transformants.

Next, eupenifeldin **23** was fed to *A. oryzae asL4+asL6* at day1 and cultured in DPY medium for different time. As a control, *A. oryzae* WT was cultured in DPY medium supplemented with eupenifeldin **23** and *A. oryzae asL4+asL6* was grown in DPY medium without eupenifeldin **23**. The cultures were extracted using ethyl acetate and analyzed by LCMS. In comparison to the UV chromatogram of the control group, no new products were generated by feeding eupenifeldin **23** to *A. oryzae asL4+asL6* mutants, and the peak of eupenifeldin **23** also disappeared in all the UV chromatogram (Figure 2.8.3). These data indicates that eupenifeldin **23** is not stable in the DPY medium or it is not the substance of the ring-contraction enzyme encoded by *asL4* and *asL6*.

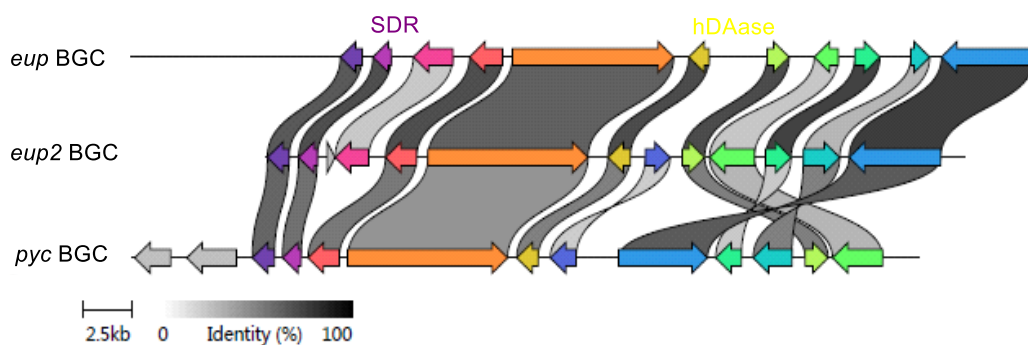


**Figure 2.8.3** Feeding eupenifeldin **23** to *A. oryzae asL4+asL6* transformants: **A**, LCMS chromatogram of *A. oryzae* WT with eupenifeldin **23** at day 6; **B**, LCMS chromatogram of *A. oryzae asL4+asL6* without eupenifeldin **23** at day 6; **C**, LCMS chromatogram of *A. oryzae asL4+asL6* with eupenifeldin **23** for 1h; **D**, LCMS chromatogram of *A. oryzae asL4+asL6* with eupenifeldin **23** at day 2; **E**, LCMS chromatogram of *A. oryzae asL4+asL6* with eupenifeldin **23** at day 4; **F**, LCMS chromatogram of *A. oryzae asL4+asL6* with eupenifeldin **23** at day 6.

## 2.9 Results-*In vitro* Activity Tests with Eup2L4, Eup2R1 and PycR1

Bioinformatic analysis of *eup2L4* from the *eup2* BGC shows that it encodes a short-chain dehydrogenase (SDR) and is homologous to *pycL2* from the *pyc* BGC (Figure 2.9.1), which indicates that it is likely to be involved in the biosynthesis of eupenifeldin **23**. Short-chain dehydrogenases are known to be NAD<sup>+</sup>- or NADP<sup>+</sup>-dependent oxidoreductases. Some members of this family are reported to facilitate the interconversion between alcohols and aldehydes or ketones with the reduction of nicotinamide adenine dinucleotide (NAD<sup>+</sup>) to NADH.<sup>155</sup> *Eup2R1* encodes an hDAase and is homologous to *pycR1* from the *pyc* BGC (Figure 2.9.1). Enzyme-catalyzed intermolecular hDA reactions, in which a 1,3-diene and a dienophile react to form cyclohexene regio- and stereoselectively are one of the most powerful synthetic strategies for the synthesis of

complex natural products.<sup>156</sup> Eup2R1 and PycR1 are proposed to be the key enzyme that catalyze hDA reactions between tropolone and the sesquiterpenoid precursors to produce tropolone sesquiterpenoids. So, we proposed that Eup2L4 catalyzes the reduction of stipitaldehyde **102**, the known tropolone intermediate in the eupenifeldin **23** biosynthetic pathway, and the product will be used for one of the substances for Eup2R1 or PycR1 in the enzymatic reactions.

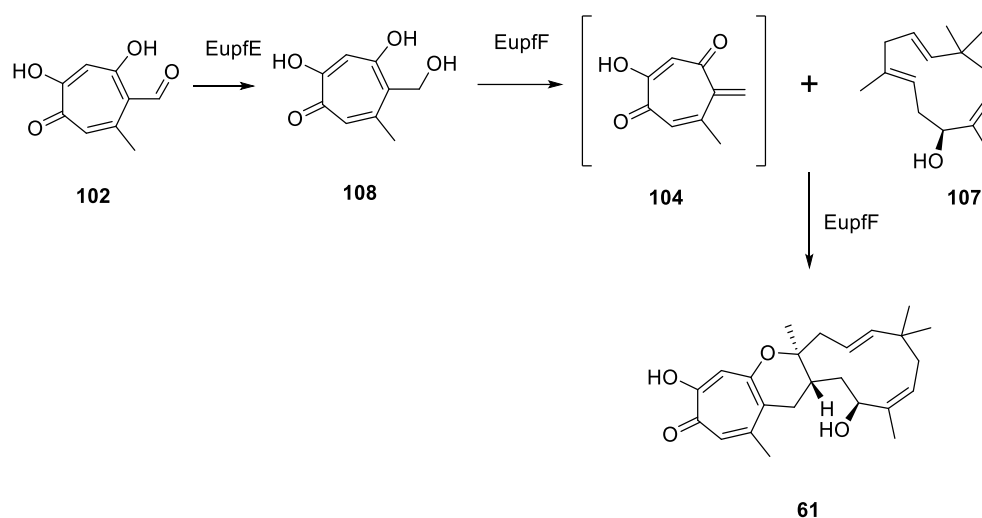


**Figure 2.9.1** Clinker analysis of BGC of eupenifeldin (*eup* and *eup 2* BGC) and pycnidione (*pyc* BGC).

In 2019, the enzymatic steps that catalyzed by short-chain dehydrogenase (SDR) EupfE and an intermolecular hDAase (EupfF) from the BGC of eupenifeldin **23** in *Penicillium janthinellum* were explored *via in vitro* assay. The results showed that EupfE converted stipitaldehyde **102** to stipitol **108** through reductive reaction. EupfF was reported to catalyze not only the dehydration of stipitol **108** to generate a reactive tropolone o-quinone methide **104** but also the stereoselective hDA reaction between o-quinone methide **104** and 2Z-humulene **107** to produce known natural product neosetophomone B **61** (scheme 2.9.2), which further served as an intermediate for the biosynthesis of eupenifeldin **23**.<sup>55</sup> For the first time, the generation of mono-substituted tropolone sesquiterpenoids *via* hDA reaction under *in vitro* conditions was achieved successfully.

While the enzymatic functions of EupfE and EupfF have been identified, the enzymatic basis of Eup2L4, Eup2R1 and PycR1 have remained unexplored. Eup2L4 is the homologous protein of EupfE while Eup2R1 and PycR1 are the homologous proteins of EupfF. In addition, we reasoned that rational extension and diversification of enzymatic substance should expand the chemical space for the synthesis of unnatural tropolone sesquiterpenoids. Therefore, it is of interest to test the *in*

*vitro* activity of Eup2L4, Eup2R1 and PycR1 with the reported natural substances and further explore their substrate selectivity to synthesize optically pure new products.



**Scheme 2.9.2** Enzymatic reactions catalyzed by EupfE and EupfF.

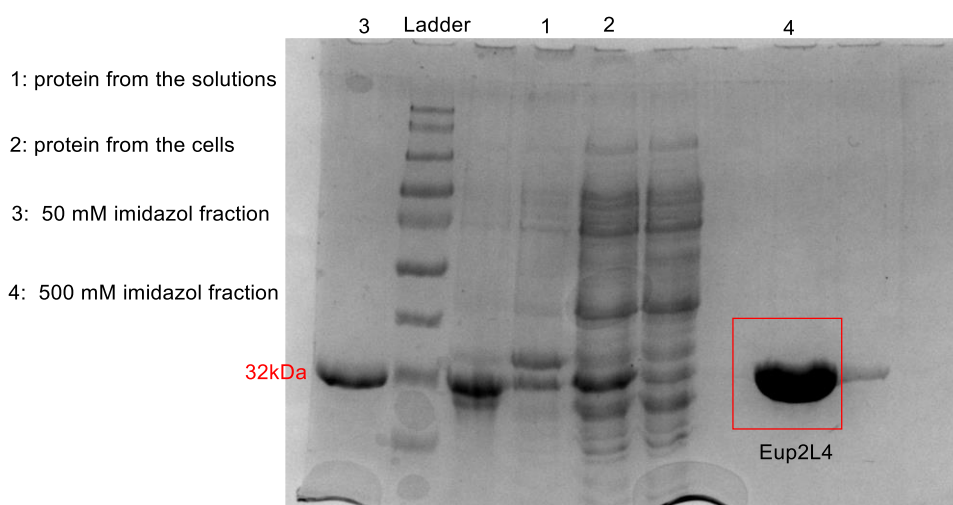
### 2.9.1 Expression and purification of protein Eup2L4 and Eup2R1

Codon-optimization of *eup2L4* and *eup2R1* were designed using the standard tool provided by Thermo Fischer Scientific. These sequences were then used for the synthesis of N-terminal his<sub>6</sub> tagged protein expression cassettes. The successful construction of the codon-optimized gene cassettes for *eup2L4* and *eup2R1* was confirmed by sequencing.

Eup2L4 was successfully overexpressed as a soluble enzyme after transforming pET100-*eup2L4* plasmid into *E. coli BL21*. A single kanamycin resistant colony was picked from a selection plate and pre-cultured overnight in LB-medium containing 50 µg/ mL kanamycin at 37 °C with 200 rpm shaking. The cells were then transferred into 2TY medium with kanamycin and grown at 37 °C and 220 rpm until an OD<sub>600</sub> between 0.4-0.6 was reached. Next, 0.1M isopropylthio-β-galactoside (IPTG) were added to induce protein expression and the bacterial culture was incubated at 16 °C and 220 rpm shaking condition overnight. The soluble protein production was extracted and analyzed using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

To get pure soluble Eup2L4, 1 L 2TY medium was used to cultivate *E. coli BL21* with pET100-

*eup2L4* plasmid using the same expression conditions described above. The cells were harvested by centrifugation and suspended in lysis buffer (pH 8) for sonification. The obtained crude lysate was further purified with Ni-NTA chromatography. Elution was done by a stepwise imidazole gradient (50 mM and 500 mM imidazole). Each fraction was collected separately and analysed by SDS-PAGE (Figure 2.9.1.1). Finally, pure Eup2L4 was founded in the 500 mM imidazole fraction based on SDS-PAGE analysis. A strong band at the size of ~32 kDa matched the expected size of the His<sub>6</sub>-Eup2L4.

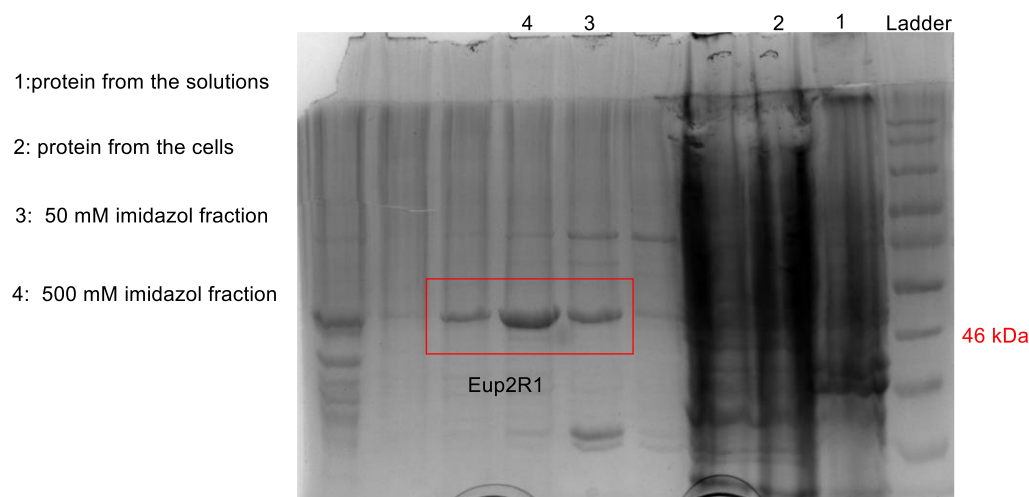


**Figure 2.9.1.1** The SDS-PAGE analysis of Eup2L4 after the purification of Ni-NTA chromatography.

Eup2R1 was also solubly expressed after transforming the synthesized codon-optimized protein expression plasmid pET100-*eup2R1* into *E. coli BL21*. A single kanamycin resistant colony was pre-cultured in LB-medium containing 50 µg/ mL kanamycin at 37 °C with 200 rpm shaking for overnight. Then, the cells were grown in TB medium with kanamycin at 37 °C and 220 rpm until an OD<sub>600</sub> between 0.4-0.6 was reached. Next, the bacterial cultures were incubated at 10 °C and 220 rpm shaking condition for 3 days. The soluble Eup2R1 protein production was extracted and analyzed using SDS-PAGE.

To get pure soluble Eup2R1, 1L TB medium was used to cultivate *E. coli BL21* with pET100-*eup2R1* plasmid in the expression condition mentioned above. The cells were harvested by centrifugation and suspended in lysis buffer (pH 8) for sonification. The obtained crude lysate was

further purified by Ni-NTA chromatography with the same protocol as Eup2L4 purification. SDS-PAGE analysis indicated that His6-Eup2R1 had an expected molecular weight of ~46 kDa (Figure 2.9.1.2).

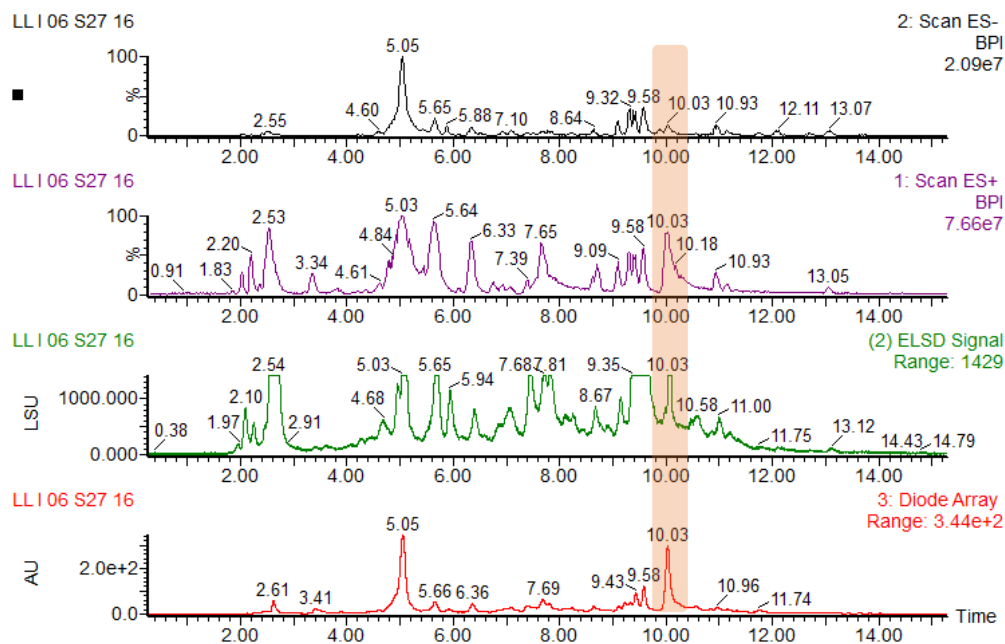


**Figure 2.9.1.2** The SDS-PAGE analysis of Eup2L4 after the purification of Ni-NTA chromatography.

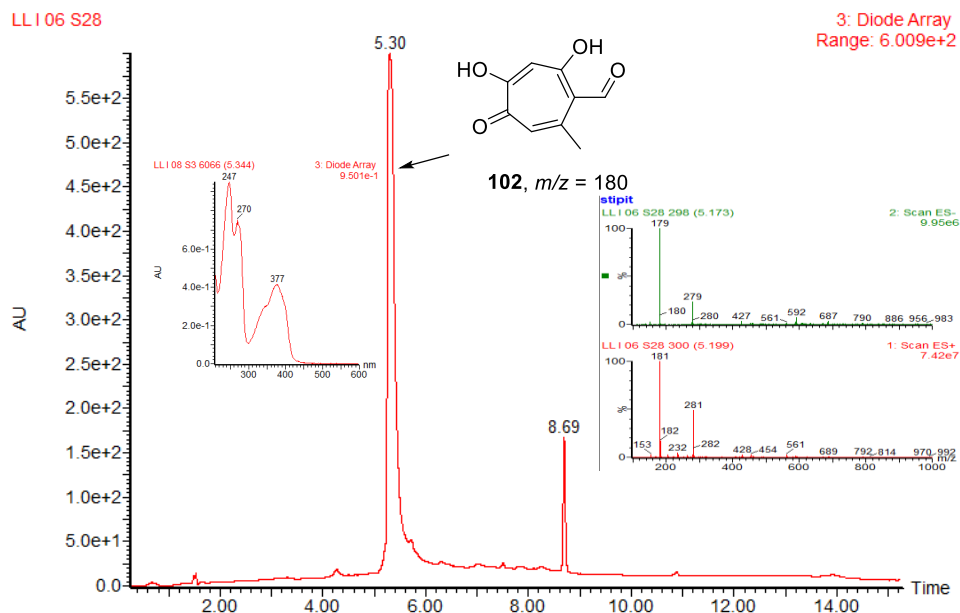
## 2.9.2 Stipitaldehyde Purification

Stipitaldehyde **102** was supposed to be the natural substance of Eup2L4, and was reported to be the main products of *T. stipitatus* *ΔtsL2*.<sup>157</sup> To obtain pure stipitaldehyde **102** for *in vitro* assay, *T. stipitatus* *ΔtsL2* (transformants are from Dr. Ahmed Al Fahad)<sup>157</sup> was regrown in sucrose medium at 28 °C, 220 rpm in flasks with metal springs inside for seven days. A large scale of fermentation (1L) was filtered and acidified to pH 3 with HCl. Then ethyl acetate was used to extract the metabolites from the culture 3 times. The organic layer was dried with MgSO<sub>4</sub> and evaporated to give solid crude residue, which was then used for preparative LCMS purification. The peak eluted at  $t_R = 10.0$  min with UV absorption  $\lambda_{max}$  247, 270 and 370 nm was supposed to be stipitaldehyde **102** (Figure 2.9.2.1), which was further confirmed by extracted ion chromatograms using 181 Da (ES<sup>+</sup> spectrum) and 179 Da (ES<sup>-</sup> spectrum). Additionally, another peak eluting at  $t_R = 5.0$  min was characterized to be a known benzaldehyde **101** based on the UV and MS spectrum analysis, the precursor of stipitaldehyde **102**.<sup>158</sup> Finally, 12 mg stipitaldehyde **102** was successfully obtained and confirmed by LCMS analysis (Figure 2.9.2.2).





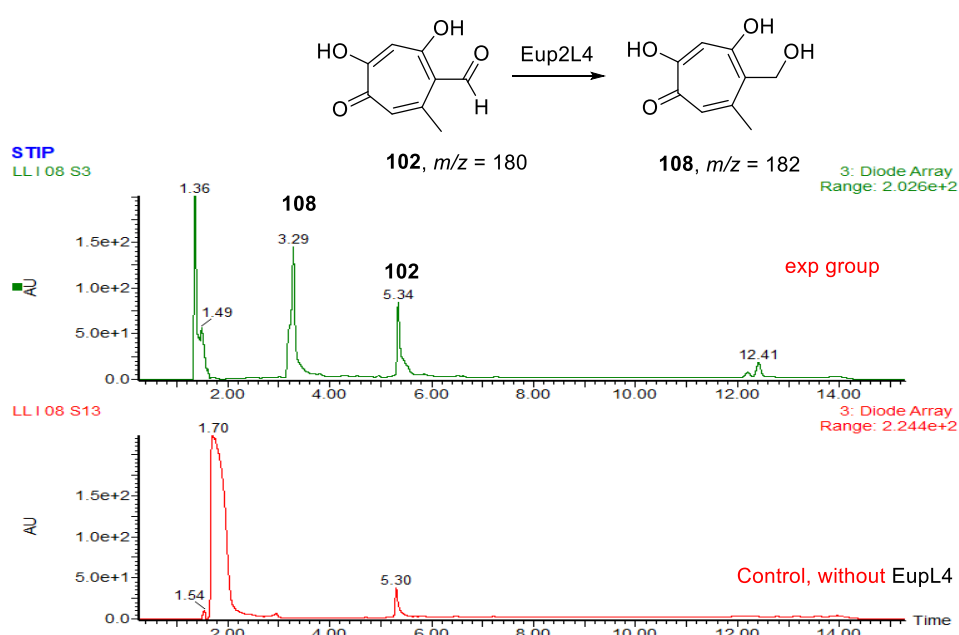
**Figure 2.9.2.1** The preparative LCMS chromatogram of *T. stipitatus*  $\Delta$ tsL2 culture extracts after 7 days' growth in flask with metal spring inside. Stipitaldehyde **102** is highlighted with orange color: **A**, ES<sup>-</sup> TIC chromatogram of the culture extracts; **B**, ES<sup>+</sup> TIC chromatogram of the culture extracts; **C**, ELSD chromatogram of the culture extracts; **D**, UV chromatogram of the culture extracts.



**Figure 2.9.2.2** LCMS analysis of the purified stipitaldehyde **102**.

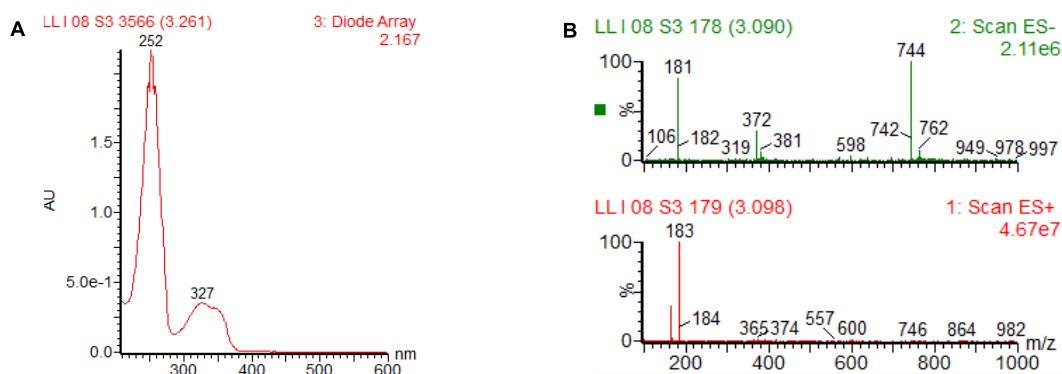
### 2.9.3 *In vitro* assay of Eup2L4

An *in vitro* assay of Eup2L4 was carried out in the same condition as the literature reported.<sup>55</sup> The freshly prepared Eup2L4 (25  $\mu$ M) was incubated with substrate stipitaldehyde **102** (2.5 mM) and NADPH (3 mM) in 50 mM bicine buffer (pH = 8) for 20 min at 26 °C. The reaction was quenched with MeCN (1:1). Control assays were carried out using denatured Eup2L4 protein under identical conditions, which meant the only difference between control group and experimental group was enzyme. The precipitated proteins were removed *via* centrifugation. The supernatants were analyzed directly by analytical LCMS.



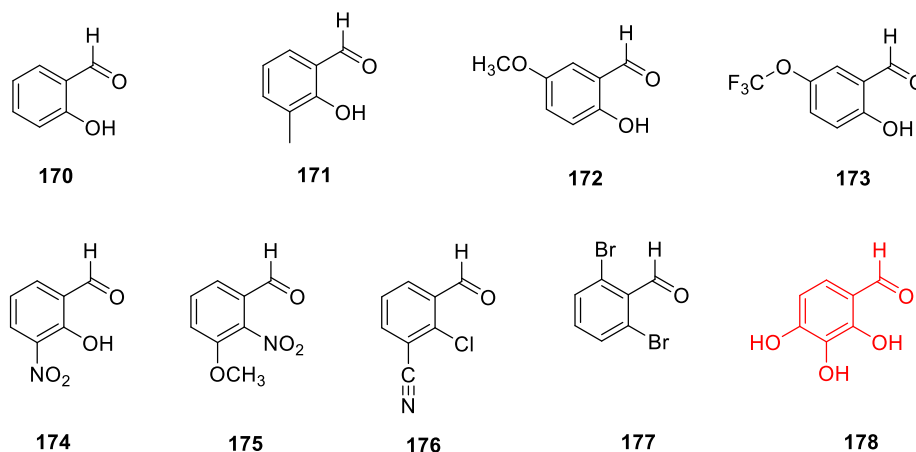
**Figure 2.9.3.1** LCMS analysis of Eup2L4 *in vitro* study with stipitaldehyde **102**.

The LCMS chromatogram of the Eup2L4 assay showed a new peak at  $t_R = 3.3$  min with UV absorption  $\lambda_{max}$  252 and 327 nm (Figure 2.9.3.1), which was identified to be stipitol **108** based on the extracted ion chromatograms using 183 Da (ES<sup>+</sup> spectrum) and 181 Da (ES<sup>+</sup> spectrum, Figure 2.9.3.2). While the LCMS analysis of the control assay only showed the presence of stipitaldehyde **102**. The results confirmed that Eup2L4 had *in vitro* assay activity that reduced stipitaldehyde **102** to stipitol **108**.



**Figure 2.9.3.2** Characterisation of stipitol **108**: **A**, UV spectrum of stipitol **108**; **B**, MS spectrum of stipitol **108**.

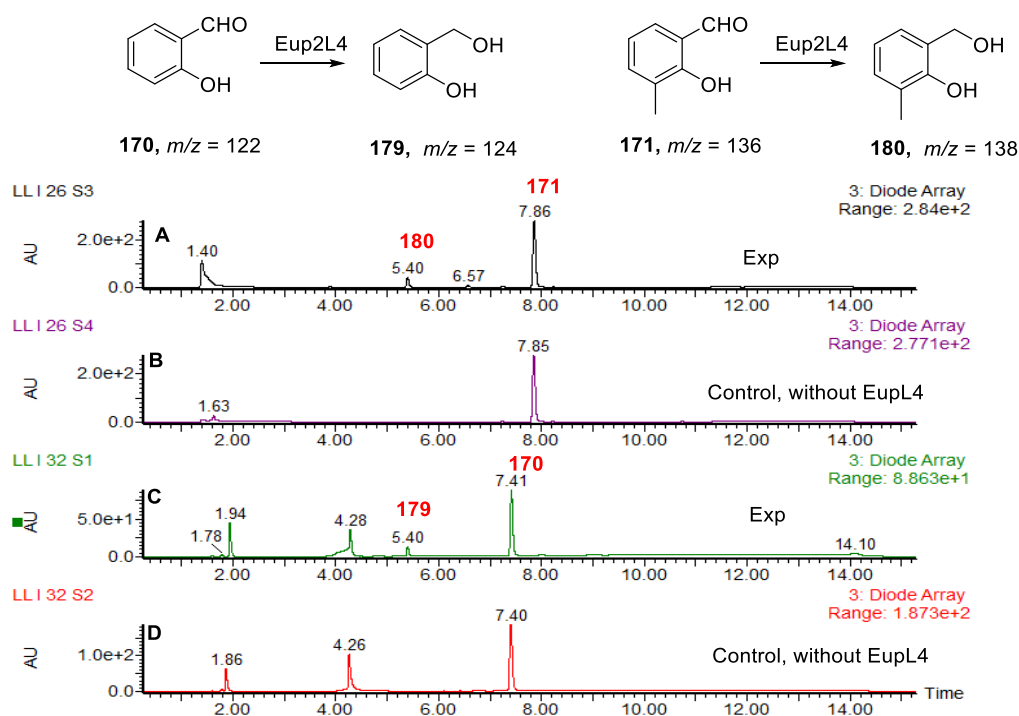
Next, we extended these initial investigations by exploring the enzyme's tolerance to a broad range of alternative aromatic aldehydes with different substituted groups. Different substrates including salicylaldehyde **170** and a number of analogous compounds **171 - 178** were obtained commercially and used for the substance of Eup2L4 in the *in vitro* assay (Figure 2.9.3.3).



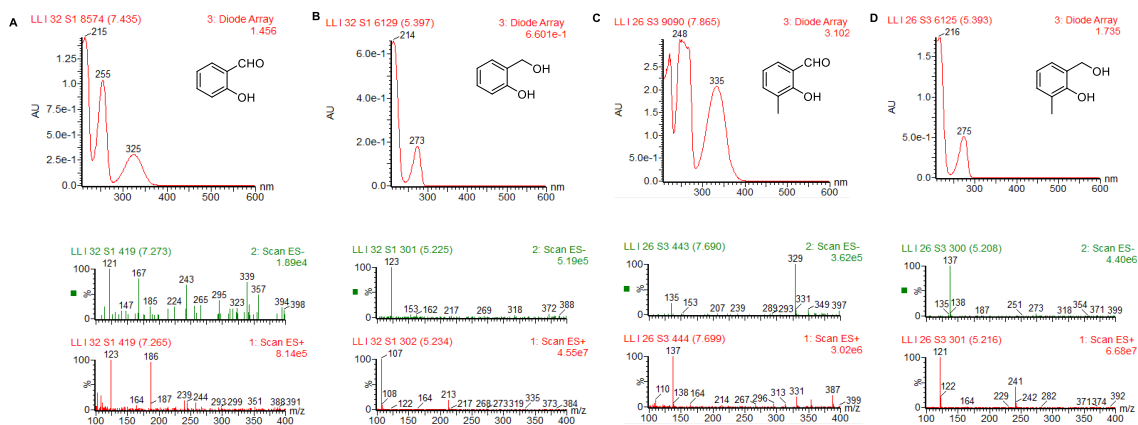
**Figure 2.9.3.3** Aromatic aldehydes used for Eup2L4 *in vitro* assay.

In order to ascertain which of the unnatural substrates Eup2L4 would be able to act upon, the freshly prepared enzyme (25  $\mu$ M) was incubated with compound **170-178** (2.5 mM) and NADPH (3 mM) in 50 mM bicine buffer (pH 8) for 20 mins. The reaction solutions were quenched with MeCN (1:1) and were directly analyzed by LCMS after removing the precipitated protein *via* centrifugation.

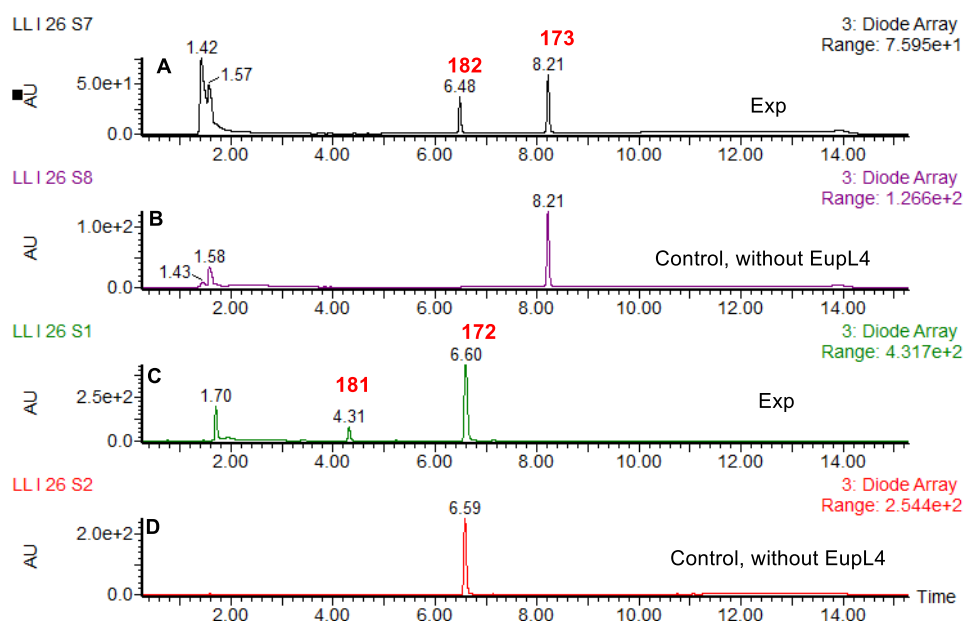
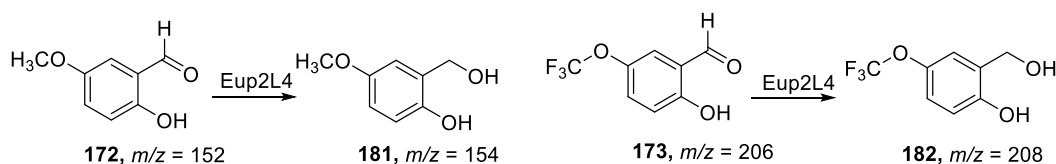
Surprising, the enzymatic reaction of compound **170-177** all showed the turnover of the substrates to the formation of the corresponding alcohol products (Figure 2.9.3.4-11). The structures of the products were confirmed based on the UV and MS data. The aldehyde group in compound **170 - 177** which showed UV absorption from  $\lambda_{max}$  300 to 360 nm, which disappeared after the reduction reaction catalyzed by Eup2L4. In addition, 2 Da more than that of the substrates indicated by MS spectra was consistent with an aldehyde reduction. However, the LCMS analysis of the Eup2L4 assay with compound **178** did not show the presence of new peak compared to the control group (Figure 2.9.3.12). In conclusion, Eup2L4 displayed broad substrate tolerances towards aromatic aldehydes and was a good candidate for enzymatic catalyzed chemo-synthetic processes.



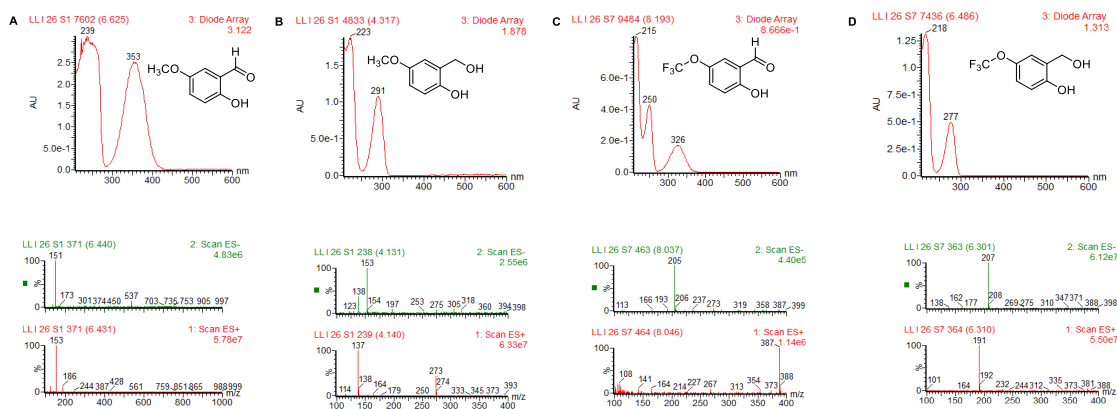
**Figure 2.9.3.4** LCMS analysis of Eup2L4 *in vitro* study with compound **170** and **171**.



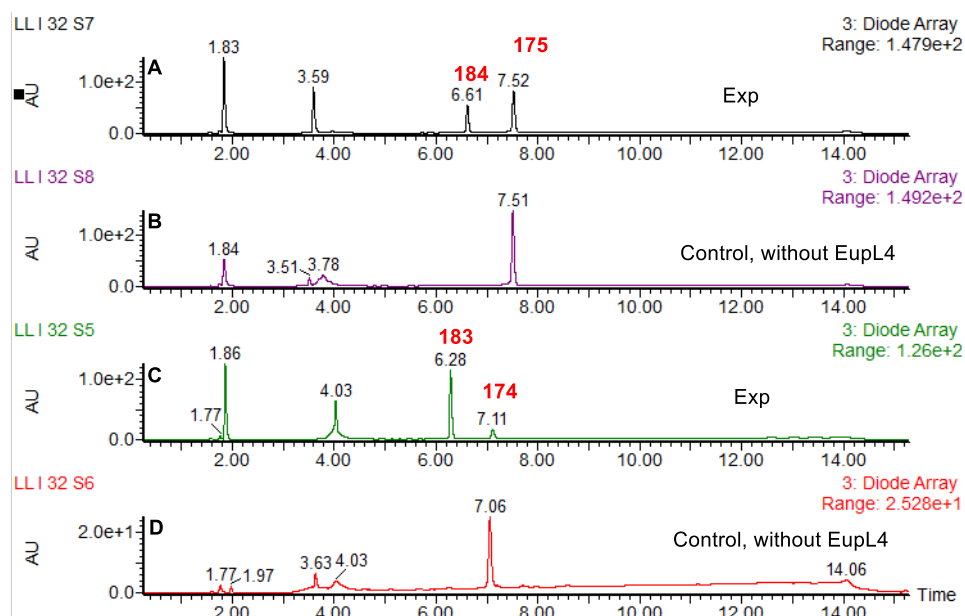
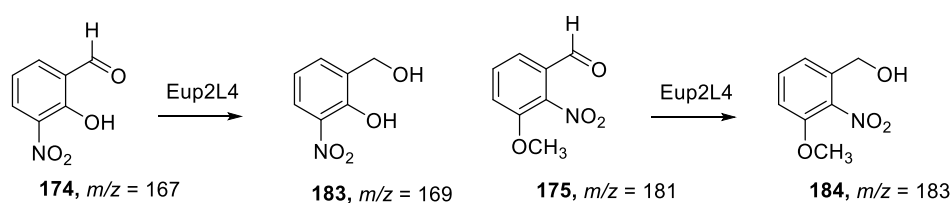
**Figure 2.9.3.5** Characterisation of compound **170**, **179**, **171** and **180**: **A**, UV and MS spectrum of compound **170**; **B**, UV and MS spectrum of compound **179**; **C**, UV and MS spectrum of compound **171**; **D**, UV and MS spectrum of compound **180**.



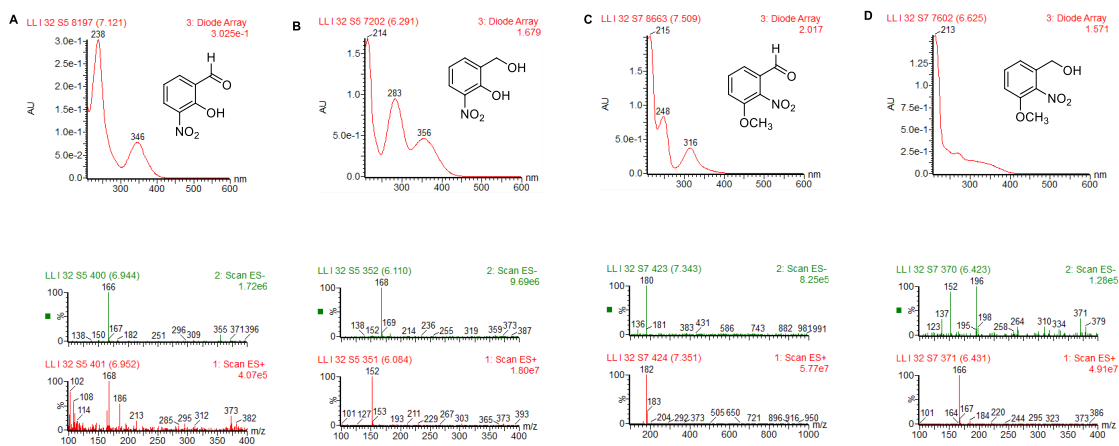
**Figure 2.9.3.6** LCMS analysis of Eup2L4 *in vitro* study with compound **172** and **173**.



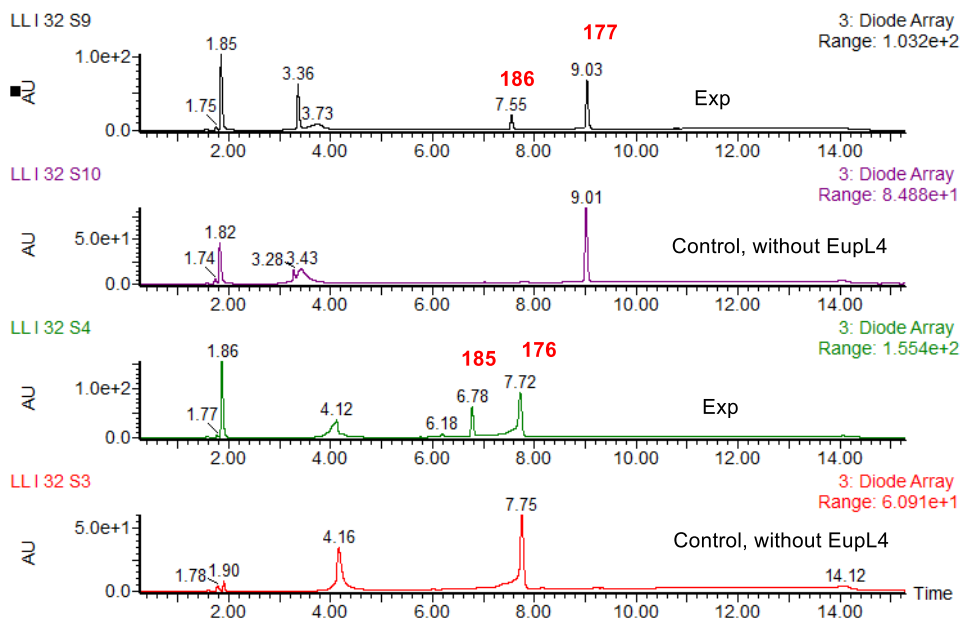
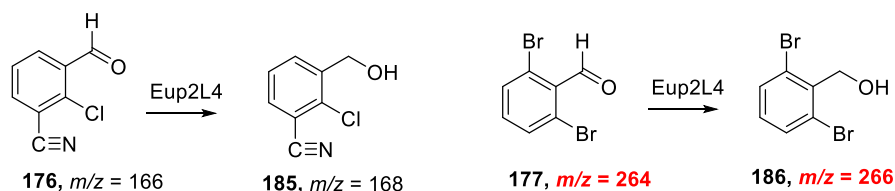
**Figure 2.9.3.7** Characterisation of compound **172**, **181**, **173** and **182**: **A**, UV and MS spectrum of compound **172**; **B**, UV and MS spectrum of compound **181**; **C**, UV and MS spectrum of compound **173**; **D**, UV and MS spectrum of compound **182**.



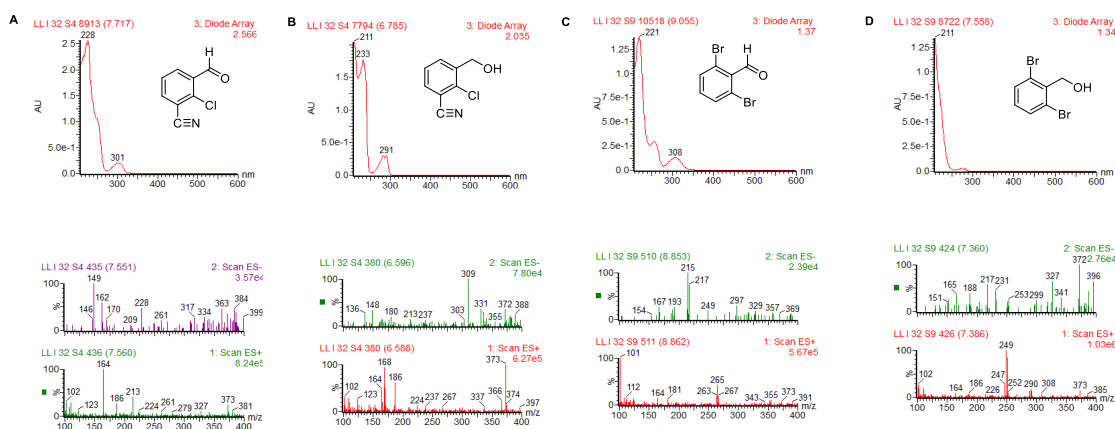
**Figure 2.9.3.8** LCMS analysis of Eup2L4 *in vitro* study with compound **174** and **175**.



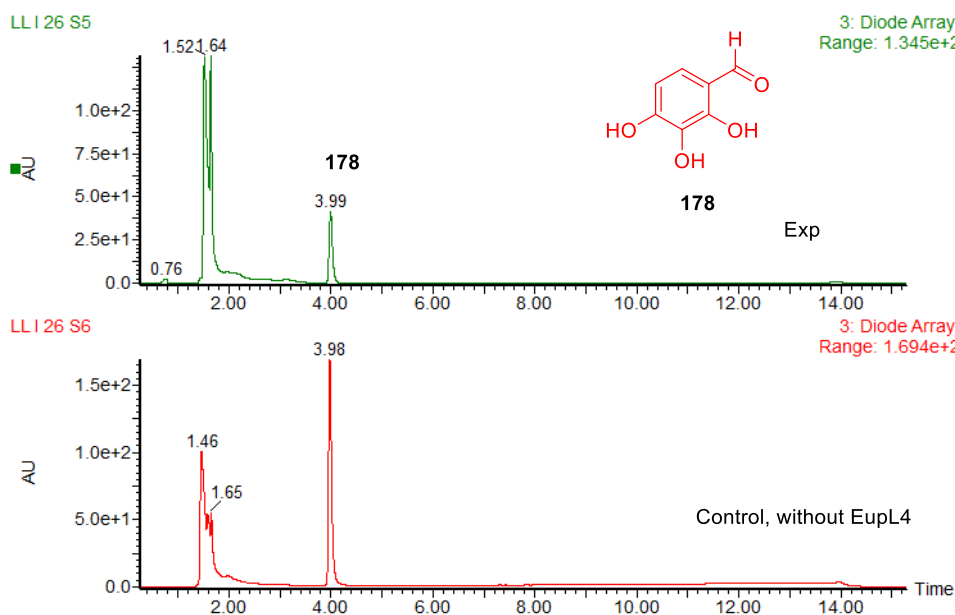
**Figure 2.9.3.9** Characterisation of compound **174**, **183**, **175** and **184**: **A**, UV and MS spectrum of compound **174**; **B**, UV and MS spectrum of compound **183**; **C**, UV and MS spectrum of compound **175**; **D**, UV and MS spectrum of compound **184**.



**Figure 2.9.3.10** LCMS analysis of Eup2L4 *in vitro* study with compound **176** and **177**.



**Figure 2.9.3.11** Characterisation of compound **176**, **185**, **177** and **186**: **A**, UV and MS spectrum of compound **176**; **B**, UV and MS spectrum of compound **185**; **C**, UV and MS spectrum of compound **177**; **D**, UV and MS spectrum of compound **186**.



**Figure 2.9.3.12** LCMS analysis of Eup2L4 *in vitro* study with compound **178**.

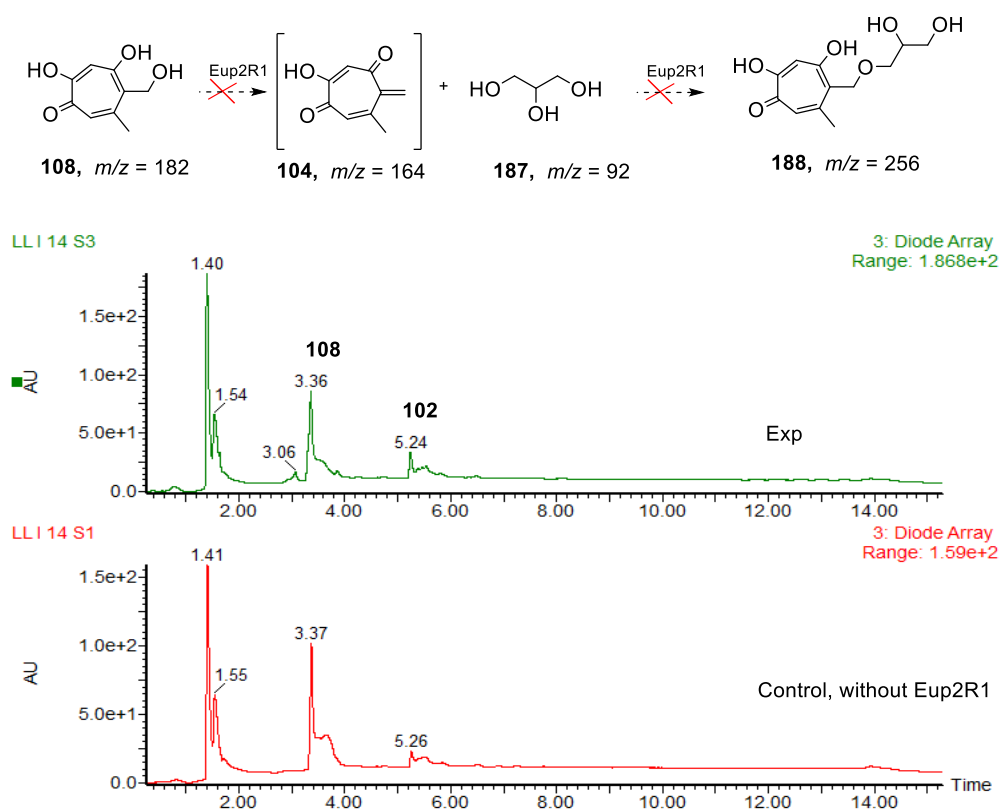
## 2.9.4 *In vitro* assay of Eup2R1

To explore if Eup2R1 catalyzed the dehydration of stipitol **108** as reported, *in vitro* assay of Eup2R1 was performed with stipitol **108** as the sole substrate in the same condition as the literature reported.<sup>55</sup> Stipitaldehyde **102** was obtained with enzymatic reaction catalyzed by Eup2L4. In this section, stipitaldehyde **102** (2.5 mM) and NADPH (3 mM) were incubated with Eup2L4 (25  $\mu$ M) and



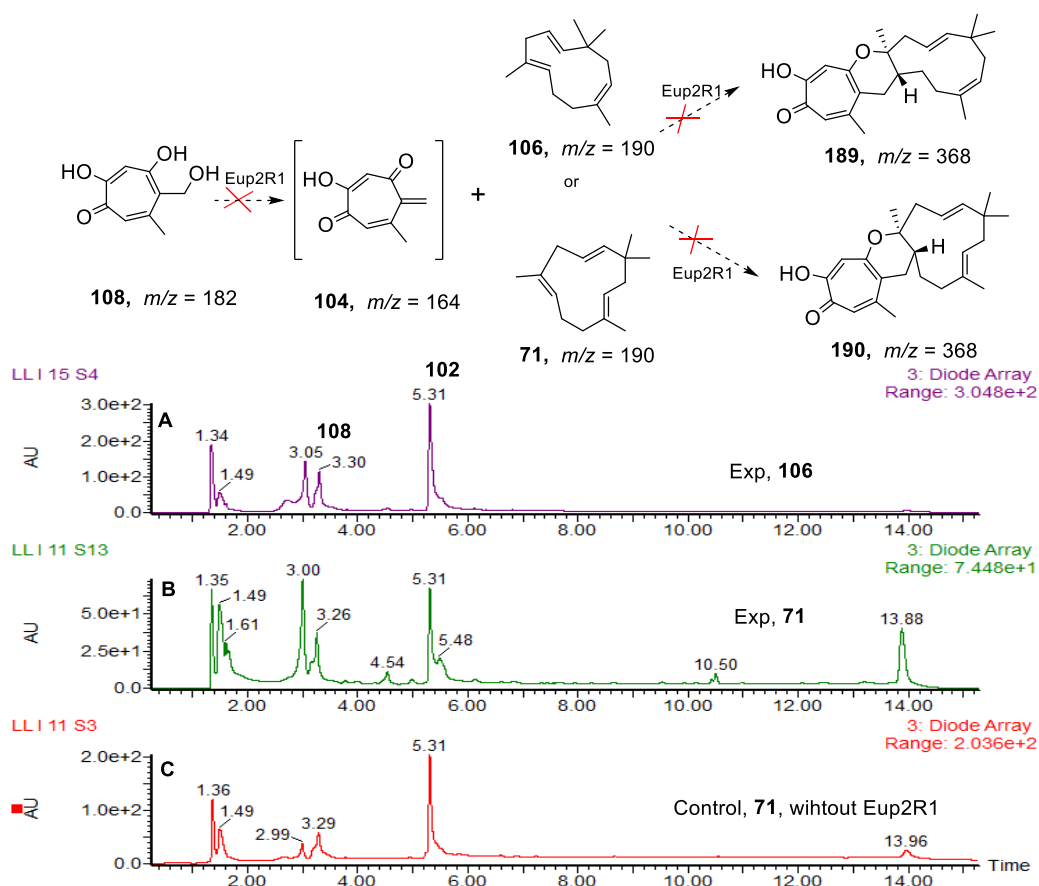
Eup2R1 (50  $\mu$ M) at 26  $^{\circ}$ C for 2 hours in 100  $\mu$ L bicine buffer (pH 8). Control assay was carried out without Eup2R1 protein in the identical condition. The reaction was quenched with MeCN (1:1). Protein was removed by heating and centrifugation while the supernatant was analyzed directly by LCMS. Since the expected product tropolone o-quinone **104** was extremely instable, 1 % glycerol **187** was added in the *in vitro* system to trap **104**. It was reported that **104** reacted spontaneously with glycerol to form addition product **188**.<sup>55</sup>

Therefore, a new peak **188** with mass of 256 in the LCMS chromatogram was expected. Unfortunately, the LCMS chromatogram of the Eup2R1 assay showed no interesting new peak compared to the control group. The peak eluting at 3.4 min and the peak eluting at 5.3 min were identified to be stipitol **108** and stipitaldehyde **102** based on UV and MS data (Figure 2.9.4.1). The results from the *in vitro* assay supported that Eup2R1 did not show the activity that catalyzes the dehydration of stipitol **108** like EupF (scheme 2.9.2).



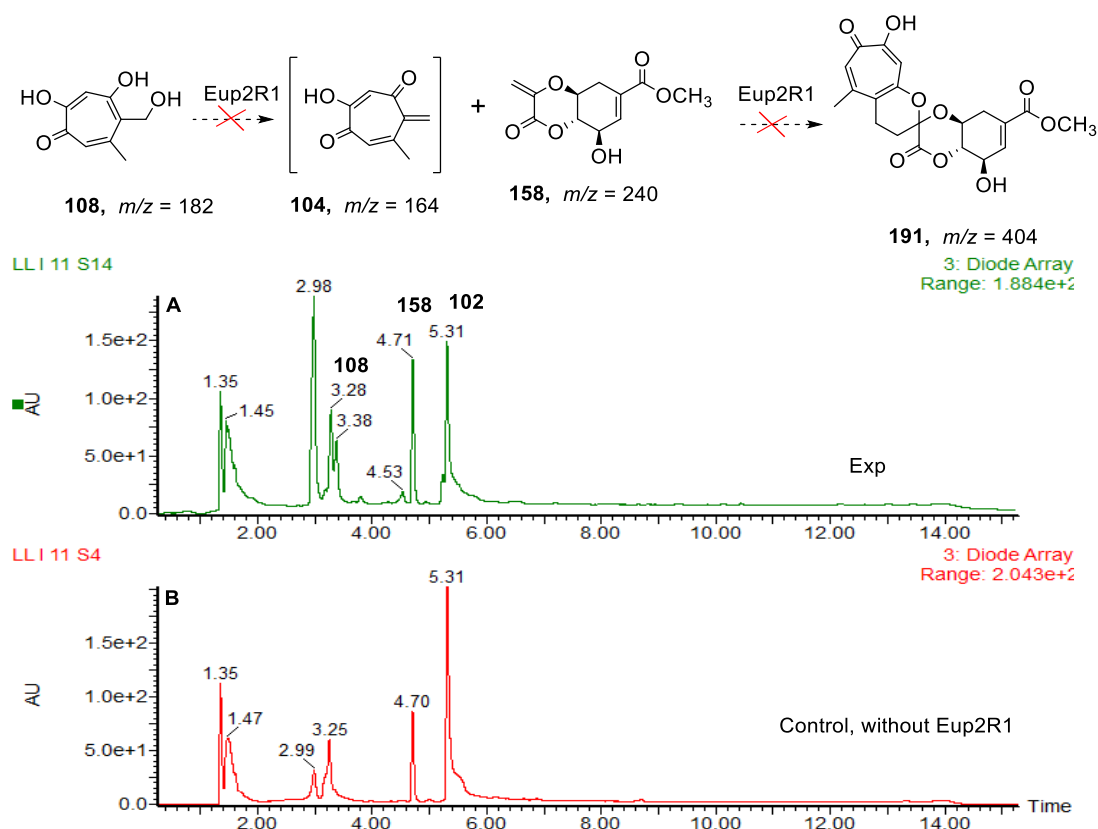
**Figure 2.9.4.1** LCMS analysis of Eup2R1 *in vitro* study with stipitol **108** and glycerol **187**.

In addition, different dienophiles were used to examine whether Eup2R1 was able to catalyze the hDA reaction, including *2E*-humulene **71**, *2Z*-humulene **106** and scytolide **158**. Our group had carried out the *in vivo* study on the DA reactions during the biosynthesis of sorbicillinoids through heterologous expression.<sup>102</sup> The result indicated that the biosynthesis of spiro-sorbicillinols A **160** and B **161** were formed though DA reactions between sorbicillinol **159** with scytolide **158**, which is catalyzed by flavin-dependent monooxygenase SorD.<sup>102</sup> In this biosynthetic chemical reaction, scytolide **158** is an active dienophile. While *2E*-humulene **71** and *2Z*-humulene **106** are the analogues of *2Z*-humulenol **107**, an active dienophile that served as the native substrate of EupfF. Scytolide **158** was purified from the crude metabolites of *P. cirsi* (Chapter 3.2.2); *2E*-humulene **71** was ordered commercially and *2Z*-humulene **106** was obtained by with enzymatic reaction catalyzed by Eup2R3 with FPP **105** as substrate.<sup>62</sup>



**Figure 2.9.4.2** LCMS analysis of Eup2R1 *in vitro* study with stipitol **108** and *2E*-humulene **71**/*2Z*-humulene **106**.

An *in vitro* reaction containing 25  $\mu\text{M}$  Eup2L4, 50  $\mu\text{M}$  Eup2R1, 2.5 mM stipitaldehyde **102**, 3 mM NADPH and 2 mM 2*E*-humulene **71** in 100  $\mu\text{l}$  50 mM bicine buffer (pH 8.0) was performed at 26  $^{\circ}\text{C}$  for 2 hours; In the second experiment group, 2*E*-humulene **71** was replaced with Eup2R3 (0.5 mg/ml),  $\text{Mg}^{2+}$  (5 mM) and FPP **105** (1 mM). In the third *in vitro* assay, the dienophile was switched to scytolide **158** (2 mM). Control assays were carried out without Eup2R1 protein under the identical conditions. The reactions were then quenched with MeCN (1:1). Proteins were removed by heating and centrifugation while the supernatants were analyzed directly by LCMS. Unfortunately, the expected product **191** was not observed in the analytical LCMS chromatograms of Eup2R1 assay (Figure 2.9.4.2-3). Hence, Eup2R1 showed no hDAase activity in all the tested *in vitro* assay.



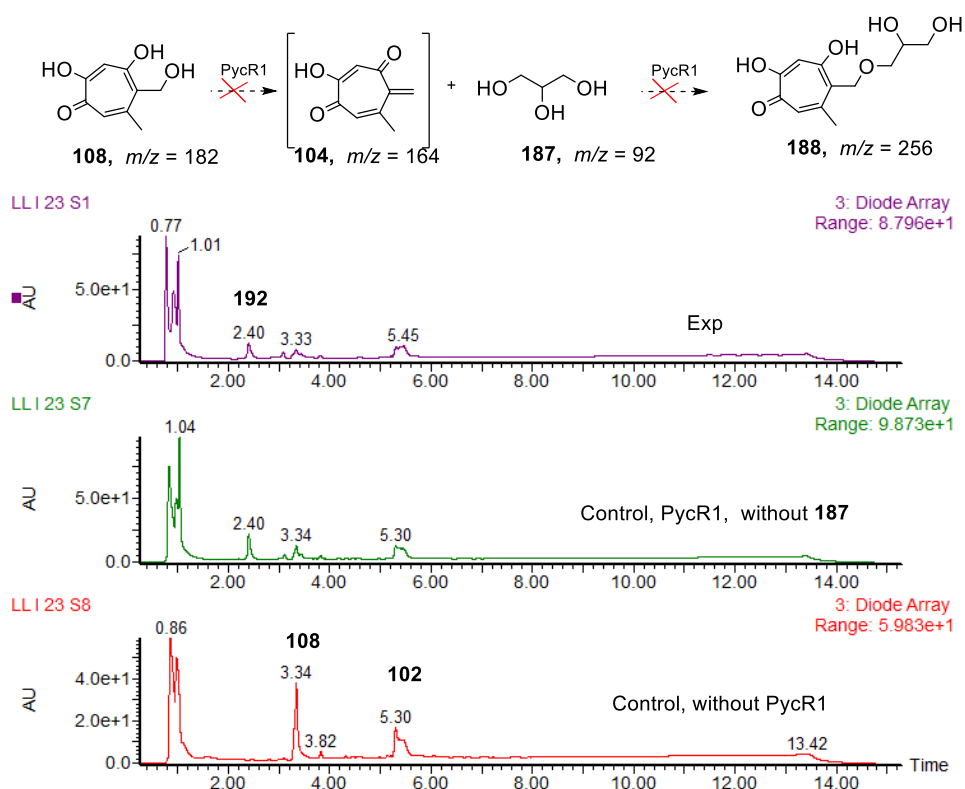
**Figure 2.9.4.3** LCMS analysis of Eup2R1 *in vitro* study with stipital **108** and scytolide **158**.

### 2.9.5 *In vitro* Assay of PycR1

Because purified Eup2R1 seemed to show no bioactivity *in vitro*, PycR1, the homolog from the pyc BGC, was tested *via in vitro* assay. Soluble PycR1 was also expressed and purified in the same protocol as Eup2R1 mentioned previously (Chapter 2.9.1), the expression plasmid pET100-*pycR1*

was provided by Dr Carsten Schotte. First of all, the proposed dehydration activity of pycR1 was tested with stipitol **108** and glycerol **187** as substrate. Next PycR1 was tested in this study to check if they could catalyze the hDA reactions between tropolone o-quinone **104** and different dienophiles.

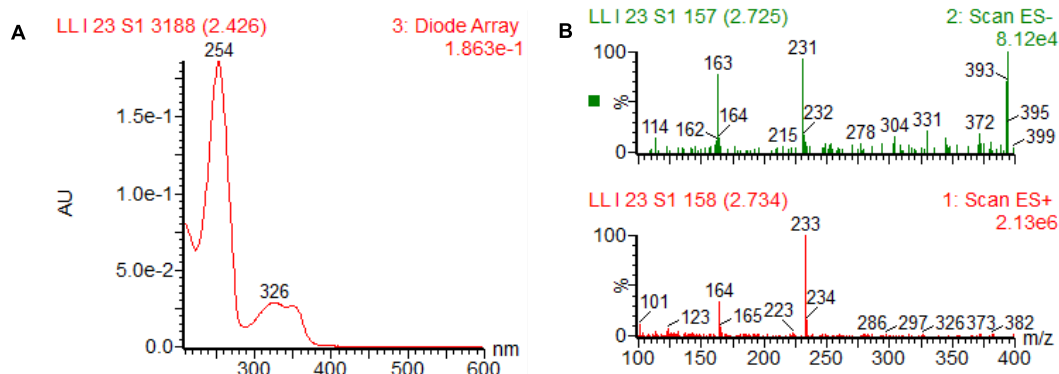
PycR1 *in vitro* assay was carried out in the same condition used for Eup2R1 (chapter 2.9.4). The freshly prepared PycR1 (50  $\mu$ M) and Eup2L4 (25  $\mu$ M) were incubated with stipitaldehyde **102** (2.5 mM), 1 % glycerol **187** and NADPH (3 mM) in 50 mM bicine buffer (pH 8) for 2 h at 26 °C. Control assay was carried out without PycR1 protein in the identical condition. The reaction was quenched with MeCN (1:1) and the precipitated protein was removed *via* centrifugation. The supernatant was analyzed directly by LCMS.



**Figure 2.9.5.1** LCMS analysis of PycR1 *in vitro* study with stipitol **108** and glycerol **187**.

The LCMS chromatogram of the PycR1 assay showed a new peak at  $t_R = 2.4$  min with UV absorption  $\lambda_{max}$  254 and 326 nm, which was similar with that of stipitol **108** (Figure 2.9.5.1). This compound **192** had a mass of 232 based on the extracted ion chromatograms using 233 Da ( $ES^+$  spectrum) and 231 Da ( $ES^-$  spectrum, Figure 2.9.5.2). The peak of stipitol **108** almost disappeared

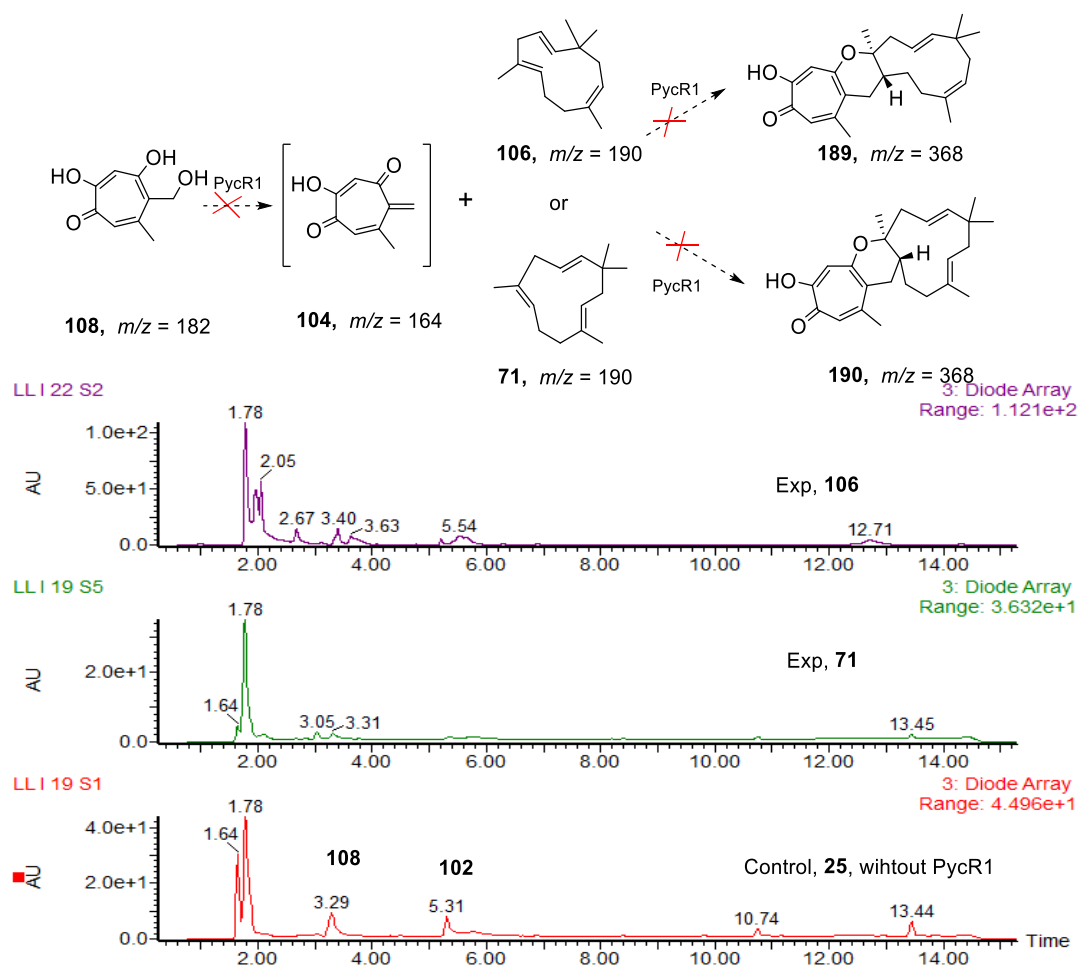
compared to the control assay without PycR1. However, the new peak also appeared in the control group without glycerol **187**. These results indicated that PycR1 used stipitol **108** as substance to give a new product instead of the expected known compound **24** with mass of 256.



**Figure 2.9.5.2** Characterisation of the new peak **192**: **A**, UV spectrum of the new peak **192**; **B**, MS spectrum of the new peak **192**.

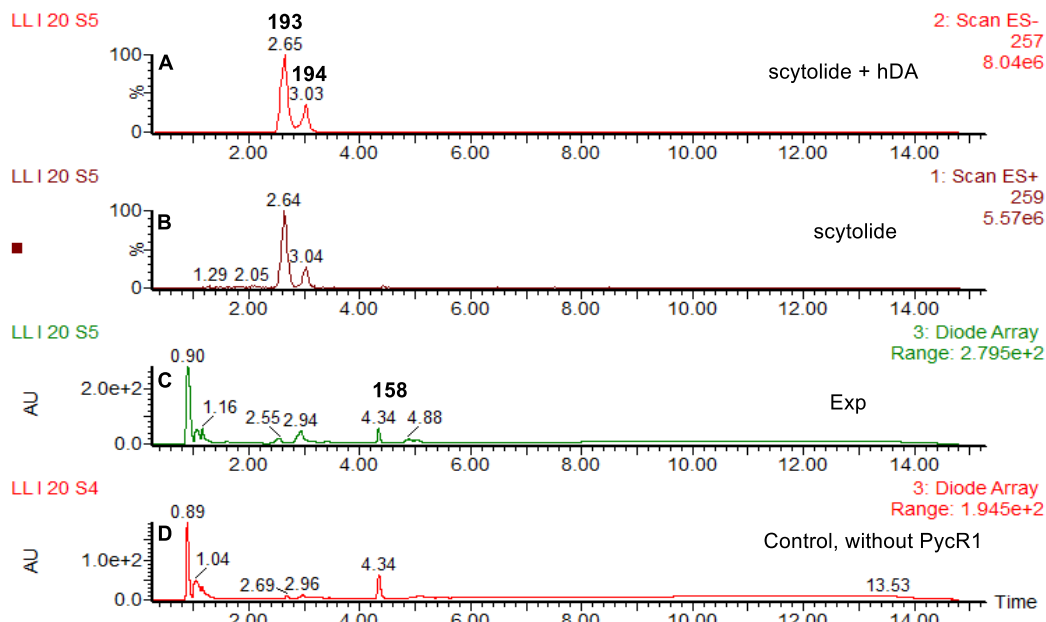
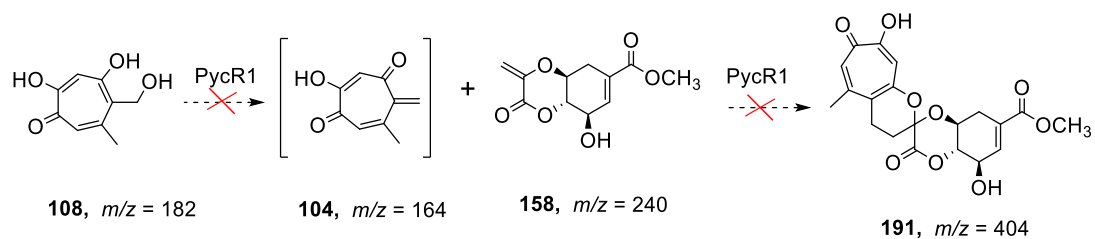
In addition, three dienophiles were tested separately for PycR1 *in vitro* assay. The experimental groups were set up as PycR1 + Eup2L4 + stipitaldehyde **102** + *2E*-humulene **71**, PycR1 + Eup2L4 + stipitaldehyde **102** + *2Z*-humulene **106**, PycR1 + Eup2L4 + stipitaldehyde **102** + scytolide **158**. All the *in vitro* assays were carried out under the same conditions as Eup2R1. Control assays were performed without PycR1 protein in the identical conditions. After the 2h reactions, the assays were stopped by adding MeCN (1:1), the protein was removed by centrifuge while the solution was analyzed by LCMS. The results indicated that there was no tropolone sesquiterpenoids formation in all the experimental groups compared with the corresponding control group (Figure 2.9.5.3-4). In addition, scytolide **158** was found to be unstable in the *in vitro* assay system, it reacted with H<sub>2</sub>O spontaneously to give two addition products compounds **193** and **194**, which were confirmed by the UV and MS data (Figure 2.9.5.5). Hence, PycR1 did not show hDAase activity in these *in vitro* assays.

In 2020, XenE, an hDAase from the BGC of xenovulene, was reported to show *in vitro* hDAase activity.<sup>159</sup> In this experiment, 2 mM stipitol **108** and 4 mM *2E*-humulene **71** were incubated with 0.4 mol % XenE in 50 mM citrate-Kpi buffer (pH 6.8) at 30 °C. After 18 h reaction, a monotropolone sesquiterpenoid dehydroxyepolone B **76** was synthesized by the hDA addition reaction.

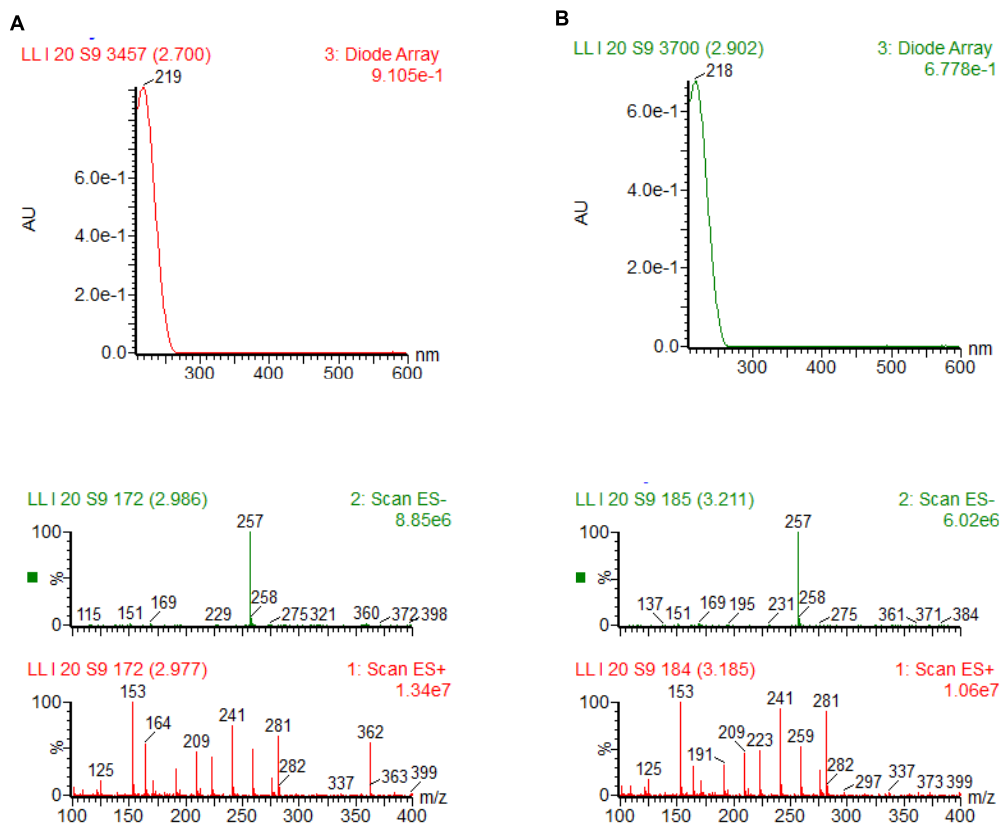


**Figure 2.9.5.3** LCMS analysis of PycR1 *in vitro* study with stipitol **108** and *2E*-humulene **71**/*ZZ*-humulene **106**.

Therefore, we decided to test the *in vitro* activity of PycR1 in citrate- $K_2HPO_4$  buffer with stipitol **108** and glycerol **187** as substances. The freshly prepared PycR1 (50  $\mu M$ ) and Eup2L4 (25  $\mu M$ ) were incubated with substrate stipitaldehyde **102** (2.5 mM), 1 % glycerol **187** and NADPH (3 mM) in 50 mM citrate- $K_2HPO_4$  buffer (pH 8) for 2 h at 26 °C. The reaction was quenched with MeCN (1:1) and the precipitated protein was removed *via* centrifugation. The supernatant was analyzed directly by LCMS. However, the LCMS chromatogram of *in vitro* assay of PycR1 in citrate- $K_2HPO_4$  buffer was the same as the one in bicine buffer, a new peak at  $t_R = 2.4$  min was detected (Figure 2.9.5.6). So, we stopped further investigation about the hDAase activity of PycR1 in citrate- $K_2HPO_4$  buffer.

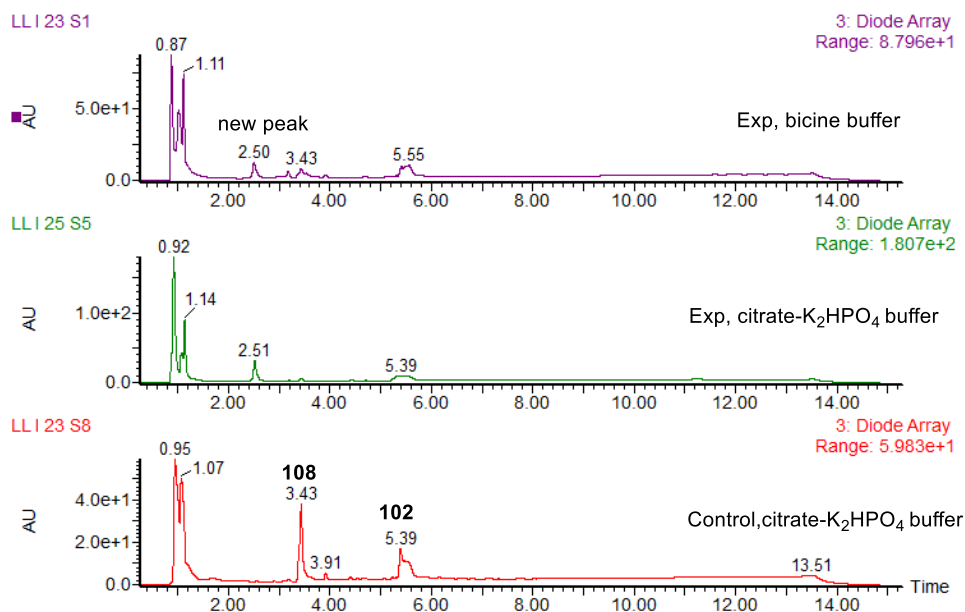


**Figure 2.9.5.4** LCMS analysis of PycR1 *in vitro* study with stipitol **108** and scytolide **158**: **A**, Extracted ion chromatogram 257 Da [M-H]<sup>-</sup>; **B**, Extracted ion chromatogram 259 Da [M+H]<sup>+</sup>; **C**, UV chromatogram of experiment group; **D**, UV chromatogram of control group.



**Figure 2.9.5.5** Characterisation of compound **193** and **194**: **A**, UV and MS spectrum of the peak at 2.6 min;

**B**, UV and MS spectrum of the peak at 3.0 min



**Figure 2.9.5.6** LCMS analysis of PycR1 *in vitro* study with stipitol **108** and glycerol **187** in citrate-K<sub>2</sub>HPO<sub>4</sub> buffer.



## 2.10 Conclusion and Discussion

In our hands, *CF236968* produces pycnidione **22** and epolone C **42** as the major compounds when cultured in PD agar for 14 days. Timecourse of metabolite production in *CF236968* indicates that three common diphenyl ethers including violaceol I **167**, cordyol C **168** and diorcinol **169** are observed at day 7, which are all decreasing intensely after the production of tropolone sesquiterpenoids. The biosynthetic pathway of diorcinol **169** has been elucidated, The BGC of diorcinol **169** is conserved in many fungal strains.<sup>160</sup> In contrast, strain *CF253087* produces eupenifeldin **23** and neosetophomone B **61** as the major compounds when cultured in PD agar for 14 days. However, the production of tropolone sesquiterpenoids is not observed in *Pleosporales sp 1987* when growing the fungi in 4 different media respectively.

Analysis of the ITS sequences suggests that *CF236968* belongs to *Leptobacillium sp.* while *CF253087* is placed in the *Phaeosphaeriaceae sp.* family. The phylogenetic analysis of *CF236968* and *CF253087* were not pushed forward, but further accurate investigations will be necessary for formally renaming the two organisms. The gDNA of *CF236968* was sequenced by using Illumina methods, which provides a draft genome of 29 Mb with an N50 of 0.5 Mb.

Bioinformatic searches for the homologous genes of *aspks1*, *asL1* and *asL3* (three tropolone biosynthetic genes) on the *CF236968* genome, enable the rapid identification of a putative BGC for pycnidione biosynthesis (the *pyc* BGC). Further searching for the homologous genes of the genes from *pyc* BGC in *CF150626* results in a putative BGC likely to be involved in eupenifeldin biosynthesis (the *eup2* BGC). Putative annotations of all the genes from the *pyc* BGC and *eup2* BGC are predicted by running protein BLAST in NCBI. They both contain three tropolone biosynthetic genes (*pycPKS*, *pycL1*, *pycL3/ eup2PKS*, *eup2L1* and *eup2L5*), a terpene cyclase gene (*pycR6/eup2R3*), an hDAase gene (*pycR1/eup2R1*), a FAD-binding monooxygenase gene (*pycR4/eup2R5*), a cytochrome P450 gene (*pycR5/eup2R6*) and a short-chain dehydrogenase gene (*pycL2/eup2L4*). Interestingly, the last three genes have no homologous gene in the BGC of xenovulene A **24**, which means that they are unique for the biosynthesis of pycnidione **22**/eupenifeldin **23**. So, it is really promising to perform further investigation on the functions of the enzymes encoded by *pycR4/eup2R5*, *pycR5/eup2R6* and *pycL2/eup2L4*.

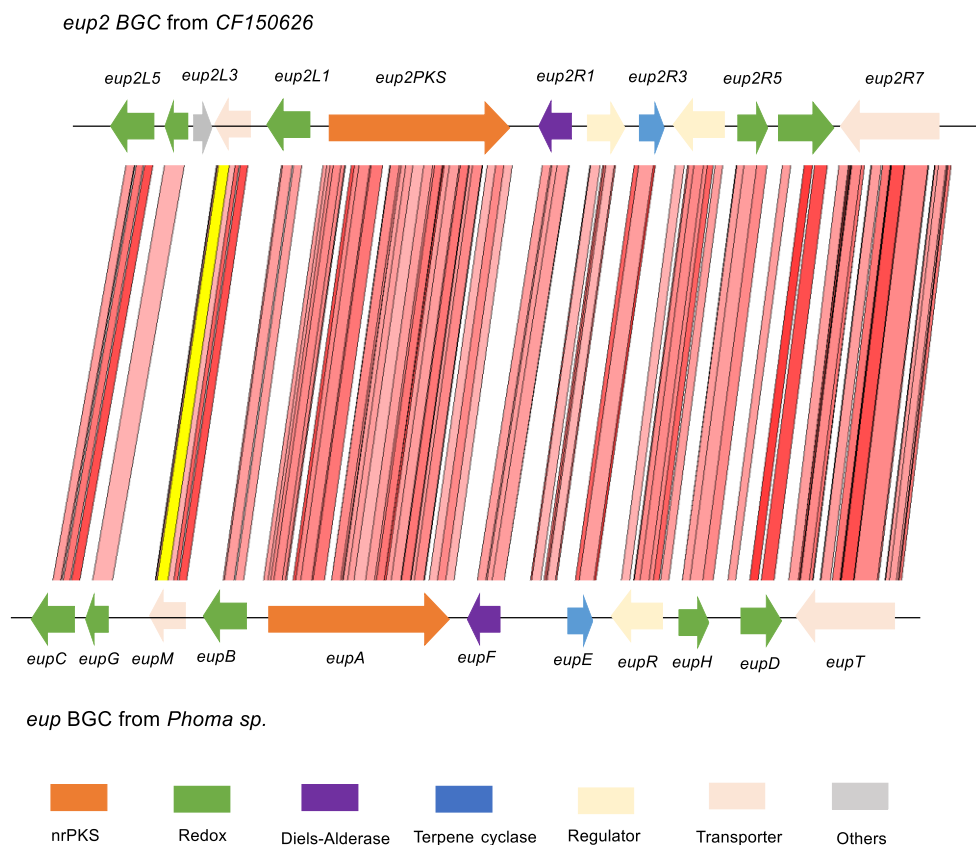
The expression levels of the genes from the *eup2* BGC were investigated using RT-PCR under both producing and non-producing conditions. It is worth noting that two tropolone biosynthetic genes (*eup2L1* and *eup2L5*), an hDAase gene (*eup2R1*), a terpene cyclase gene (*eup2R3*), and an FAD-binding monooxygenase gene (*eup2R5*) are only expressed under the producing condition, which support the conclusion that the *eup2* BGC is strongly linked to the production of eupenifeldin **23**. Besides, expression of the short-chain dehydrogenase encoded by *eup2L4* is observed under both producing and non-producing conditions, while the expression of a cytochrome P450 monooxygenase (P450) encoded by *eup2R6* is not observed under either producing condition or non-producing conditions. These data do not provide clear evidences for their participation in the biosynthetic pathway of eupenifeldin **23**.

In order to knock out genes from the *eup2* BGC, the transformation protocol for *CF150626* was explored. The best conditions for production of enough protoplasts were obtained by digesting young mycelium (2 days old) with *Trichoderma* lysing enzymes (10 mg/ml) and driselase from *Basidiomycetes* (5 mg/ml) in 1.1 M NaCl solution at 30 °C, 110 rpm for 3 h. Since the fungus is sensitive to hygromycin B, a pTH-GS-eGFP plasmid containing a *hph* gene and an *eGFP* gene was chosen as exogenous DNA for the transformation. The attempts for a PEG-mediated transformation protocol based on the *phoma* sp. produce a few promising transformants.

At the same time, the *eup* BGC from *Phoma* sp., the homologous BGC of *eup2* BGC (Figure 2.10.1), was identified to be the BGC of eupenifeldin **23** in the published literature.<sup>54</sup> Deletion of the genes that encoded the nr-PKS (*eupA/eup2PKS*), non-heme Fe (II)-dependent dioxygenase (*eupC/eup2L5*) or a ABC-transporter (*eupT/eup2R7*) respectively result in the abolition of eupenifeldin **23** production with no interesting intermediate, which of course provide no clear information about the function of the targeted genes.<sup>54</sup> Because of this published experimental evidence while our own studies were ongoing, we stopped further investigation of the transformation protocol for *CF150626*.

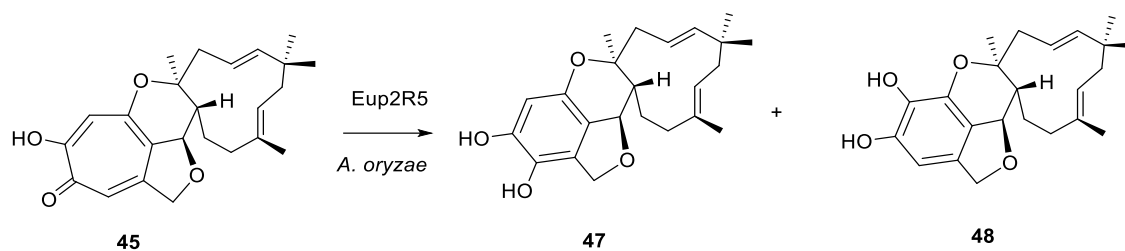
As for the origin of the monobenzopyranyl moiety in noreupenifeldin B **64**, there are two different proposals. The monobenzopyranyl moiety might derive from either orsellinaldehyde or the ring-contraction of a tropolone sesquiterpenoid (TS). AsL4 and AsL6 are FAD-dependent oxygenases previously shown to catalyze regioselective oxidative ring contractions of xenovulene B **45** to produce xenovulene A **24** (Scheme 2.8).<sup>20</sup> In order to check if noreupenifeldin B **64** is the direct ring-

contraction product of eupenifeldin **23**, *asL4* and *asL6* were expressed in *A. oryzae*. Next, eupenifeldin **23** was fed to *A. oryzae asL4+asL6* transformants cultured in DPY medium at day1. However, metabolite analysis at different days shows neither new product peaks nor the presence of eupenifeldin **23**, which indicates that eupenifeldin **23** is not stable in the DPY medium.



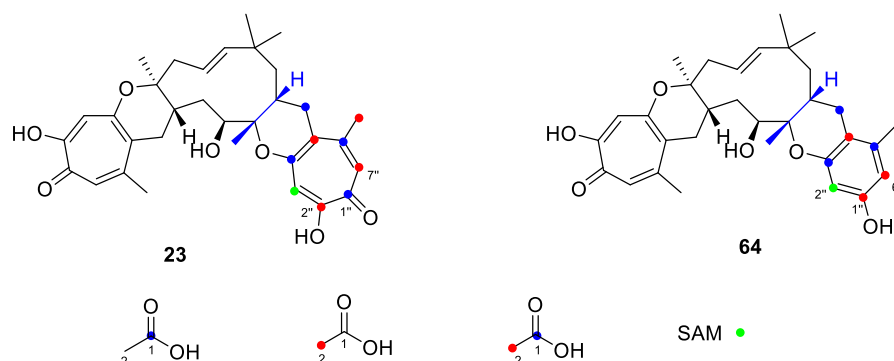
**Figure 2.10.1** Artemis analysis of the *eup2* BGC from *CF150626* and the *eup* BGC from *Phoma* sp..

In a parallel work, the homologous gene *eup2R5* was co-expressed with xenovulene B **45** producing genes in *A. oryzae* by Dr Carsten Schotte. Metabolite analysis of the transformants revealed the formation of compounds **47** and **48**, with 12 Da less than that of xenovulene B **45** (Scheme 2.10.1),<sup>35</sup> which indicates that the FAD-binding monooxygenase Eup2R5 shows *in vivo* ring-contraction activity and supports the proposal that noreupenifeldin B **64** is derived from a ring-contraction of eupenifeldin **23** catalyzed by Eup2R5.



**Scheme 2.10.1** The ring-contraction of of xenovulene B **45** catalyzed by Eup2R5.

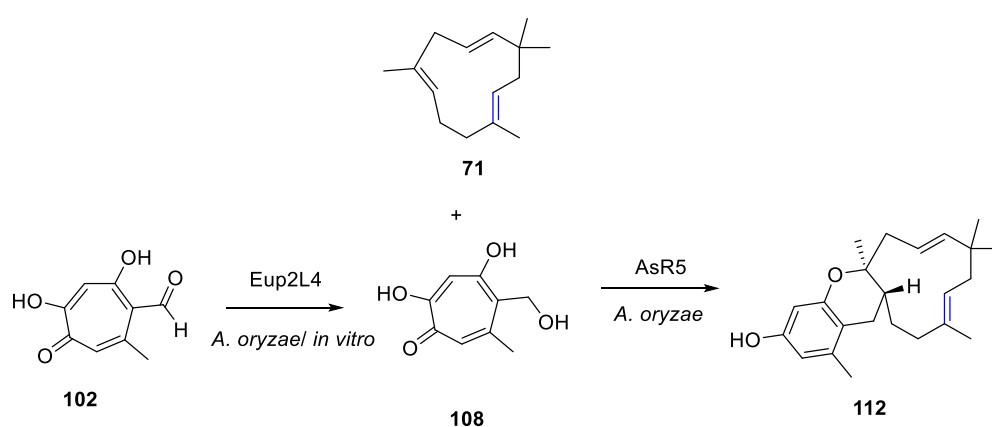
Additionally, isotope feeding experiments done by Dr Carsten Schotte with [1-<sup>13</sup>C]- and [2-<sup>13</sup>C]-labelled acetates in *CF150626* also support the conclusion that the monobenzopyranyl moiety is derived from the ring-contraction of a tropolone.<sup>35</sup> The observed pattern of label incorporation in the benzopyranyl moiety of noreupenifeldin B **64** shows that carbons C-6'' and C-1'' are both derived from [2-<sup>13</sup>C]-acetate instead of an intact acetate units [1-<sup>13</sup>C]- and [2-<sup>13</sup>C]-labelled acetates, and C-2'' is not labelled due to its origin from the methyl of methionine (Figure 2.10.2). These results are identical to the conclusion based on the Simpson's earlier isotopic labelling studies on xenovulene A **24** in *Acremonium strictum* (Scheme 1.2.5.1).<sup>161</sup>



**Figure 2.10.2** Isotopic labelling of eupenifeldin **23** and noreupenifeldin B **64** in *CF150626*.

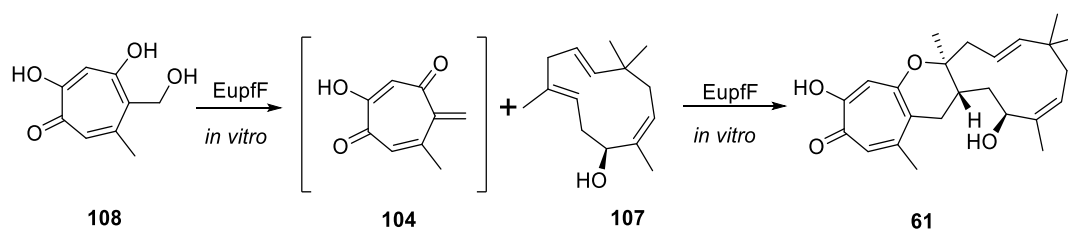
The aim of *in vitro* assay was to explore the function of enzymes encoded by *eup2L4*, *eup2R1* and *pycR1* and their role in the biosynthesis of eupenifeldin **23** and pycnidione **22**. Based on the bioinformatic analysis of the *pyc* BGC and the *eup2* BGC, we propose that short-chain dehydrogenase Eup2L4 catalyzes the reduction of stipitaldehyde **102**, and hDAase Eup2R1/PycR1 could take the product together with humulene as substrates to form the TS. Thus, these three enzymes were expressed in *E. coli BL21* for the *in vitro* assays. The substrate selectivity of the enzymes is also explored in this project.

According to our *in vitro* study on Eup2L4, the result suggests that it catalyzes the reduction of stipitaldehyde **102**. This is consistent with the *in vitro* assay results of EupfE performed by Hu and co-workers.<sup>55</sup> On the other hand, *eup2L4* is co-expressed with three tropolone biosynthetic genes (*aspks1*, *asL1* and *asL3*), a terpene cyclase gene (*asR6*), and an hDAase gene (*asR5*) from the BGC of xenovulene A **24** in *A. oryzae* by Dr Carsten Schotte. Metabolite analysis confirms the production of a new mono-TS compound **112** (scheme 2.10.2).<sup>35</sup> This result shows that Eup2L4 is responsible for the reduction of stipitaldehyde **102** during the biosynthesis of eupenifeldin **23**. In addition, Eup2L4 displays broad substrate tolerances towards different aromatic aldehydes.



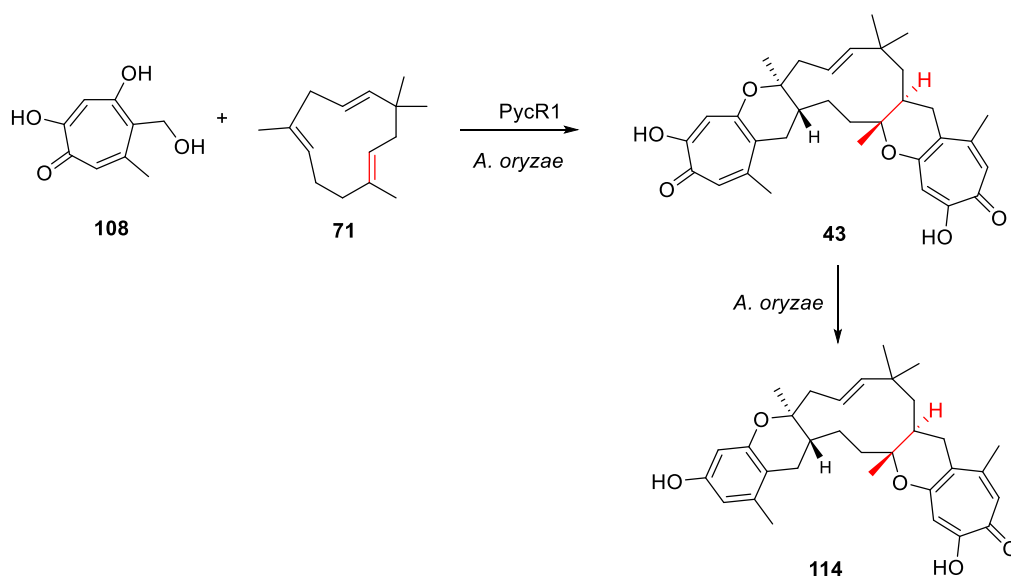
**Scheme 2.10.2** The reduction of stipitaldehyde **102** catalyzed by Eup2L4.

The *in vitro* assay of Eup2R1 shows no activity of dehydration or hDAase activity. Additionally, *eup2R1* is co-expressed with a terpene cyclase gene (*eup2R3*), a short-chain dehydrogenase gene (*eup2L4*) and stipitaldehyde producing genes from the BGC of xenovulene A **24** in *A. oryzae* by Dr Carsten Schotte. Metabolite analysis of transformants by LCMS shows no production of any tropolone sesquiterpenoids.<sup>35</sup> These results indicate that Eup2R1 exhibits neither *in vitro* nor *in vivo* activity, which is different from the observed *in vitro* assay results of the homologous protein EupfF performed by Hu and co-workers.<sup>55</sup> They infer that the hDAase EupfF catalyzes the dehydration of stipitol **108** to form active tropolone *o*-methide **104**, then catalyzes the hDA reaction between tropolone *o*-methide **104** and 2Z-humulene **107** to give neosetophomone B **61** (Scheme 2.10.3).



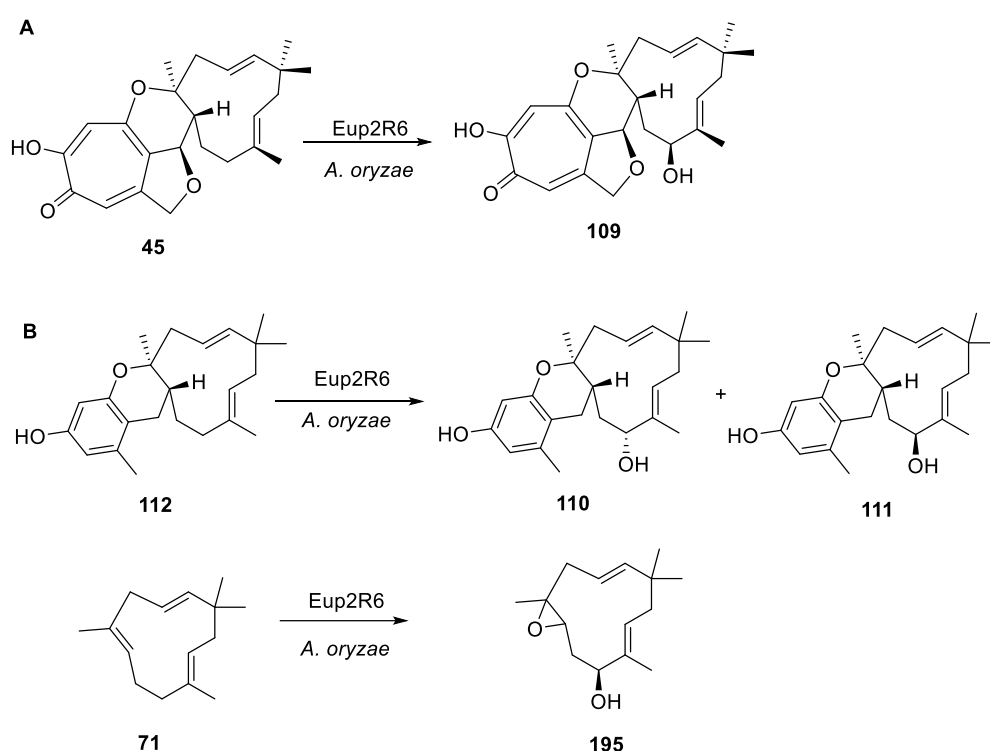
**Scheme 2.10.3** The dehydration and hDA reaction catalyzed by EupfF as observed by Hu and co-workers.<sup>55</sup>

The *in vitro* assay for PycR1 indicates that it uses stipitol **108** as its substrate to give a new peak **192** with the mass of 232. Due to the low concentration of the product, it is difficult to purify the compound **192** for NMR characterization. Unfortunately, PycR1 also shows no hDAase activity when incubated with stipitol **108** and three different dienophiles including *2E*-humulene **71**, *2Z*-humulene **106** and scytolide **158**. However, co-expression of *pycR1* with a terpene cyclase gene (*pycR6*), and the stipitaldehyde producing genes from the BGC of xenovulene A **24** in *A. oryzae* leads to the production of dehydroxypycnidione **43** and compound **114**. (Scheme 2.10.4).<sup>35</sup> This work confirms that PycR1 catalyzes the hDA addition between stipitol **108** and *2E*-humulene **71** during the biosynthesis of pycnidione **22**.



**Scheme 2.10.4** The hDA reaction catalyzed by PycR1.

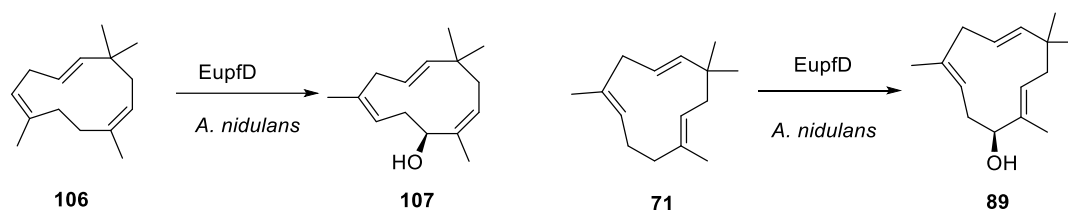
The function of Eup2R6, a cytochrome P450 monooxygenase, was investigated by Dr Carsten Schotte via heterologous expression in *A. oryzae*. Co-expression of *eup2R6* with the xenovulene B **45** producing genes shows the production of 10-hydroxyxenovulene B **109** with a 16 Da more than that of xenovulene B **45**. (Scheme 2.10.5A) Further inclusion of *eup2R6* in the expression system producing compound **112** results in the formations of three additional compounds.<sup>35</sup> Compounds **110** and **111** are confirmed to be the hydroxylated derivatives of **112** while compound **195** is assigned to be the 1,2-epoxy-10-hydroxy derivative of *2E*-humulene **71** and named phomanoxide B. (Scheme 2.10.5B)



**Scheme 2.10.5** The hydroxylation reactions catalyzed by Eup2R6.

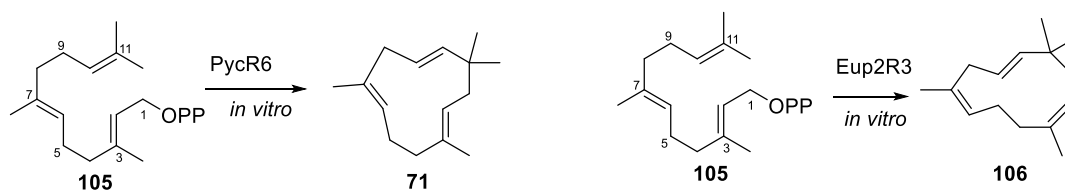
On the other hand, the homologous enzyme EupfD was identified to be responsible for the hydroxylation of both *2E*-humulene **71** and *2Z*-humulene **106** by Hu and co-workers.<sup>55</sup> Co-expression of *eupfD* and *eupfG* (a terpene cyclase gene) from the *eupf* BGC in *A. nidulans* leads to the production of *2Z*-humulenol **107**, a hydroxylated product of *2Z*-humulene **106**, while *2E*-humulenol **89** is obtained by feeding *2E*-humulene **71** to *A. nidulans* transformants expressing *eupfD* (Scheme 2.10.6).<sup>55</sup> All the studies of Eup2R6 and EupfD indicates that the cytochrome P450

enzymes catalyzes a necessary hydroxylation reaction during the biosynthesis of eupenifeldin **23**. And Eup2R6 displays broad substrate tolerances towards both humulenes and TS.



**Scheme 2.10.6** The hydroxylation reactions catalyzed by EupfD.

The function of Eup2R3, a putative humulene synthase, was investigated by Dr Carsten Schotte *via* the *in vitro* assay.<sup>62</sup> The result suggests that Eup2R3 is responsible for the formation of 2Z-humulene **106**, which is consistent with the results of hetero-expression of the putative humulene gene *eupfG* in *A. nidulans* performed by Hu and co-workers.<sup>55</sup> In addition, homologous gene *pycR6* from the *pyc* BGC was co-expressed with hDAase gene (*pycR1*), and stipitaldehyde producing gene from the BGC of xenovulene A **24** in *A. oryzae*, which yields two new di-TS, containing *syn*-relationship at the eastern dihydropyran ring junction (Scheme 2.10.4).<sup>35</sup> This study indicates that PycR6 is responsible for 2E-humulene **71** formation during the biosynthesis of pycnidione **22** while Eup2R3 catalyzes the production of 2Z-humulene **106** during the biosynthesis of eupenifeldin **23** (Scheme 2.10.7).

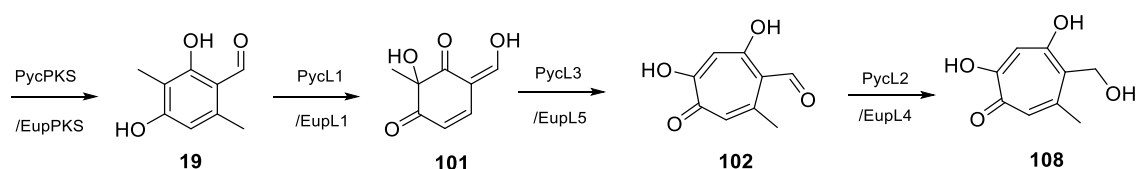


**Scheme 2.10.7** The terpene cyclization reactions catalyzed by PycR6 and Eup2R3.

Based on all the mentioned investigations on the functions of the biosynthetic proteins from the BGCs of pycnidione **22** and eupenifeldin **23**, we propose the biosynthetic pathways for both pycnidione-type TS and eupenifeldin-type TS, which share common early-stage steps to produce common tropolone precursor, stipitol **108** (Scheme 2.10.8). In the first step, 3-methylorcinolaldehyde **19** is synthesized from acetyl-CoA catalyzed by nr-PKS encoded by *pycPKS/eup2PKS*. The second



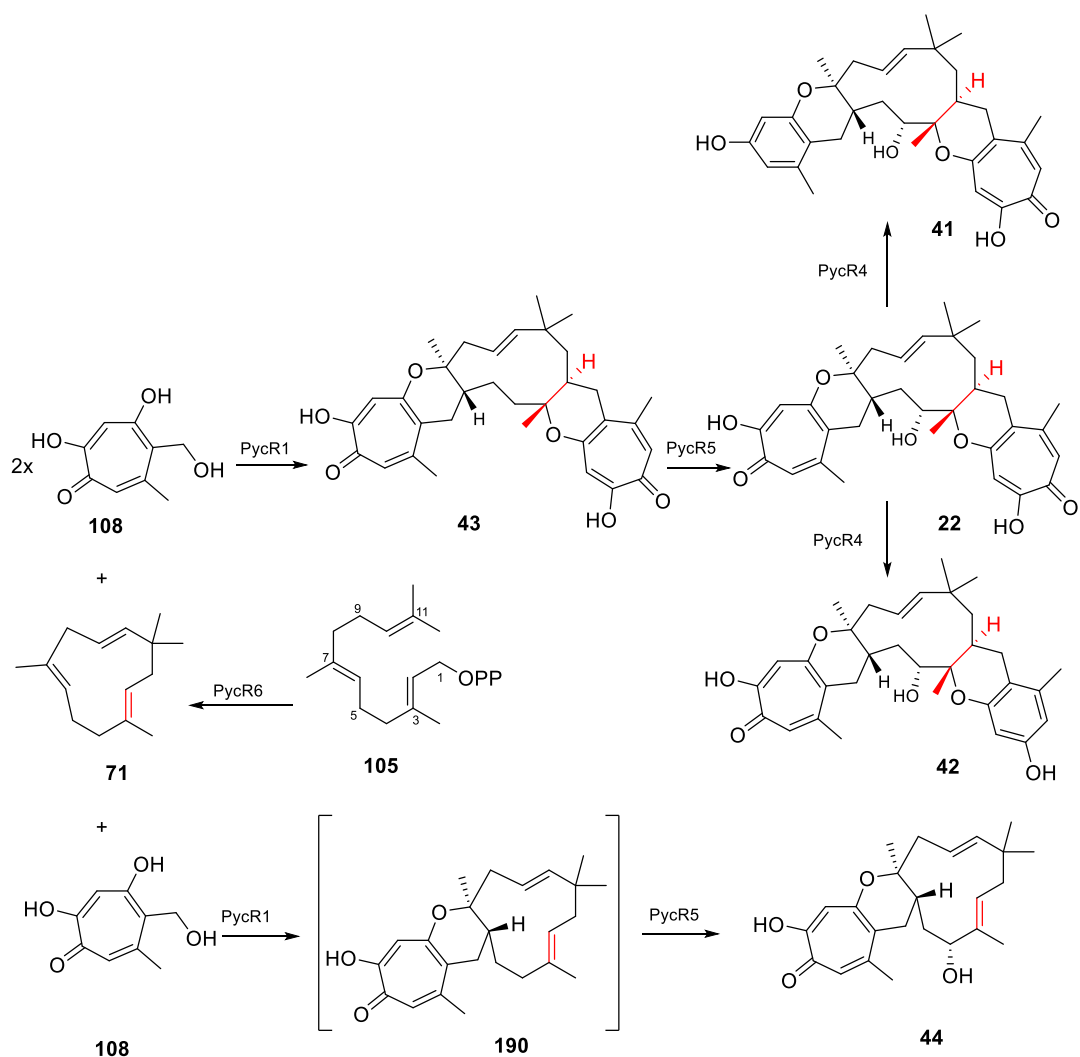
reaction is oxidative dearomatization of **19** catalyzed by the FAD dependent monooxygenase encoded by *pycL1/eup2L1*. The product **101** further undergoes the methyl-hydroxylation and ring expansion catalyzed by non-haem iron II oxygenase (encoded by *pycL3/ eup2L5*) to produce stipitaldehyde **102**. It is then reduced to stipitol **108** by the short-chain dehydrogenase (encoded by *pycL2/eup2L4*)).



**Scheme 2.10.8** Early-stage chemical steps of the biosynthetic pathway for eupenifeldin **23** and pycnidione **22**.

The biosynthesis of pycnidione **22** and eupenifeldin **23** diverges after the formation of terpene precursor catalyzed by the humulene synthases. The terpene cyclase encoded by *pycR6* converts FPP **105** into *2E*-humulene **71**, which leads to the biosynthesis of the pycnidione series TS (Scheme 2.10.9). In the next biosynthetic process, the hDA reaction between stipitol **108** and *2E*-humulene **71** is catalyzed by hDAase PycR1 to produce dehydroxypycnidione **43**, a natural co-metabolites of pycnidione **22**.<sup>34</sup> Hydroxylation of dehydroxypycnidione **43** at C-10 then forms pycnidione **22**, which is catalyzed by a cytochrome P450 PycR5. Formation of the epolone A **41** and epolone C **42** involve a ring contraction on the tropolone moiety of pycnidione **22** catalyzed by FAD-binding monooxygenase encoded by *pycR4*.

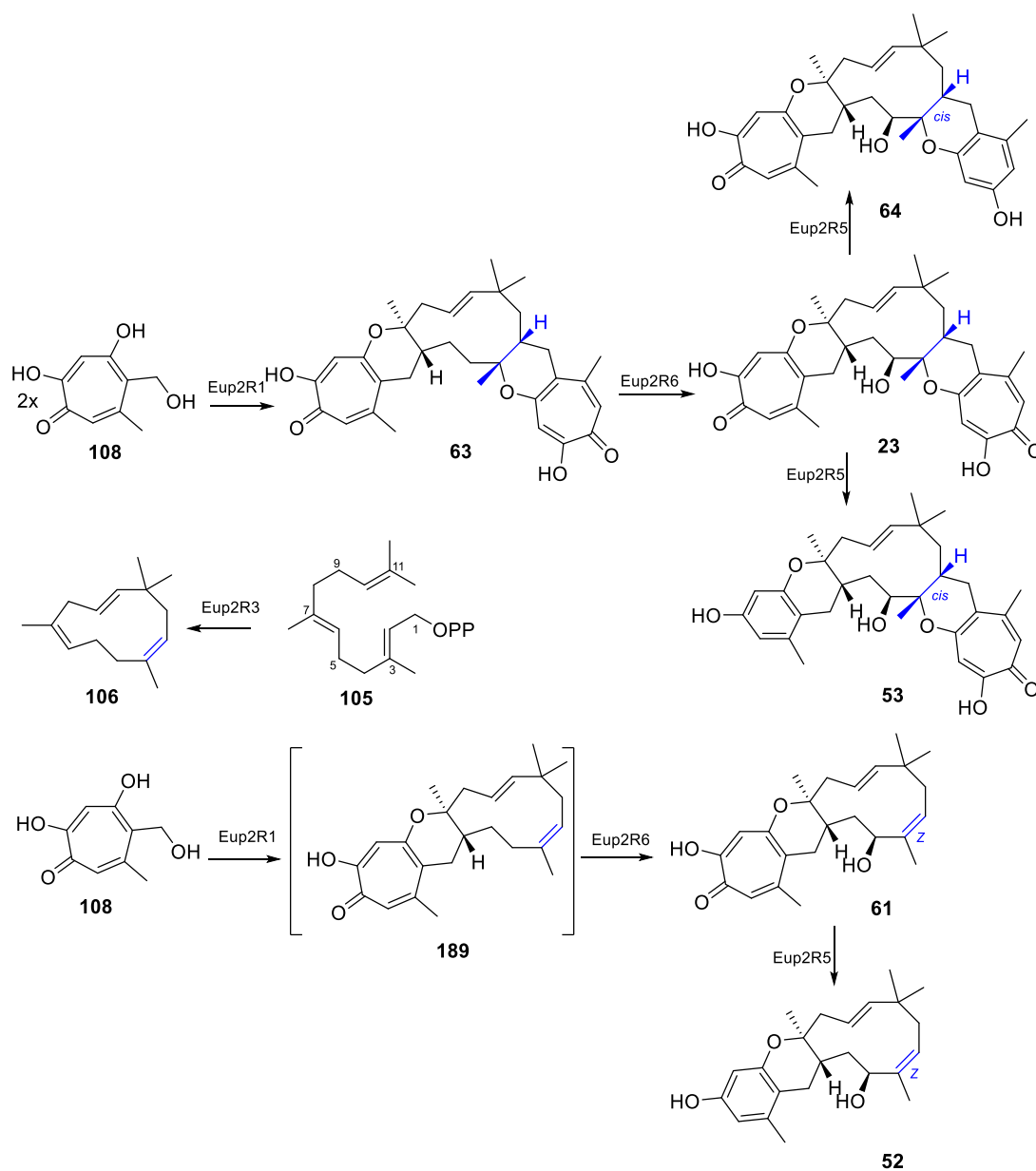
While the terpene cyclase Eup2R3 converts FPP **105** into *2Z*-humulene **106** that serves as the precursor of eupenifeldin series TS (Scheme 2.10.10). In the next biosynthetic process, the hDA reaction between stipitol **108** and *2Z*-humulene **106** is catalyzed by hDAase Eup2R1 to produce dehydroxyeupenifeldin **63**.<sup>35</sup> Hydroxylation of dehydroxyeupenifeldin **63** at C-4 then forms eupenifeldin **23**, which is catalyzed by a cytochrome P450 Eup2R6. Formation of noreupenifeldin **53** and noreupenifeldin B **64** involve a ring contraction on the tropolone moiety of eupenifeldin **23** catalyzed by the FAD-binding monooxygenase Eup2R5.



**Scheme 2.10.9** Proposed late-stage chemical steps of the biosynthetic pathway for pycnidione-type TS

Based on the proposed biosynthetic pathway, we conclude that pycnidione-type TS are derived from 2E-humulene **71** and the 2E-alkene configuration is mirrored in the final pathway product (2E-alkene in mono-TS and *trans*-fusion at the C-2/C-3 ring junction in di-TS). Eupenifeldin-type TS are derived from 2Z-humulene **106**, so they display a 2Z-alkene configuration in mono-TS and *cis* fusion at the C-2/C-3 ring junction in di-TS. Of special interest hereby is that we propose that it is the different stereoselectivity of the humulene synthases (Eup2R3 and PycR6) that sets the stereochemistry of the later-formed pycnidione-type TS and eupenifeldin-type TS. Additionally, 4R hydroxyl group in pycnidione-type TS and 4S hydroxyl group in eupenifeldin-type TS are determined by different stereoselectivity of the cytochrome P450 monooxygenase (Eup2R6 and PycR5). A few natural mono-TS, such as neosetophomone B **61** and epolone B **44**,<sup>33,41</sup> have been

isolated as the co-metabolites of eupenifeldin **23** and pycnidione **22**, indicating that the hDAases (Eup2R1 and Pyc1) are probably able to produce both mono- and di-tropolone sesquiterpenoids.



**Scheme 2.10.10** Proposed late-stage chemical steps of the biosynthetic pathway for eupenifeldin-type TS.

Taken together, the *pyc* BGC for the biosynthesis of the pycnidione series of TS and the *eup* BGC for the biosynthesis of the eupenifeldin series of TS are identified. Even beyond that, five functional proteins, including a terpene cyclase (PycR6/Eup2R3), an hDAase (Pyc1/Eup2R1), a FAD-binding monooxygenase (PycR4/Eup2R5), a cytochrome P450 (PycR5/Eup2R6) and a short-chain dehydrogenase (PycL2/Eup2L4) were found to be necessary for the biosynthesis of TS based on the

*in vitro* assay and hetero-expression in *A. oryzae*. Additionally, new and detailed biosynthetic routes for both pycnidione-type TS and eupenifeldin-type TS are proposed based on the identified gene functions from the *pyc* BGC and the *eup* BGC respectively. However, the chemical steps and the intermediates in the biosynthetic pathway for both eupenifeldin **23** and pycnidione **22** are still unclear. In the future, biosynthesis of pycnidione **22** could be investigated by co-expression of *pycR5* and dehydroxypycnidione producing genes in *A. oryzae*. While it is also interesting to check if the hDAase encoded by *pycR1* could take 2Z-humulene **106** as substrate by replacing *pycR6* with *eup2R3* in *A. oryzae* transformants that producing dehydroxypycnidione **43**. The stereoselectivity of the PycR5 and Eup2R6 (cytochrome P450) also deserves further investigation.

### 3 Biosynthesis of scytolide in fungi

Scytolide is a secondary metabolite of produced by *Scytalidium uredinicola*, *Phyllosticta cirsii* and *Trichoderma reesei*.<sup>99,100</sup> It is a natural herbicide and is already identified to be a key precursor of spiroorbicillinols in the biosynthetic pathway of sorbicillinoids by the Cox group (chapter 1.4).<sup>102</sup> Based on its structural similarity to EPSP and shikimate, we propose that scytolide is derived from the shikimate pathway, which is a primary metabolic pathway by which microorganisms and plants produce the aromatic amino acids and a number of other aromatic compounds which are essential for sustaining the basic functions of living organisms. Besides, a vast number of secondary metabolites of plants and microorganisms are also from this pathway and a majority of them are derived from chorismate, the direct product of the shikimate pathway. However, there are also some natural products that are derived from the intermediates in the shikimate pathway, which exists more frequently in microorganisms (chapter 1.3). Up to now, no experimental research about the chemical synthesis or biosynthesis of scytolide has been reported.

#### 3.1 Project aims

Although scytolide **158** was reported from several different fungi and proposed to derive from the shikimic acid pathway, nothing is known about its biosynthetic pathway. One of the major aims is to identify the biosynthetic gene cluster and clarifying the chemical reactions involved in scytolide biosynthesis.

For this reason, the secondary metabolite profile of two known scytolide producers *P. cirsii* and *T. reesei* *Atmus53* will be analyzed in detail first. The methods to understand the biosynthesis involve traditional isotope feeding experiments for identifying the biosynthetic origin of the carbon atoms, and modern genome mining methods aimed at finding the putative BGCs. Finally, molecular biology tools will allow reverse transcription PCR, knockout experiments, and heterologous expression, to characterize the role and function of the genes and chemical reactions catalyzed by their respective encoded proteins.

Feeding <sup>13</sup>C-labeled precursors will reveal the origin of the building blocks needed in the biosynthesis. The building blocks of scytolide might be phosphoenol pyruvate, which is formed from

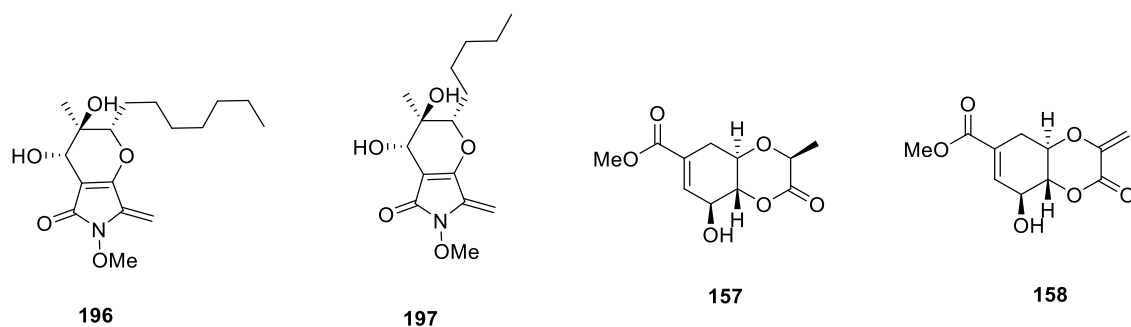
glycerol via the glycolytic pathway. For this reason, we plan to feed [2- <sup>13</sup>C]-glycerol **187** to test if the C-3 and C-5 atom of scytolide will be enriched. Bioinformatics analysis of the total genome sequences from *P. cirsii* and *T. reesei*  $\Delta$ tmus53 will be performed to mine the putative BGC and functional genes linked to scytolide production. In the research of the core enzymes involved in the biosynthesis, special regard is held for the shikimate pathway related enzymes. And the likely presence of a SAM-dependent *O*-methyltransferase within the cluster will also be indicative to justify the presence of the methyl ester.

RT-PCR allows rapid characterisation of the putative BGC and functional genes that involves scytolide biosynthesis without requiring genetic modification. However, it does not provide conclusive evidence of a causal link between gene expression and metabolite production. The putative BGC will be further confirmed by knock-out experiments. Disruption of targeted genes from the BGC will prevent scytolide biosynthesis and lead to accumulation of intermediates. The results will help us understand how scytolide is formed and disclose the chemical reactions catalyzed by potentially functional enzymes. In addition, heterologous expression experiments in *A. oryzae* will be deployed to examine the function of key enzymes in more detail.

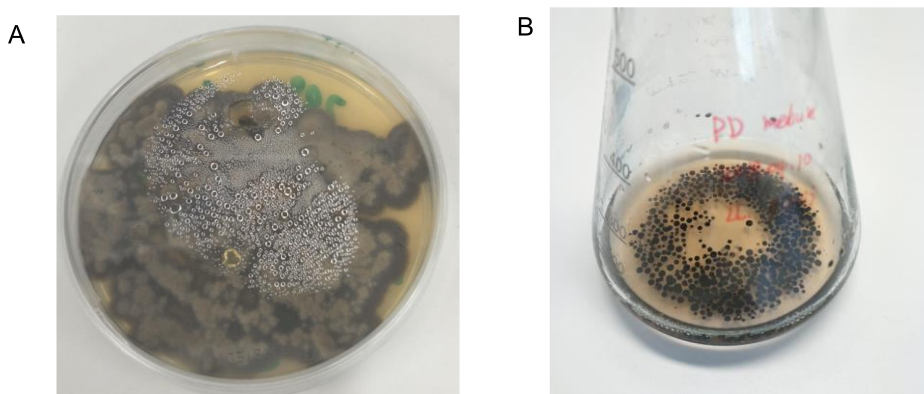
## **3.2 Results-Analysis of Scytolide 158 Produced by *P. cirsii* and *T. reesei* $\Delta$ tmus53**

### **3.2.1 Analysis of Secondary Metabolites from *P. cirsii***

The fungus *Phyllosticta cirsii* was obtained from Dr. Francesco Trenti. He received the organism as a 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). *P. cirsii* was isolated from infected Canada thistle around 20 years ago and evaluated as a biocontrol agent of this noxious perennial weed. It has been shown to produce different phytotoxic metabolites with potential herbicidal activity when grown in liquid cultures. Different classes of metabolites including phyllostictine A **196**, phyllostictine B **197**, phyllostin **157** and scytolide **158** have been isolated and identified chemically from this fungus (Figure 3.2.1.1).<sup>162,163,100</sup>

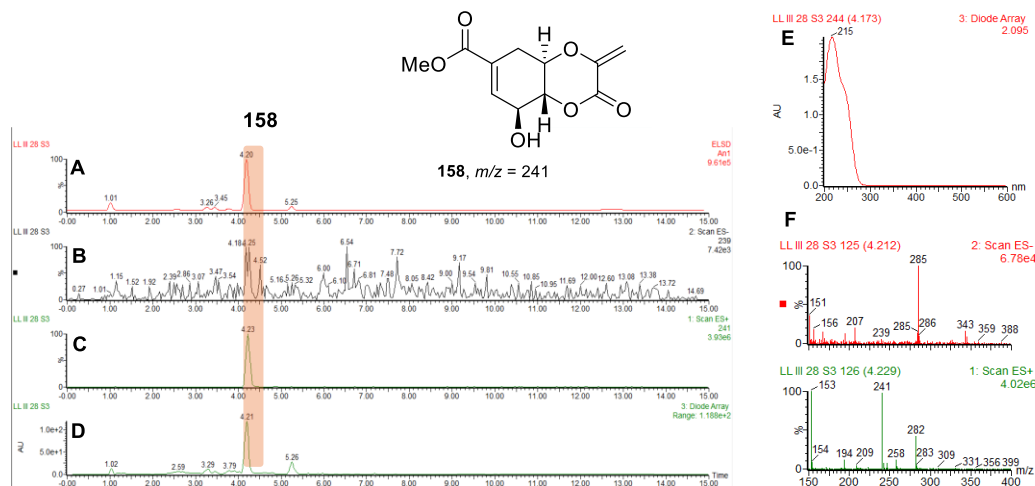


**Figure 3.2.1.1** Representative natural products from *P. cirsii*



**Figure 3.2.1.2** Morphologies of *P. cirsii*: **A**, *P. cirsii* growing on PD agar; **B**, *P. cirsii* growing in liquid PD medium.

In our hands, a different chemical profile was obtained using PD medium (Figure 3.2.1.3 A). Fermentation at 28 °C and 110 rpm for 7 days was suitable for the production of scytolide **158**. The fungal cultures without mycelium were extracted with the same volume of ethyl acetate two times. The organic layer was dried with MgSO<sub>4</sub> and rotary evaporated to give crude metabolites. The crude extract was dissolved in HPLC grade acetonitrile to a concentration of 10 mg/ml and then submitted for LCMS analysis. The HPLC chromatogram of wild type *P. cirsii* showed a single dominant peak at  $t_R = 4.2$  min with UV absorption  $\lambda_{max}$  215 nm (Figure 3.2.1.3 E). Extracted ion chromatograms using 241 Da (ES<sup>+</sup>) led to the identification of scytolide **158** (Figure 3.2.1.3 F). However, compound **158** was known to have [M + Na]<sup>+</sup> molecular ion 263 Da in the literature.<sup>100</sup>

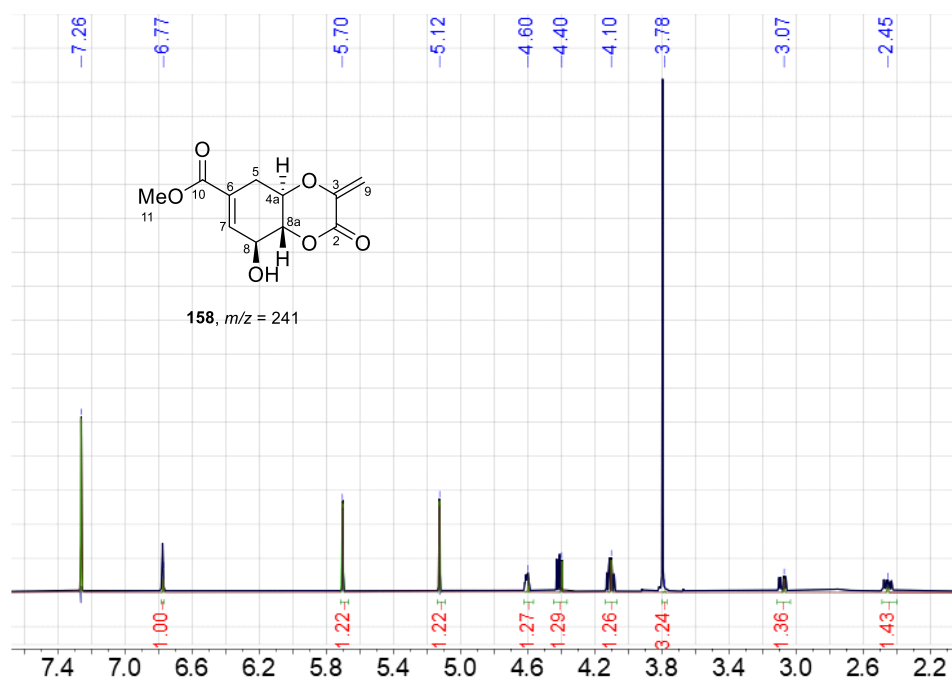


**Figure 3.2.1.3** The LCMS chromatogram of *P. cirsii* culture extracts after 7 days' growth in PD medium: **A**, ELSD chromatogram of *P. cirsii* culture extracts; **B**, Extracted ion chromatograms  $[M + H]^+$ ; **C**, Extracted ion chromatograms  $[M - H]^-$ ; **D**, UV chromatogram of *P. cirsii* culture extracts; **E**, UV spectrum of **158**; **F**, Mass spectrum of **158**.

### 3.2.2 Structural Identification of Scytolide **158**

To confirm the assignment of the compound eluting at 4.2 min as scytolide **158**, the peak was targeted for purification and NMR analysis. A total of 1 L of *P. cirsii* cultures were grown in PD medium for 7 days and the mycelium were then separated by vacuum filtration. The supernatants were combined together and then extracted with ethyl acetate 2 times. The organic extract was dried over anhydrous  $MgSO_4$  and evaporated to afford 500 mg crude extract. In total, 10 mg scytolide **158** was purified by mass-directed preparative HPLC for  $^1H$  and  $^{13}C$  NMR analysis. The data were identical to the literature data of **158** (Table 3.2.2). In the literature, the planar structure and relative configuration of scytolide **158** were determined by X-ray analysis, while the absolute configuration of scytolide **158** was assigned to be 4a*R*,8*S*,8a*R* by combined analysis of ECD and VCD data.<sup>99,100</sup> It is worth noting that the structure assigned to **158** was first confirmed by extracted ion chromatograms using 241 Da ( $[M + H]^+$ ) instead of 263 Da ( $[M + Na]^+$ ) in the  $ES^+$  spectrum.





**Figure 3.2.2**  $^1\text{H}$  NMR spectrum of scytolide **158**.

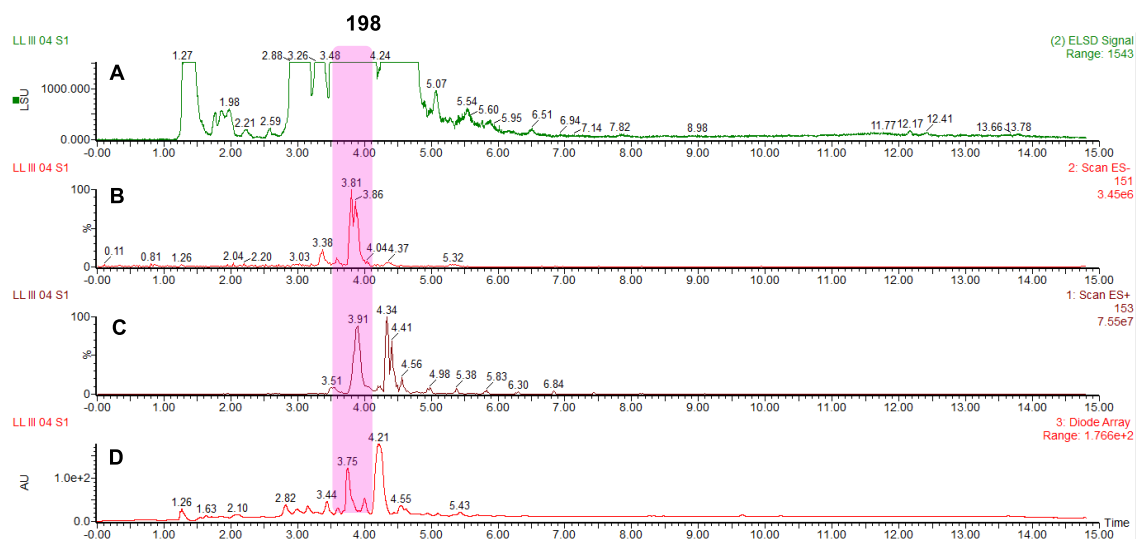
Position	$\delta_{\text{C}}/\text{ppm}$ <b>158</b>	$\delta_{\text{H}}/\text{ppm}$ <b>158</b>	$\delta_{\text{C}}/\text{ppm}$ <b>158</b> literature	$\delta_{\text{H}}/\text{ppm}$ <b>158</b> literature
2	159.2		160.0	
3	146.2		146.0	
4a	70.2	4.07 td ( $J = 10.1, 6.5$ Hz)	70.2	4.07 ddd ( $J = 10.0, 8.0, 6.6$ Hz)
5	29.7	3.04 dddd, ( $J = 17.7, 6.5, 1.8, 0.9$ Hz) 2.41 dddd, ( $J = 17.8, 9.9, 4.0, 2.9$ Hz)	29.6	3.04 dd ( $J = 17.0$ and $6.6$ Hz) 2.42 m
6	128.1		127.4	
7	137.0	6.73 m	139.4	6.76 br s
8	70.2	4.56 ddt ( $J = 8.0, 4.0, 2.0$ Hz)	69.2	4.51 m
8a	82.5	4.37 dd, ( $J = 10.3, 8.0$ Hz)	82.7	4.37 dd ( $J = 10.0, 8.0$ Hz)
9	105.3	5.66 d ( $J = 1.6$ Hz) 5.09 d ( $J = 1.6$ Hz)	105.0	5.67 d ( $J = 0.7$ Hz) 5.10 d ( $J = 0.7$ Hz)
10	165.7		165.7	
11	52.5	3.76 s	52.3	3.76 s

**Table 3.2.2** NMR data for scytolide **158** in  $\text{CDCl}_3$  (600 MHz); Literature data was also measured in  $\text{CDCl}_3$ .

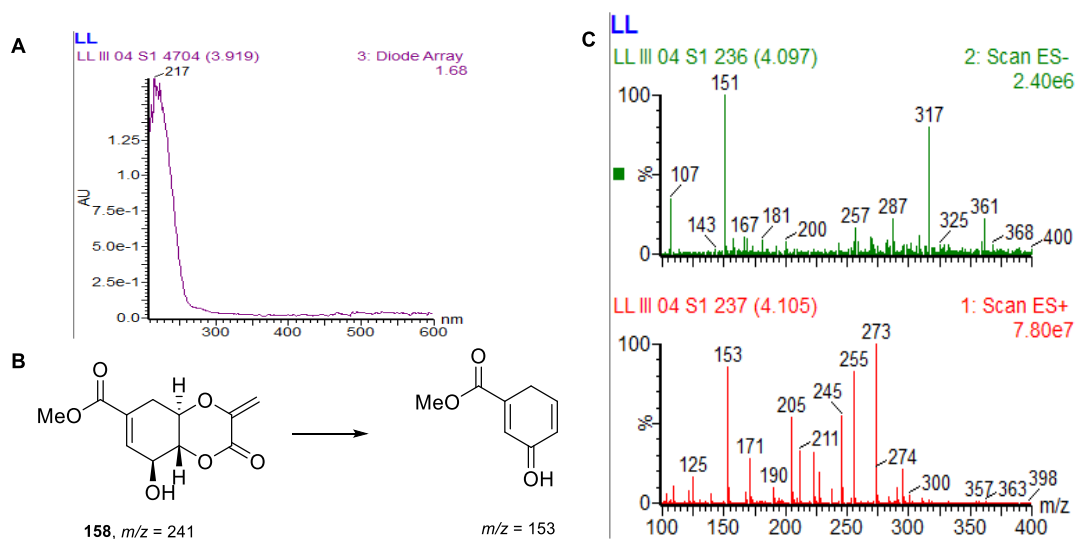
### 3.2.3 Structural Identification of Methylscytolide **198**

A new minor peak was observed at the retention time of 3.7 min after storing the crude metabolites of *P. cirsii* in methanol for 3 days (Figure 3.2.3.1). It gave UV absorption  $\lambda_{\text{max}}$  217 nm that are similar to scytolide **158** (Figure 3.2.3.2 A). The mass spectra of the new peak show some similarity with that of scytolide **158** in that they both contain the 153 Da ion signal in the  $\text{ES}^+$  spectrum and 151 Da ion signal in the  $\text{ES}^-$  spectrum (Figure 3.2.3.2 B). These chemical characteristics did not appear to correspond to a known *P. cirsii* metabolite. So, the new peak was proposed to be a new

compound similar to scytolide **158**, which was then purified by preparative-LCMS using a reverse phase column (0 - 7 min: 10 % - 57 %; 7 - 10 min: 57 %). 5 mg of this new compound **198** was isolated, which enabled full 1D and 2D NMR spectrum.



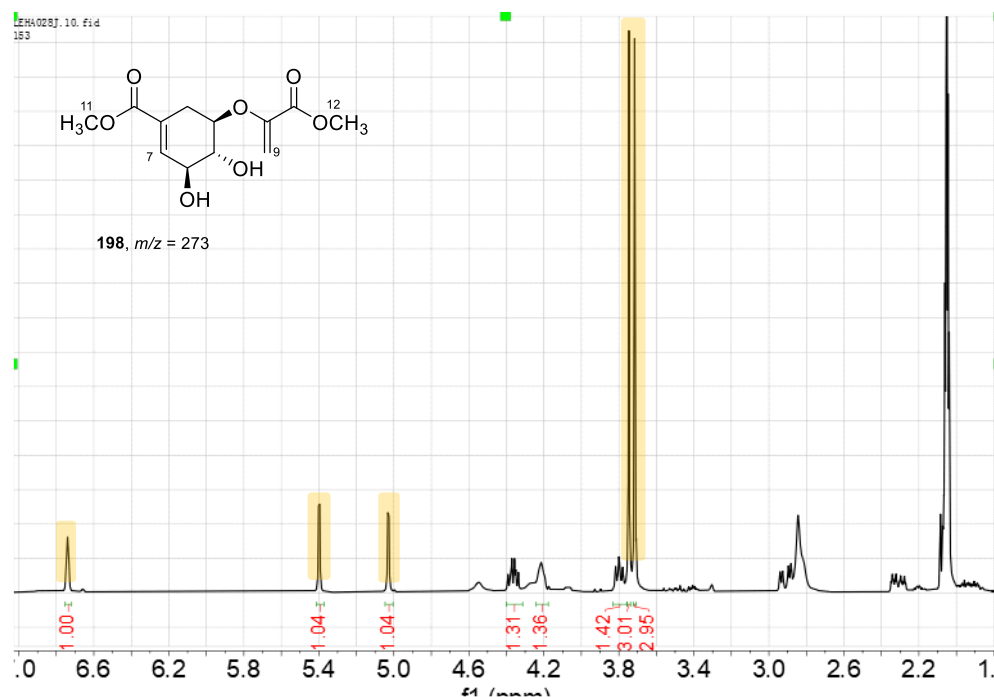
**Figure 3.2.3.1** The LCMS chromatogram of the crude metabolites of *P. cirsii* in methanol for 3 days: **A**, ELSD chromatogram of the crude metabolites; **B**, Extracted ion chromatograms  $[M+H]^+$ ; **C**, Extracted ion chromatograms  $[M-H]^-$ ; **D**, UV chromatogram of the crude metabolites.



**Figure 3.2.3.2** Characterisation of the new peak eluting at 3.7 min: **A**, UV spectrum of the new peak; **B**, Putative structural basis for 153 Da in the ES<sup>+</sup> spectrum and 151 Da in the ES<sup>-</sup> spectrum; **C**, Mass spectrum of the new peak **198**.

The <sup>1</sup>H NMR spectrum of **198** revealed the presence of two singlet *O*-methyl [ $\delta_H$  3.72 (s, H<sub>3</sub>-11), 3.75 (s, H<sub>3</sub>-12)] and three olefinic protons [ $\delta_H$  6.74 (td, H-7), 5.40 (d, H-9), 5.03 (d, H-9)] (figure 3.2.3.3). The similar characteristic signals were also observed in the <sup>1</sup>H NMR spectrum of scytolide

**158.** The  $^{13}\text{C}$  NMR spectra of **198** showed 12 carbon resonances including two *O*-methyls, one methylene, three *O*-methines and six quaternary carbons (Table 3.2.3). These data suggested that **198** might contain one more *O*-methyl group than scytolide **158**.

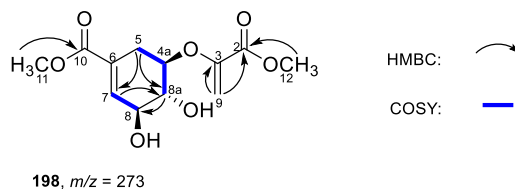


**Figure 3.2.3.3**  $^1\text{H}$  NMR spectrum of new peak; The most identifiable signal was highlighted in color.

HMBC correlations from H-12 to C-2 and from H-9 to C-2 and C-3 indicated that one *O*-methyl was connected to C-2, while the second *O*-methyl was determined to be attached to C-10 based on HMBC correlations from H-11 to C-10. The large carbon chemical shift difference ( $>5$  ppm) between **198** and scytolide **158** at C-4a, C-8a, C-2, C-3, C-9 supported the cleavage of 6-member lactone ring. The  $^1\text{H}$ - $^1\text{H}$  COSY correlation of H-7/H-8, H-4a/H-5 and H-4a/H-8a, together with HMBC correlations from H-5 to C-4a, C-8a and C-7, H-4a and H-7 to C-8a and H-8a to C-8 and C-4a established the core 6-membered ring that are similar to scytolide **158**, which was also supported by the similar carbon chemical shift between **198** and scytolide **158** at C-5, C-6, C-7, C-8, C-10.

The relative and absolute stereochemistry at positions C-4a, C-8a and C-8 were assumed to be identical to scytolide, but not further determined. Collectively, the structure of **198** was definitely assigned with the molecular formula of  $\text{C}_{12}\text{H}_{16}\text{O}_7$ , the corresponding  $[\text{M} + \text{H}]^+$  molecular ion signal 273 was also observed in  $\text{ES}^+$  spectrum. So, the major compound eluting at 3.8 min was shown to be

a methylation product of scytolide **158**, which was then named as methylscytolide **198**. It was previously reported as a key intermediate in the chemical synthesis of 5-enolpyruvylshikimate 3-phosphate, but no spectroscopic (NMR or UV) characterisation was available.<sup>164</sup>



Position	$\delta_C$ /ppm <b>198</b>	$\delta_H$ /ppm <b>198</b>	COSY	HMBC
2	164.3			H-9, H-12
3	151.0			H-9
4a	76.8	4.36, td, $J = 8.3, 5.5$ Hz	H-5, H-8a	H-5, H-8a
5	28.7	2.89, m 2.31, ddt, $J = 17.7, 8.0, 2.7$ Hz		
6	127.6			
7	140.0	6.74, td, $J = 2.6, 1.2$ Hz	H-8	H-5
8	71.9	4.22, brs		H-8a
8a	74.6	3.80, dd, $J = 8.5, 6.6$ Hz		H-4a, H-7, H-5
9	97.0	5.40, d, $J = 2.6$ Hz 5.03, d, $J = 2.6$ Hz		
10	167.0			H-11
11	52.1	3.72, s		
12	52.5	3.75, s		

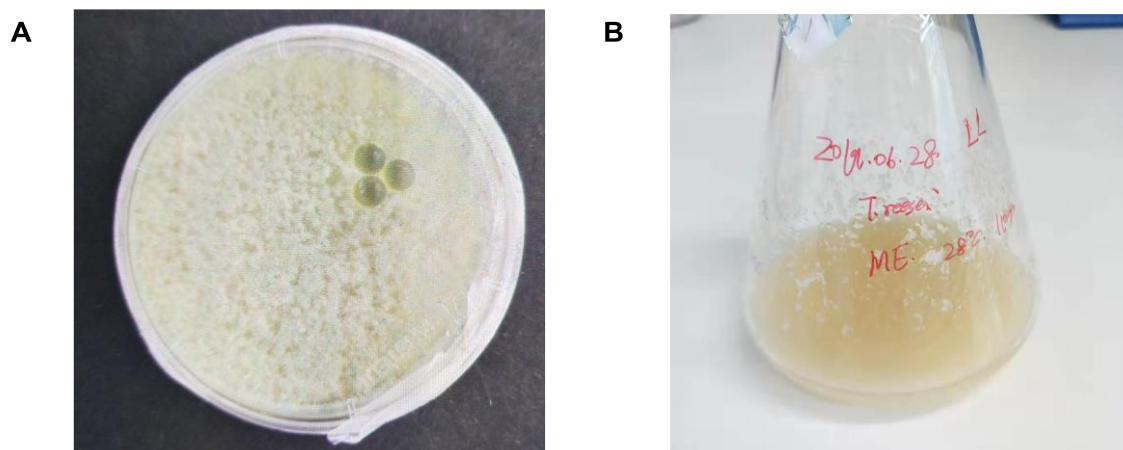
**Table 3.2.3** NMR data for methylscytolide **198** in acetone- $d_6$  (500 MHz).

### 3.2.4 Analysis of Secondary Metabolites from *T. reesei* $\Delta$ tmus53

*Trichoderma reesei* is well known for efficient production of cellulases, and it secretes a typical yellow pigment during growth. The yellow pigment is not favorable for industrial enzyme production because it makes the downstream process more complicated and thus increases operating costs. Recently, these yellow pigments were identified to be a diverse group of sorbicillinoids, including spirosorbicillinols A **160** and B **161**.<sup>165</sup> In 2017, a BGC in *T. reesei* which is responsible for the biosynthesis of sorbicillinoids was reported. This BGC contains two PKS gene, *sorA* and *sorB*. Knocking out the two polyketide synthases in the BGC, both individually and simultaneously, abolished the production of sorbicillinoids in *T. reesei*.<sup>166</sup>

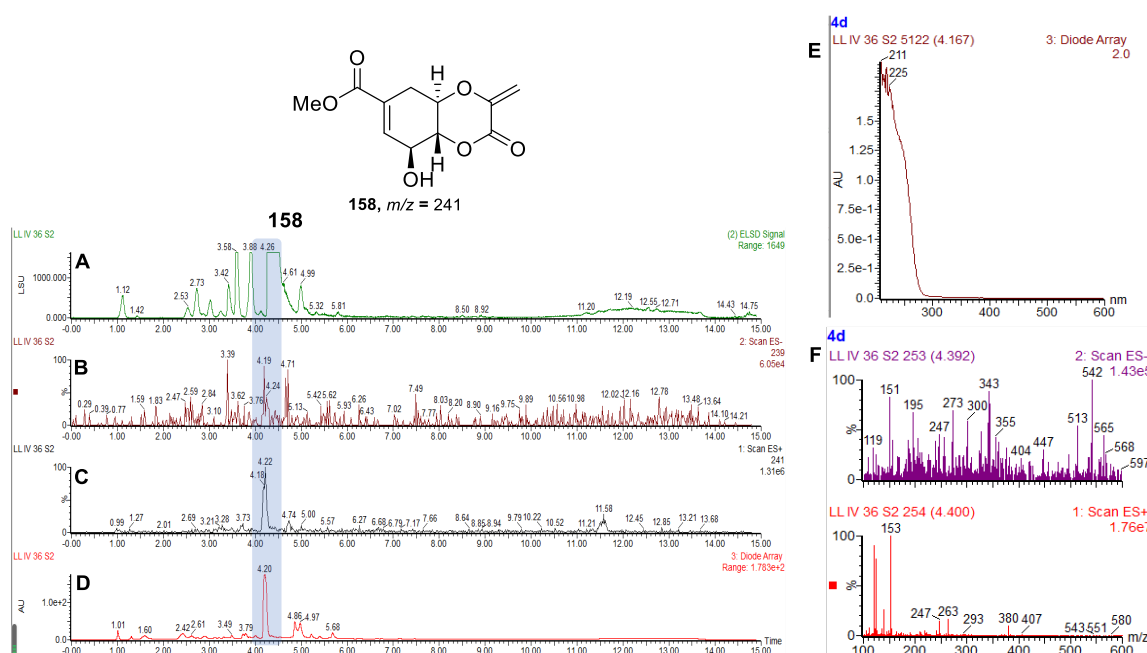
Isolation of spirosorbicillinols from *Trichoderma* species was reported previously and its biosynthesis was proposed to be a result of DA reaction between sorbicillinol **159** and scytolide **158**.<sup>167</sup> Knockout of *sorA* in *T. reesei*  $\Delta$ tmus53 experiment within our group showed not only the

abolition of sorbicillinoid production but also the accumulation of a precursor that was identified as scytolide **158**. In addition, spirosorbicillinols A **160** and B **161** were produced by feeding scytolide **158** to *A. oryzae* strains expressing sorA, sorB, sorC and sorD simultaneously, which confirmed the previous proposal that they were derived from sorbicillinol **159** and scytolide **158**.<sup>102</sup>



**Figure 3.2.4.1** Morphologies of *T. reesei*  $\Delta$ tmus53 : **A**, *T. reesei*  $\Delta$ tmus53 growing on ME agar; **B**, *T. reesei*  $\Delta$ tmus53 growing on ME medium.

*T. reesei*  $\Delta$ tmus53 show higher rates of homologous integration than wild type strain during gene knock out experiments, because MUS53 (a homolog of human Lig4) is required for nonhomologous end joining (NEHJ).<sup>168</sup> *T. reesei*  $\Delta$ tmus53 was cultured in ME medium at 28 °C and 110 rpm for 4-5 d. The organism was obtained from Ms. Eman F. Bassiony, and she received the organism as a gift from the Mach-Aigner group (Institute of Chemical Engineering, TU Wien, Vienna, Austria). Cells were removed by filtration and the supernatant was extracted with ethyl acetate two times. Organic extracts were dried and dissolved in acetonitrile to a concentration of 10 mg/mL for HPLC analysis. To find scytolide **158**, extracted ion chromatograms (EIC) of cultures were analyzed at 241 Da (ES<sup>+</sup> mode) corresponding to the expected [M + H]<sup>+</sup> molecular ion. A major peak was observed at the  $t_R = 4.2$  min with UV absorption  $\lambda_{max}$  211 nm.



**Figure 3.2.4.2** The LCMS chromatogram of *T. reesei*  $\Delta$ *tmus53* culture extracts after 5 days' growth in ME medium: **A**, ELSD chromatogram of *T. reesei*  $\Delta$ *tmus53* culture extracts; **B**, Extracted ion chromatograms [M + H]<sup>+</sup>; **C**, Extracted ion chromatograms [M - H]<sup>-</sup>; **D**, UV chromatogram of *T. reesei*  $\Delta$ *tmus53* culture extracts; **E**, UV spectrum of **158**; **F**, Mass spectrum of **158**.

### 3.2.5 Timecourse of Metabolite Production from *P. cirsii*

Time-course production experiments were performed to determine the time of maximum production of scytolide **158**. For such experiments, 7 flasks PD media (100 ml in one flask) were inoculated using a spore-suspension of *P. cirsii* and grown under the same conditions (28 °C and shaking 110 rpm). Analysis of cultures with LCMS started from day 3 until day 11 (Figure 3.2.5.1). The relative concentration of scytolide **158** was indicated using DAD peak area, which was only used as an internal comparison of the accumulation of targeted compound (Figure 3.2.5.2). The investigation showed that production of scytolide **158** had begun by day 5, increasing intensely to reach its peak by day 9. After that, a new peak appeared, which was identified to be phyllostictine A **196**, already known to be a metabolite of *P. cirsii* based on comparison of UV absorption and ESI ionisation patterns (Figure 3.2.5.3).<sup>142</sup>

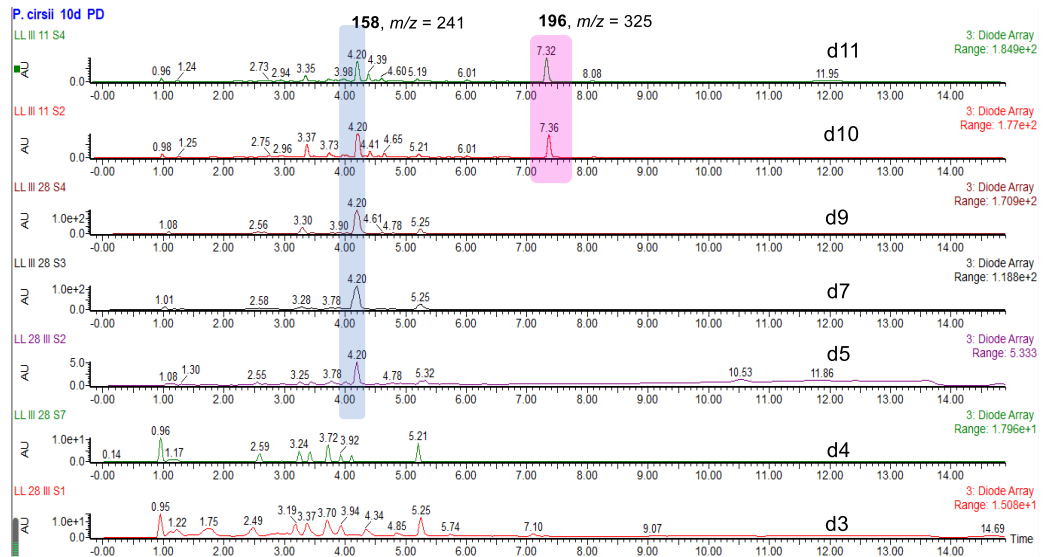


Figure 3.2.5.1 LCMS chromatogram of a time-course secondary metabolites production from *P. cirsi*.

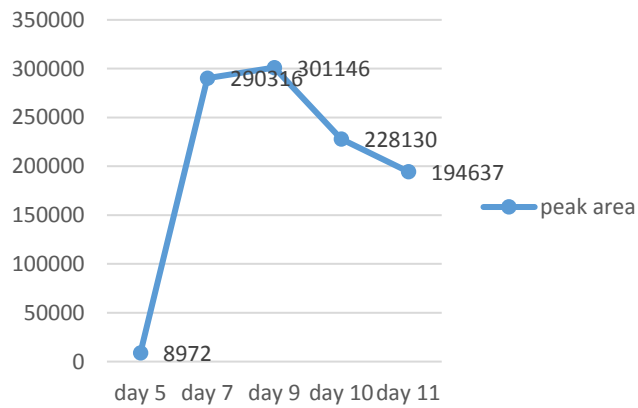


Figure 3.2.5.2 Production curve of scytolide 158 production in *P. cirsi* WT over 11 days.

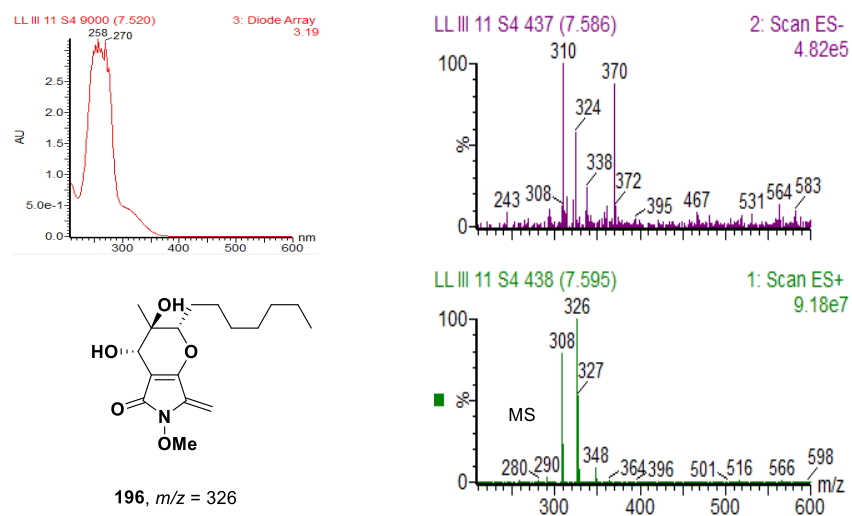
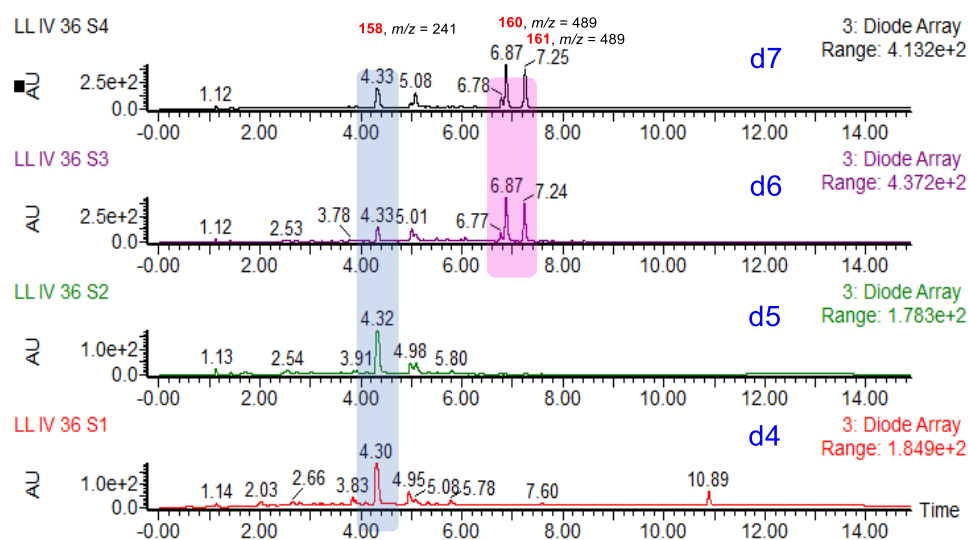


Figure 3.2.5.3 UV and MS spectrum of phyllostictine A 196.

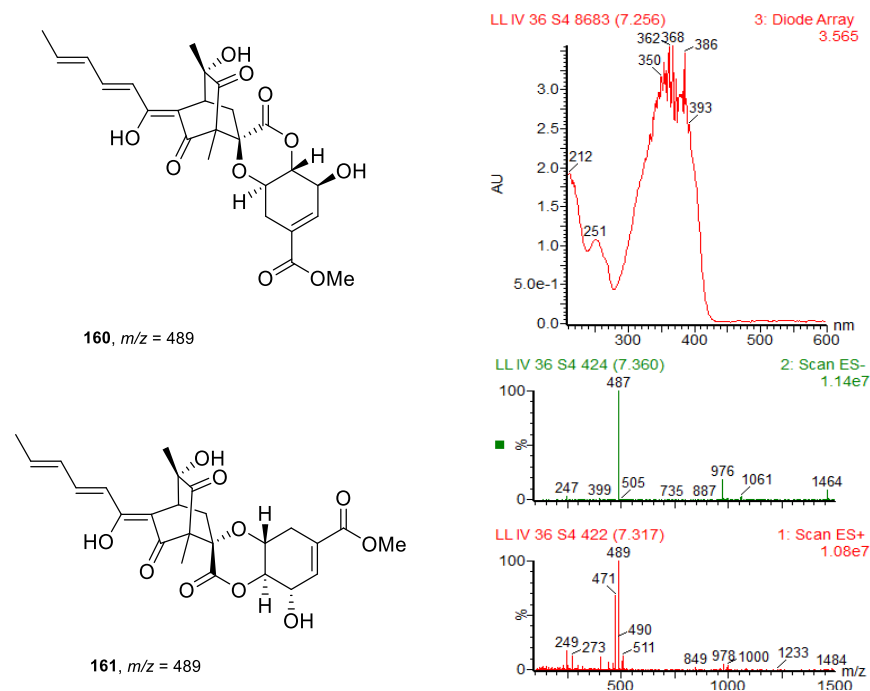
### 3.2.6 Timecourse of Metabolite Production from *T. reesei* $\Delta$ tmus53

In addition, a time-course study was also performed in *T. reesei*  $\Delta$ tmus53 to monitor the change of metabolites profile in a liquid culture. Four Erlenmeyer flasks of ME medium with *T. reesei*  $\Delta$ tmus53 spore-suspension were grown under the same conditions (28 °C and shaking 110 rpm). The timing of accumulation of scytolide **158** was investigated by analysis of culture flasks extracted from day 4 to day 7. The HPLC chromatograms showed that the production of scytolide **158** started on day 4 and was at high level on day 5. However, on the chromatogram of day 6 the peak of scytolide **158** became small, while peaks of other compounds eluting at 6.8 min and 7.2 min appeared (Figure 3.2.6.1). Those peaks had molecular ions of  $m/z$  488 and UV absorption  $\lambda_{max}$  362 nm corresponded to spirosorbicillinols A **160** and B **161** respectively (Figure 3.2.6.2), which are consistent with the fact that scytolide **158** is the precursor of spirosorbicillinol A **160** and B **161**.<sup>102</sup>



**Figure 3.2.6.1** LCMS chromatogram of a time-course secondary metabolites production from *T. reesei*  $\Delta$ tmus53.





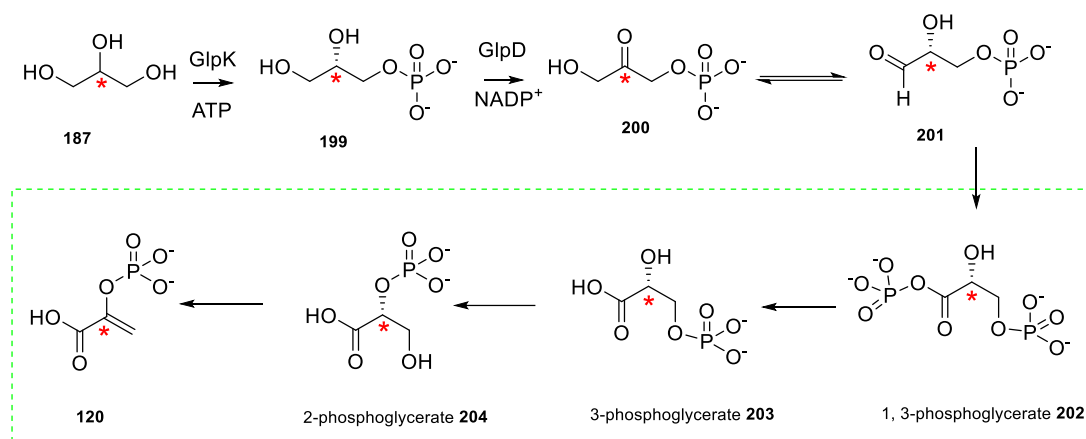
**Figure 3.2.6.2** UV and MS spectrum of spirosorbicillinols A **160** and B **161**.

### 3.3 Results-Glycerol Feeding Experiments

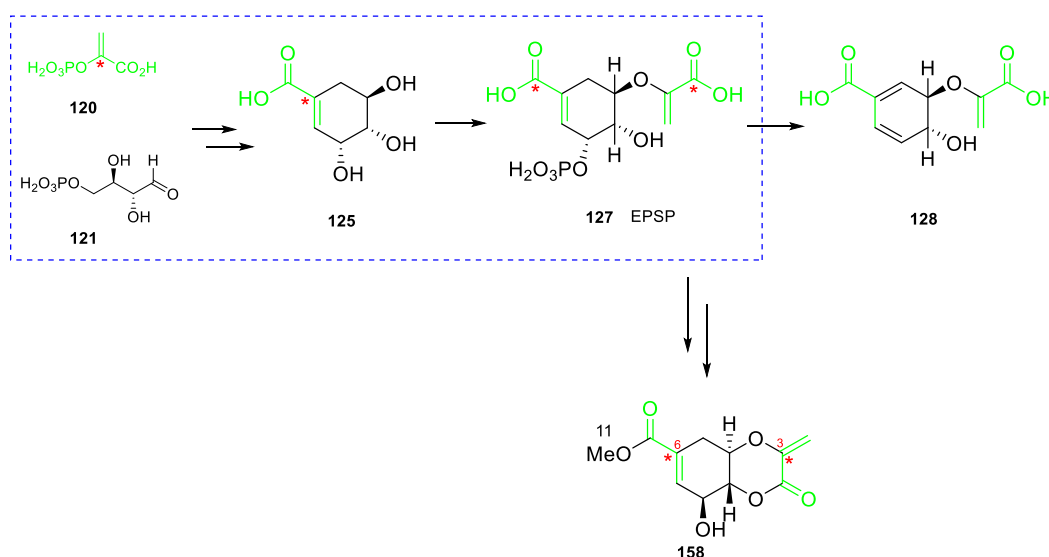
In recent years, isotopolog profiling has proven to be a powerful approach to determine the biosynthetic origin of secondary metabolites from a number of fungi. The method is based on  $^{13}\text{C}$ -incorporation experiments using completely or selectively  $^{13}\text{C}$ -labelled precursors such as acetate, methionine or glycerol. Due to the isotopic labels, specific  $^{13}\text{C}$ -enrichments are generated in downstream metabolic products which can be determined by  $^{13}\text{C}$ -NMR spectroscopy. By comparison of these NMR datasets, the differential substrate usage and the metabolic pathways can be clarified.<sup>169</sup>

Glycerol **187** can be considered as a favorable carbon source for the production of aromatic amino acids including phenylalanine **129** and tyrosine **130**, which are derived from the shikimic acid pathway.<sup>170</sup> The shikimate pathway is initiated by the condensation reaction between phosphoenolpyruvate **119** (PEP) and erythrose-4-phosphate **121** (E4P). PEP **120** formation from glycerol **187** can occur via the lower glycolytic pathway: glycerol **187** is activated by an ATP-dependent glycerol kinase (GlpK) to yield glycerol-3-phosphate **199** (G3P). It is then oxidized at C-2 by a membrane-bound ubiquinone-8-dependent G3P dehydrogenase (GlpD) to dihydroxyacetone phosphate **200** (DHAP) which is, in-turn, isomerized to glyceraldehyde 3-phosphate **201** (GA3P);

Both DHAP **200** and GA3P **201** are further metabolized in the glycolysis pathway (Scheme 3.3.1).<sup>171,172</sup> The product of this pathway, PEP **120**, is then used as the start unit in the shikimate pathway (Scheme 3.3.2).



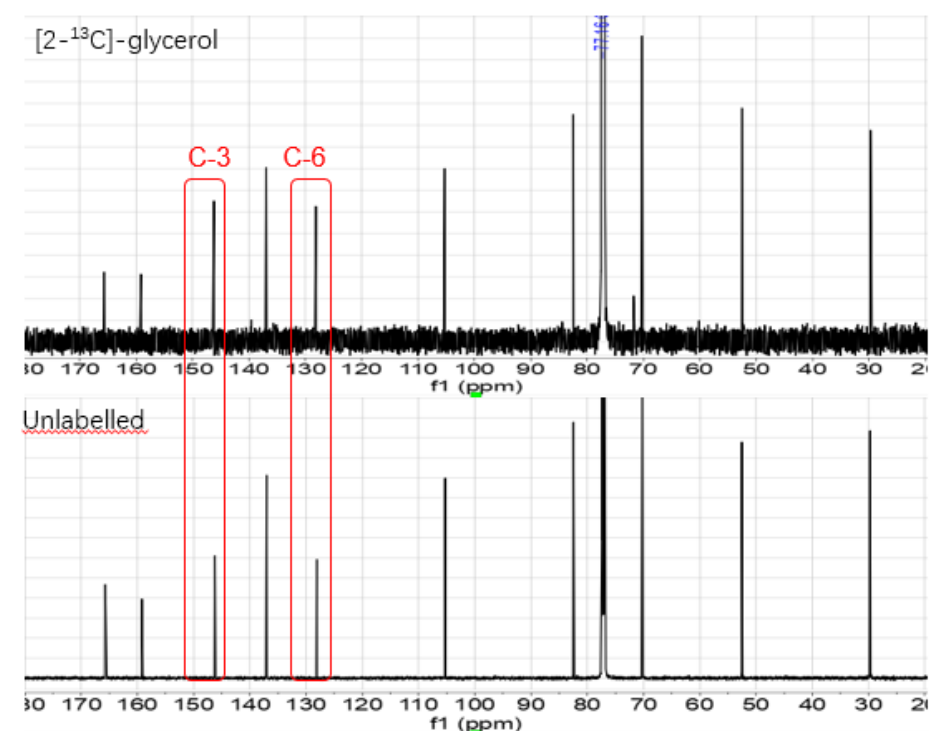
**Scheme 3.3.1** Biosynthesis of phosphoenolpyruvate from glycerol **187**; Glycolysis pathway are highlighted with green dotted line.



**Scheme 3.3.2** Proposed biosynthetic pathway of scytolide **158**. Shikimic acid pathway intermediates are highlighted with a blue dotted line.

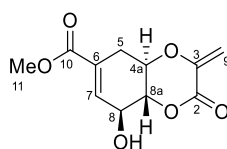
Scytolide **158** is structurally similar to EPSP **127**, which is the key intermediate on the pathway from shikimate **125** to chorismate **128** (Scheme 3.3.2). To determine if the origin of scytolide **158** is from the shikimate pathway, feeding experiments were done by supplementing producing cultures with [2-<sup>13</sup>C]-glycerol **187** to a final concentration of 10 mM. The labelled building block was supplemented at the beginning of scytolide biosynthesis based on time-course evidence as the

following way: *P. cirsii* was grown in 5 flasks of PD medium (100 ml per flask); 465 mg [2-<sup>13</sup>C]-glycerol **187** was diluted in 2 ml H<sub>2</sub>O; 100 μl labeled precursor was added to each flask at days 4, 5, 6 and 7 respectively. The crude metabolites were extracted from 500 ml cultures using ethyl acetate, which was further purified by preparative-HPLC to obtain 1.5 mg labelled scytolide. To determine the level and position of the incorporation, the samples were analyzed by <sup>13</sup>C NMR and compared to the unlabelled scytolide in carbon signal intensity (Figure 3.3.1). Peak intensity enhancement at the same carbon indicated the incorporation of <sup>13</sup>C.



**Figure 3.3.1** NMR signals of scytolide with labelled precursor compared to unlabelled scytolide; Signals with enhanced intensity are highlighted with red rectangle.

Scytolide carbon signal intensity was indicated by peak height, and was further normalized by dividing its value over the intensity of C-10 (ester group) of the same NMR dataset acquisition. The ester group was used as reference, as it lays at an unlabeled position in the feeding experiment. Peak enhancement was estimated by calculating the ratio between the normalized carbon peak intensity of labelled scytolide over that of unlabelled scytolide. Significance was set at a fold increase with threshold 1.5 (Table 3.3.1).

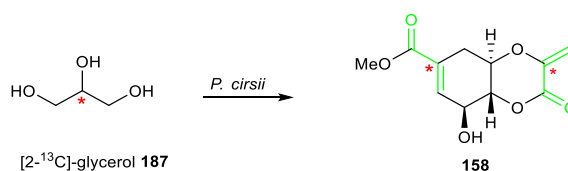


158,  $m/z = 241$

C	Normalized peak intensity (unlabelled)	Normalized Peak intensity (labelled)	Peak enhancement
2	0.8	1.0	1.2
3	1.3	2.0	1.5
4a	3.3	4.4	1.3
5	2.6	3.0	1.1
6	1.3	2.0	1.5
7	2.2	2.6	1.2
8	2.6	3.1	1.2
8a	2.7	3.2	1.2
9	2.1	2.5	1.1
10	1	1	1
OMe	2.5	3.3	1.3

**Table 3.3.1:** NMR carbon signal peak enhancement of labelled scytolide; Significance was set as threshold fold  $\geq 1.5$ .

A metabolite derived from the shikimate pathway would be expected to have the moiety derived from PEP **120** and erythrose-4-phosphate **121**, while the moiety derived from PEP **120** should have the incorporation of  $^{13}\text{C}$  for the specific carbon (Scheme 3.3.2). In particular, we expected  $[2-^{13}\text{C}]$ -glycerol **187** feeding to enrich C-3 and C-6 NMR signals in scytolide, because C-3 and C-6 are most likely derived from PEP **120** by the action of the enzymes from the shikimate pathway respectively and  $[2-^{13}\text{C}]$ -glycerol **187** is converted to C-2 labelled PEP **120** via the glycolytic pathway. Excitingly, the results from feeding  $[2-^{13}\text{C}]$ -glycerol **187** were in accord with the proposal, which confirmed that scytolide was derived from the shikimate pathway, probably EPSP **127** (Scheme 3.3.3).

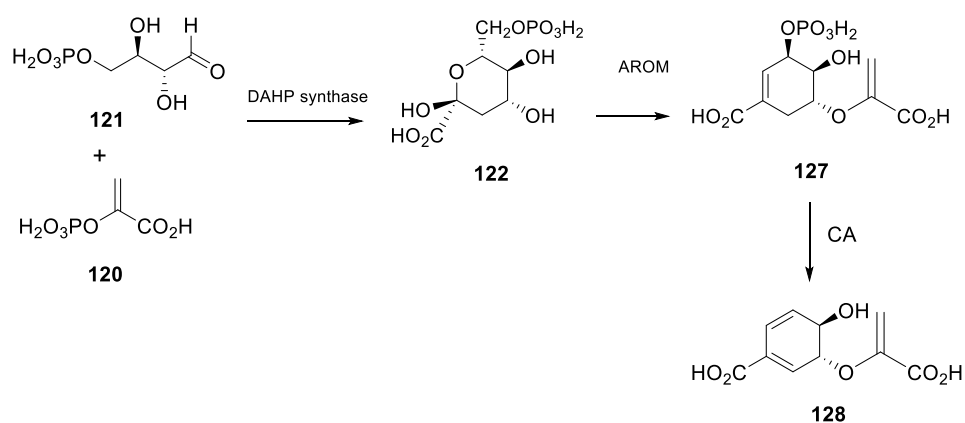


**Scheme 3.3.3** Incorporation of  $[2-^{13}\text{C}]$ -glycerol **187** into scytolide **158**.

### 3.4 Results-Genome of *P. cirsi* and *T. reesei* $\Delta$ tmus53

#### 3.4.1 Shikimic Acid Pathway in Fungi

The shikimate pathway begins with the condensation of PEP **120** and E4P **121** to form 3-deoxy-D-arabino-heptulosonate-7-phosphate (DAHP), and requires another six enzymatic steps to produce chorismate **128** (Scheme 3.4.1). The enzymes catalysing the shikimate pathway in bacteria are monofunctional and are encoded by genes that are dispersed throughout the genome. In fungi, the enzymes catalysing steps two to six are fused into a single pentafunctional polypeptide known as the AROM protein, which converts DAHP **122** to EPSP **127** in the shikimate pathway.<sup>173</sup> Comparison of the predicted amino acid sequences encoded by AROM locus in fungi with those of shikimate pathway genes from an array of bacterial species strongly suggests that they are divergently related to common ancestral sequences and that the gene encoding pentafunctional AROM polypeptide is evolved by multiple gene fusions in the shikimate pathway.<sup>173,174</sup>

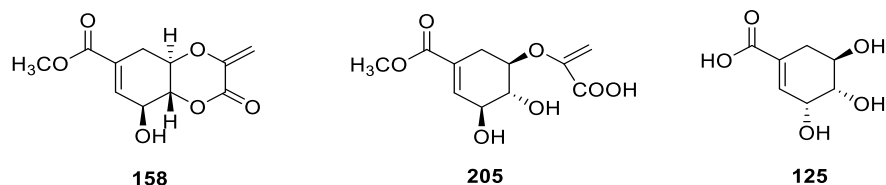


**Scheme 3.4.1** The shikimate pathway consists of seven chemical reaction steps in fungi.

#### 3.4.2 *In silico* Analysis of Putative BGC for Scytolide

Based on the structural analysis, scytolide **158** might be derived biogenetically *via* lactonization of the enol pyruvate of methyl 3-epishikimate **205** originated from the shikimate pathway (Figure 3.4.2.1). Scytolide **158** was first reported in 1993 as a metabolite produced when *Scytalidium uredinicola* is grown on malt extract.<sup>99</sup> Up to now, it is also produced by *P. cirsi* and *T. reesei*  $\Delta$ tmus53 *AsorA*.<sup>100,102</sup> In previous work of the Cox group, total genome sequence of *P. cirsi* was

already performed by Paired End Illumina sequencing at the Center for Biotechnology, Bielefeld, Germany and annotated in the GenDBE platform.<sup>142</sup> The genome sequence of *T. reesei* has been publically available since 2008 under accession number AAIL00000000.



**Figure 3.4.2.1** Structures of scytolide **158**, methyl 3-epishikimate **205** and shikimate **125**.

In many fungi, the enzymes catalyzing the shikimic acid pathway are fused into a single polypeptide (the pentafunctional AROM protein). To find the possible biosynthetic gene cluster of scytolide, the translated amino acid (AA) sequences of three pentafunctional AROM polypeptides from *Acremonium sp*, *Penicillium freii* and *Fusarium proliferatum ET1* were used as templates in a BLAST search of the protein encoded by the *T. reesei* *Atmus53* genome. The protein blast results from NCBI indicated **three shikimate pathway gene homologues**. For simplicity, these three genes are assigned generic names, *scy TA*, *scy TB* and *scy TC* (Table 3.4.2.1).

Locus tag / Size	Generic name	Identity % ( <i>Acremonium sp</i> )	Identity % ( <i>Penicillium freii</i> )	Identity % ( <i>Fusarium proliferatum ET1</i> )	Putative function
TRIREDRAFT_ 82516 / 4728 bp	<i>Scy TC</i>	70 %	61 %	70 %	Pentafunctional Aromatic polypeptide
TRIREDRAFT_ 5296 / 2142 bp	<i>Scy TA</i>	28 %	31 %	34 %	type I 3- dehydroquase- like protein
TRIREDRAFT_ 65869 / 2547 bp	<i>Scy TB</i>	26 %	24 %	25 %	quinic 5- dehydrogenase

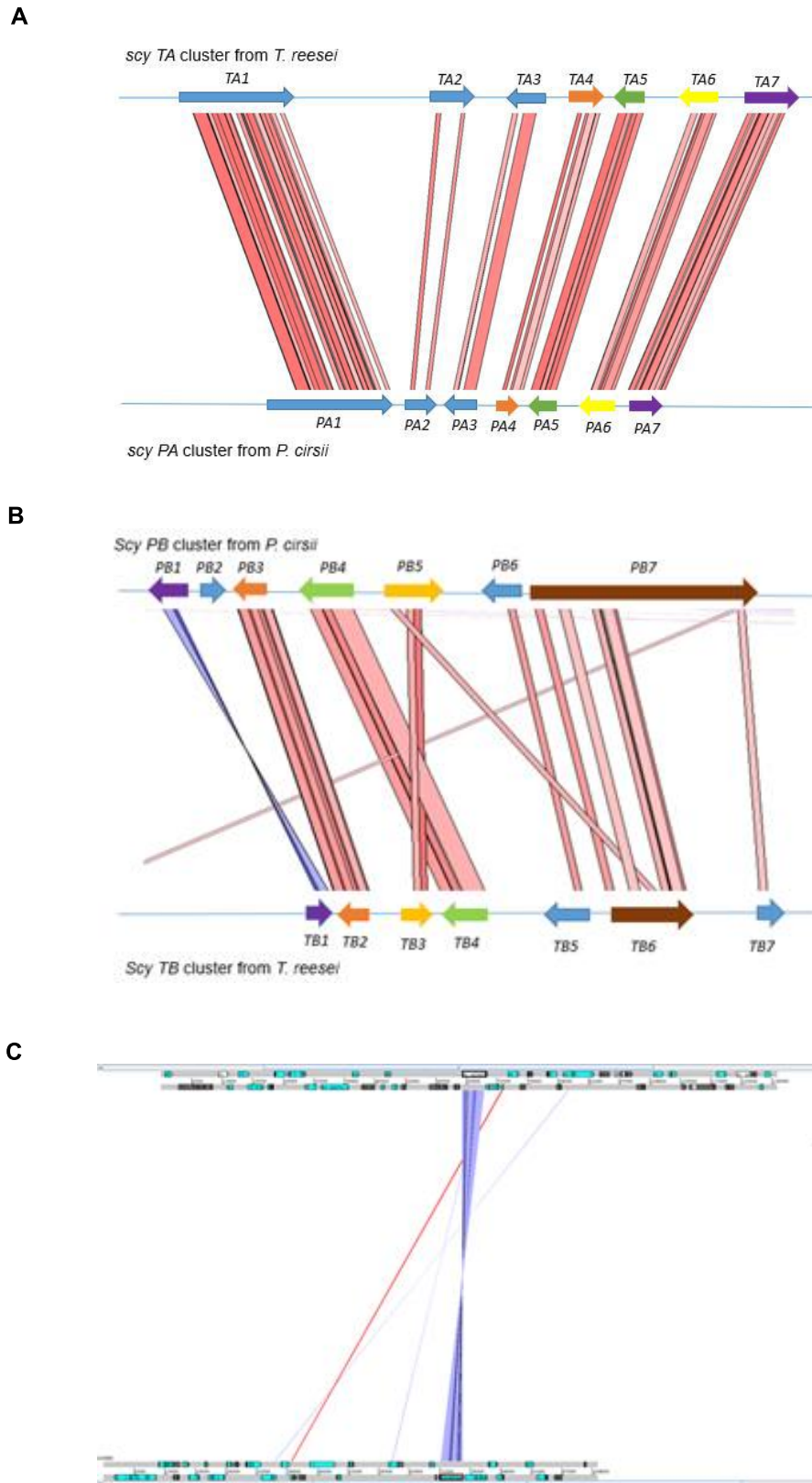
**Table 3.4.2.1:** Three homologous genes of pentafunctional AROM polypeptides in *T. reesei* *Atmus53*

Usually, genes responsible for secondary metabolite production are clustered in fungi. Here, a cluster was defined as extending from the core shikimate pathway gene in both directions until enough genes were included. Thus, three BGC candidates were defined based on genes *scy TA*, *scy TB* and *scy TC*. With three putative cluster from *T. reesei* *Atmus53* tentatively identified, clustered genes surrounding *scy TA-TC* were blasted for the putative protein function. Due to same metabolite from *T. reesei* *Atmus53* and *P. cirsii*, gene clusters involved in scytolide biosynthesis in both *T. reesei*

*Atmus53* and *P. cirsii* are predicted to show considerable homology. Thus, translated AA similarity of genes from *scy TA-TC* cluster were searched in the whole genome of *P. cirsii* using Custom BLAST in Geneious. Both the *scy TA* and *scy TB* clusters were found to have clustered homologous genes in the *P. cirsii* genome (Figure 3.4.2.2 A-B ), which were named as *scy PA* and *scy PB* clusters respectively. However, no clustered homologous gene of *scy TC* cluster was observed, so *scy TC* gene cluster was quickly discounted from potential involvement in scytolide biosynthesis (Figure 3.4.2.2 C). *Scy TC* is predicted as the primary metabolism pentafunctional aromatic polypeptide that belongs to shikimate pathway.

Putative cluster genes were assigned generic names according to their position from left to right. For example, the leftmost gene in the *scy PA* cluster was dubbed *PA1*, and immediately to the right became *PA2*. Both *scy PA* and *scy PB* clusters consisted of seven different open reading frames. All the genes from *scy PA* cluster had homologues in *scy TA* cluster, whereas only 5 genes from *Scy PB* cluster were found in proximity to *scy TB* cluster. From this analysis the *scy PA* cluster was the most promising BGC for scytolide **158** (Table 3.4.2.2).

Putative annotations of all the genes from *scy PA* and *scy PB* clusters were predicted by running protein Blast in NCBI. In the *scy PA* cluster, the enzyme encoded by *PA5* is assigned as DAHP synthase, which catalyzes the reaction of PEP **120** and D-erythrose-4-phosphate **121** to DAHP **122** and phosphate. The protein encoded by *PA4* is a putative member of the methyltransferase family, a class of enzymes known to catalyze methylation of diverse substrates. The enzyme encoded by *PA6* is predicted to be a non-heme iron  $\alpha$ -ketoglutarate dependent oxygenase, a class of multifunctional oxygenase. *PA7* is proposed to encode an acetyl-CoA synthetase-like protein while the protein encoded by *PA1* is predicted to be a transporter protein. Another two genes are characterized as unknowns (Table 3.4.2.3).



**Figure 3.4.2.2:** Homologous gene clusters analysis between *P. circsii* and *T. reesei*  $\Delta$ *mus53* : **A**, Artemis analysis of *scy TA* and *scy PA* clusters; **B**, Artemis analysis of *scy TB* and *scy PB* clusters; **C**, Artemis analysis of *scy TC* and *scy PC* clusters.



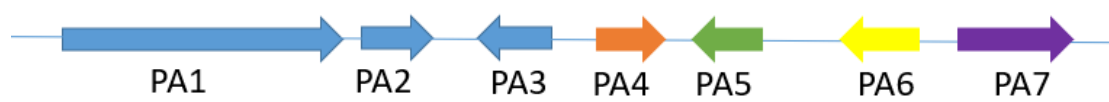
Generic name in <i>scy PA</i> cluster	Generic name in <i>scy TA</i> cluster	Identity	Generic name in <i>scy PB</i> cluster	Generic name in <i>scy TB</i> cluster	Identity
<i>PA1</i>	<i>TA1</i>	62.5 %	<i>PB1</i>	<i>TB1</i>	67.6 %
<i>PA2</i>	<i>TA2</i>	35.7 %	<i>PB3</i>	<i>TB2</i>	54.7 %
<i>PA3</i>	<i>TA3</i>	42.1 %	<i>PB4</i>	<i>TB4</i>	68.8 %
<i>PA4</i>	<i>TA4</i>	57.6 %	<i>PB5</i>	<i>TB3</i>	60.8 %
<i>PA5</i>	<i>TA5</i>	54.0 %	<i>PB7</i>	<i>TB6</i>	43.6 %
<i>PA6</i>	<i>TA6</i>	62.9 %			
<i>PA7</i>	<i>TA7</i>	54.5 %			

**Table 3.4.2.2:** Identity of amino acid sequence between two homologous genes in both *scy TA* / *scy PA* clusters and *scy TB* / *scy PB* clusters

DAHP synthase is known to catalyze the first step in the shikimate pathway. Usually there are two or three differently regulated DAHP synthase isoenzymes in fungi, which are repressed by either tyrosine **130** or phenylalanine **129**.<sup>175</sup> However, the putative DAHP synthase (*PA5*) from the *scy PA* cluster was not the homologous gene of DAHP synthase in the shikimate pathway. It is proposed that the protein encoded by *PA5* is regulated by some unknown secondary metabolites and presumably makes DAHP **122** using PEP **120** and D-erythrose-4-phosphate **121**. DAHP **122** could then be used as the substrate for the pentafunctional AROM protein from primary metabolism, which produces EPSP **127** that might serve as the closest intermediate on the biosynthetic pathway to scytolide **158**. After that, three tailoring chemical reactions are necessary for converting EPSP **127** to scytolide **158**: Dephosphorylation of EPSP might be catalyzed by one of the unknown proteins from the *scy PA* cluster, and the configuration variation of hydroxyl group at C-8 indicates that an epimerization reaction happens simultaneously or later than dephosphorylation reaction. The second reaction might be an intramolecular lactone ring cyclization catalyzed by ATP dependent enzyme encoded by *PA7* to produce intermediate **207**. In the third reaction step, methyltransferase encoded by *PA4* converts **207** into scytolide **158** by adding a methyl to the carboxyl group. Combined with the result of the Artemis analysis between *scy PA* cluster and *scy TA* cluster that all 7 genes showed considerable sequence homology, *scy PA* cluster emerged as the frontrunner to encode scytolide biosynthesis. In final, a putative scytolide biosynthetic pathway based on *scy PA* cluster was proposed (Scheme 3.4.2.1).

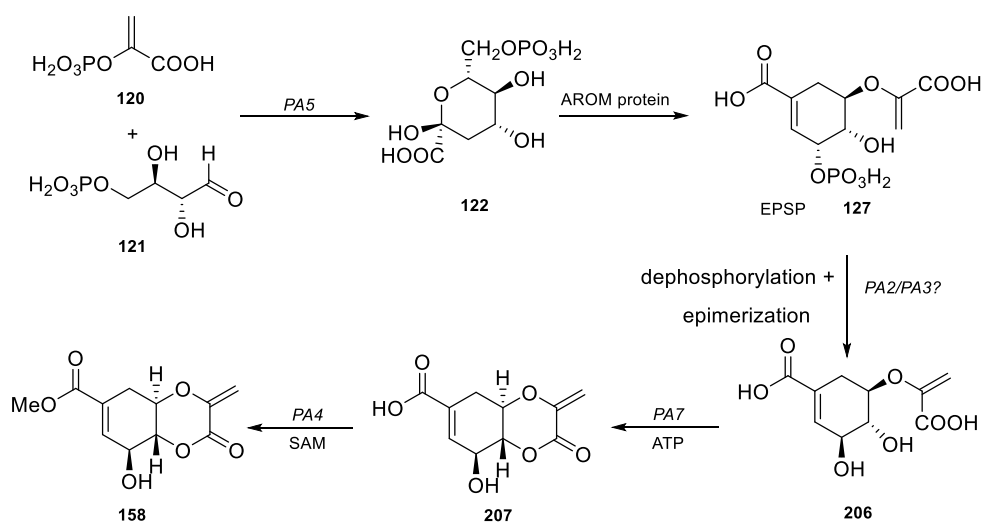
In the *scy PB* cluster, the enzyme encoded by *PB1* is assigned as a dehydroquinase, which converts 3-dehydroquinone **123** to 3-dehydroshikimate **124**. Protein encoded by *PB3* is related to 3-dehydroshikimate dehydratase, this enzyme catalyzes the dehydration of **124**. The enzyme of *PB5* is predicted to be related to quinate dehydrogenase, which belongs to the family of oxidoreductases.

*PB4* is predicted to be similar to quinate permease, an integral membrane transporter that imports quinic acid to be catabolized as a carbon source. The protein encoded by *PB2* is predicted to be related to inositol monophosphatase. Transcription of *scy PB* cluster might be positively mediated by the quinic acid utilization activator encoded by *PB6* and negatively mediated by quinate repressor protein, the product of the *PB7* gene (Table 3.4.2.4). All the proteins encoded by the *scy PB* cluster corresponded to expected enzymes from the known fungal quinate pathway.<sup>176</sup> Thus, the *scy PB* cluster is probably involved in the catabolism of quinate rather than scytolide biosynthesis in *P. cirsii*.




Name and Locus_tag	Annotation	Near Pblast hit
<i>PA1</i> Pc_g4409	putative ATP-binding multidrug cassette transport protein	<i>Fusarium napiforme</i> FNAPI_6401
<i>PA2</i> Pc_g4410	uncharacterized protein	<i>Macroventuria anomochaeta</i> BU25DRAFT_449737
<i>PA3</i> Pc_g4411	Hypothetical protein	<i>Macroventuria anomochaeta</i> BU25DRAFT_344824
<i>PA4</i> Pc_g4412	methyltransferase family protein	<i>Purpureocillium lilacinum</i> VFPBJ_02445
<i>PA5</i> Pc_g4413	DAHPh synthetase	<i>Glonium stellatum</i> AOQ84DRAFT_433044
<i>PA6</i> Pc_g4414	non-heme iron $\alpha$ -ketoglutarate dependent oxygenase	<i>Fusarium tjaetaba</i> FTJAE_10257
<i>PA7</i> Pc_g4415	acetyl-CoA synthetase-like protein, ATP-binding enzyme	<i>Macroventuria anomochaeta</i> BU25DRAFT_452728

**Table 3.4.2.3:** Annotations of *scy PA* cluster in *P. cirsii*



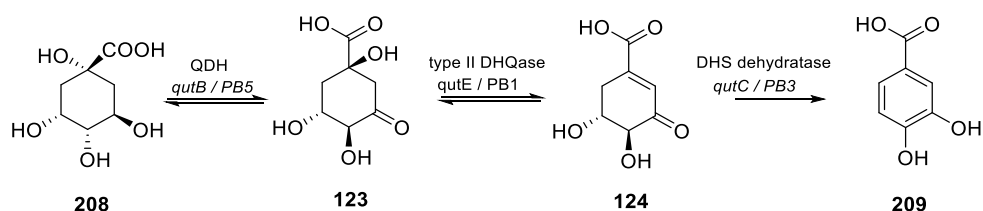
**Scheme 3.4.2.1** Putative scytolide 158 biosynthetic pathway based on *scy PA* cluster



Name Locus_tag	Annotation	Near Pblast hit
<i>PB1</i> (Pc_g3314)	Dehydroquinase	<i>Ampelomyces quisqualis</i> BDU57DRAFT_519147
<i>PB2</i> (Pc_g3315)	inositol monophosphatase	<i>Setomelanomma holmii</i> EK21DRAFT_59764
<i>PB3</i> (Pc_g3316)	3-dehydroshikimate dehydratase	<i>Ophiobolus disseminans</i> CC86DRAFT_428322
<i>PB4</i> (Pc_g3317)	Quinate permease	<i>Alternaria arborescens</i> AA0111_g6073
<i>PB5</i> (Pc_g3318)	quinate dehydrogenase	<i>Stemphylium lycopersici</i> TW65_91092
<i>PB6</i> (Pc_g3319)	quinic acid utilization activator	<i>Decorospora gaudefroyi</i> BDW02DRAFT_386065
<i>PB7</i> (Pc_g3320)	quinate repressor protein	<i>Parastagonospora nodorum</i> SN15 JI435_114350

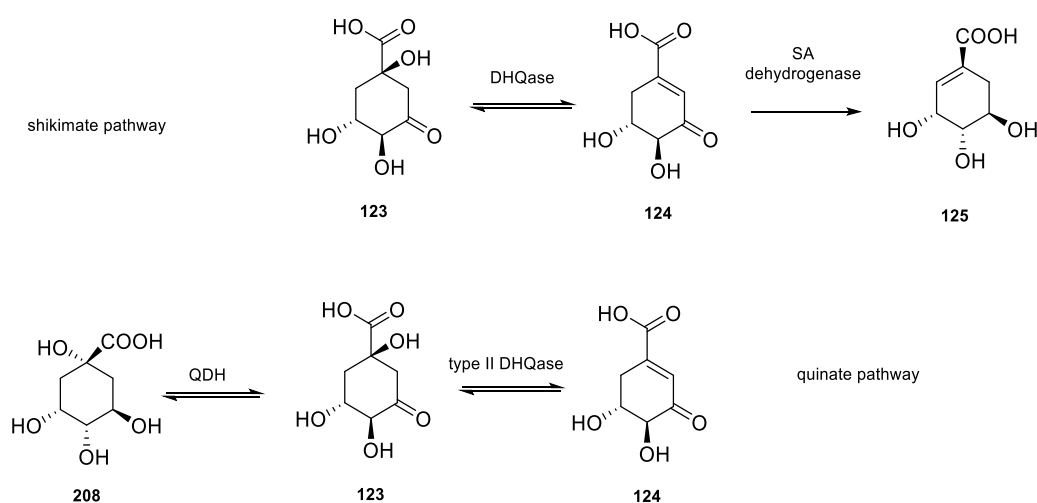
**Table 3.4.2.4.** Annotation of *scy PB* cluster in *P. cirsii*

Quinate **208** is an abundant carbon source for many fungi. It is a metabolite reversibly formed from a side branch of the shikimate pathway from 3-dehydroquinate **123** catalyzed by quinate dehydrogenase (QDH).<sup>177</sup> The molecular mechanisms involved in the catabolism of quinate have been intensively studied by genetic and molecular techniques in the filamentous fungi *Neurospora crassa* and *Aspergillus nidulans*. A cluster of eight genes (designated *qut* BGC) have been discovered from *A. nidulans*, including *qutA* (encodes an activator protein), *qutB* (encodes quinate dehydrogenase), *qutC* (encodes a dehydroshikimate dehydratase), *qutD* (encodes a quinate permease), *qutE* (encodes catabolic type II 3-dehydroquinase), *qutG* (encodes a protein homologous with bovine myo-inositol monophosphatase), *qutH* (encodes a protein that has a low level similarity with a DNA polymerase from Hepatitis B virus) and *qutR* (encodes a repressor protein).<sup>178</sup> The quinate pathway consists of three reaction steps (Scheme 3.4.2.2), beginning with the reversible reaction from quinic acid **208** to 3-dehydroquinate **123** catalyzed quinate dehydrogenase. 3-Dehydroquinate **123** is then converted to 3-dehydroshikimate **124** (SA) by the catalytic action of type II 3-dehydroquinase (type II DHQase), this reaction also exists in the shikimate pathway. The last reaction step is the production of protocatechuic acid **209** (PCA) catalyzed by dehydroshikimate dehydratase (DHS dehydratase).



**Scheme 3.4.2.2** The quinate pathway consists of three chemical reaction steps in fungi.

3-Dehydroquinic acid **123** and dehydroshikimate **124** are two common metabolites shared between the shikimate pathway and the quinate pathway in many fungi and some bacteria. There are two types of 3-dehydroquinase enzyme that interconverts DHQ **123** and DHS **124**, one is from a component of a pentafunctional polypeptide encoded by *AROM* locus while the another one is from a single-function polypeptide encoded by a gene in the *qut* *BGC*. Unexpectedly, these two isoenzymes show no sequence similarity, strongly suggesting convergent evolution.<sup>173</sup> However, quinate dehydrogenase from the quinate pathway and the shikimate dehydrogenase in *AROM* protein appear to be divergently related to a common ancestral sequence. Most surprising is the finding that the three C-terminal enzyme activities of the *AROM* protein (specifying shikimate kinase, 3-dehydroquinase and shikimate dehydrogenase) are clearly related to the *QUTR* protein active in the quinate pathway, which are consistent with the sequence similarities among *AROM* locus, *TB6* and *PB7* genes found in our bioinformatic analysis.<sup>66,176,179</sup>

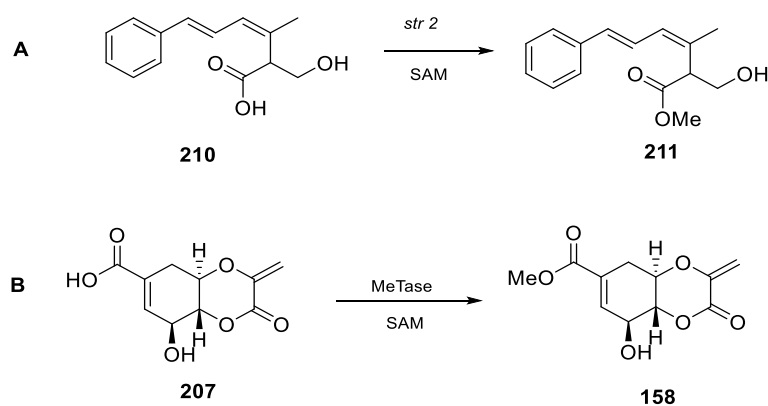


**Scheme 3.4.2.2** Two common metabolites shared between shikimate pathway and quinate pathway.

### 3.4.3 *In silico* Analysis of Methyltransferase

Since many of the genes from the shikimate pathway are not clustered in microorganism, it is also possible that the genes responsible for the biosynthesis of scytolide **158** are not clustered. For this reason, we took an alternative approach for the discovery of potential genes involved in its biosynthesis. Considering that a methyltransferase must be necessary for the methyl group in scytolide **158**, the gene that encodes this functional enzyme might exist randomly in the whole genome of both *P. cirsii* and *T. reesei*  $\Delta$ tmus53 but show considerable sequence homology. Methyltransferases (MeTs) are common tailoring enzyme during the biosynthesis of natural products. They form a large family of enzymes that methylate a diverse set of targets, ranging from the three major biopolymers to small molecules. Most of these MeTases use the cofactor *S*-adenosyl-*L*-Methionine (SAM) as a methyl group donor. They perform this function by binding the methyl donor SAM and catalyzing the covalent transfer of a methyl group from SAM to their substrates.<sup>180</sup>

181



**Figure 3.4.3.1** Chemical reaction catalyzed by methyltransferases.

To find all the possible methyltransferases encoded by the *P. cirsii* and *T. reesei*  $\Delta$ tmus53 genomes, the translated amino acid (AA) sequence of *str2* from *Strobilurus tenacellus* was used as a template in a BLAST search of homologous genes. The gene *str2* encodes an *O*-methyltransferases (*O*-MeT) that selectively catalyzes the methylation of the carboxyl group of desmethylbolineol **210** to bolineol **211** (Figure 3.4.3.1A).<sup>182</sup> The protein blast results from NCBI indicated 25 *O*-MeT gene homologues

in *T. reesei* *Δtmus53* and 19 *O*-MeT gene homologues in *P. cirsii*. It was worth noting that *PA4/TA4* were not included in the blast results, which were predicted to encode methyltransferases in chapter 3.4.2. All the scattered methyltransferase genes found in *P. cirsii* and *T. reesei* *Δtmus53*, together with *PA4/TA4* were analyzed for % identities using multiple sequence alignment in Geneious. 11 genes that encode putative methyltransferases from *P. cirsii* were found to show relatively high % identity to 17 putative methyltransferases in *T. reesei* *Δtmus53*. Among them *PA4* and *TA4* show the highest homology with 55.7% identity, and they both show low similarity to other methyltransferase genes. So, no specific and promising methyltransferase that might encode a methylation step from the putative intermediate **207** to scytolide **158** (Figure 3.4.3.1B) was found in this section.

	PA4	P.C MeT 6	P.C MeT 1	P.C MeT 8	P.C MeT 4	P.C MeT 2	P.C MeT 3	P.C MeT 7	P.C MeT 5	P.C MeT 9	P.C MeT 10
TA4	55.7	19	20	15.1	16	16.7	17.6	19.6	20.4	18.4	19.7
MeT 7	17.3	37.3	24.2	23	22.3	27.1	24.7	24.1	24	24.6	16.9
MeT 17	14.9	17.5	25.2	20.3	19.6	24.5	26.7	24.3	31.2	36.9	33.3
MeT 8	15.3	18.2	26.5	28	23.9	27.8	26.5	28.8	27.5	24.9	19.3
MeT 4	14.6	16.5	40.3	31.3	32.6	36.7	34.7	37.6	29.6	28	17.5
MeT 13	13	17.2	24	22.6	23.5	25.1	25.6	26.7	26.3	22	16.3
MeT 3	14.8	21.5	27.1	29.7	27.7	30.1	29.5	31.9	31	28	18.1
MeT 1	16.8	20.8	29.7	32.3	30.8	35.5	34.7	33.5	33	31.5	18.9
MeT 11	15.1	18.5	33	28.9	29.3	32	30.8	30.9	29.7	24.4	18.2
MeT 10	13.8	20.1	28.3	35.8	27.9	34.1	29.2	30.6	26.1	26.9	16.5
MeT 5	14.5	18.9	28.1	38	28.2	31.9	29.5	32.7	28.8	29.4	16.5
MeT 6	19.7	21.9	31.5	34.2	30.2	33.1	32.2	34.3	34.2	28.9	20.9
MeT 12	14.7	16.7	32.6	23.4	37.5	38.8	36.1	28.9	31.6	24.9	15.4
MeT 2	16.5	23.2	31.1	30.8	33.6	38.8	40.2	36.6	34.4	27.6	18.9
MeT 15	17.2	20.1	26.4	27.8	28.5	31.2	32.7	30.4	37.5	24.8	16.9
MeT 14	13.7	19.9	19.7	22	21.6	21.4	22.5	23.4	23.9	29.7	14.8
MeT 9	16.3	17.7	30.1	22.7	27.8	31.8	27.9	25.1	25.6	22.2	17.1

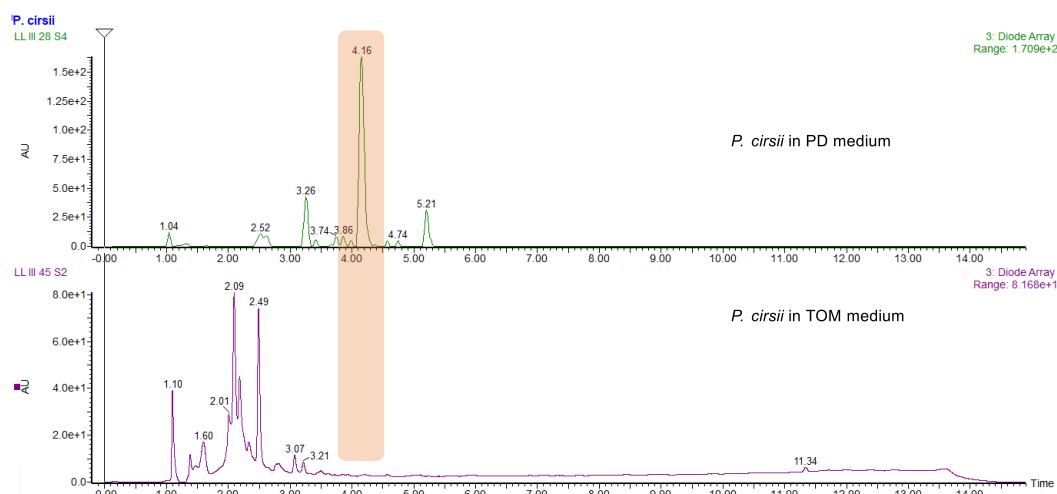
**Figure 3.4.3.2** Putative *O*-MeTs in *P. cirsii* and % identity to their closest relatives to putative *O*-MeTs in *T. reesei* *Δtmus53*; Those with low identity shaded green, while those with inferred homology shaded red.

### 3.5 Results-RT-PCR Analysis for the Interesting Genes from *P. cirsii*

In order to more clearly link the putative genes found by bioinformatic analysis with the biosynthesis of scytolide **158**, the expression levels of these genes under producing and non-producing condition were investigated by RT-PCR. It evaluated whether the observed differences of one gene expression level in two conditions were significant or in the range of natural variation.

The non-producing condition of *P. cirsii* was explored by growing the fungi in different liquid media. TOM medium was found to suppress scytolide production, which was then selected as non-producing medium for comparative transcriptome analysis (Figure 3.5.1). The producing condition

were previously described in chapter 3.2.

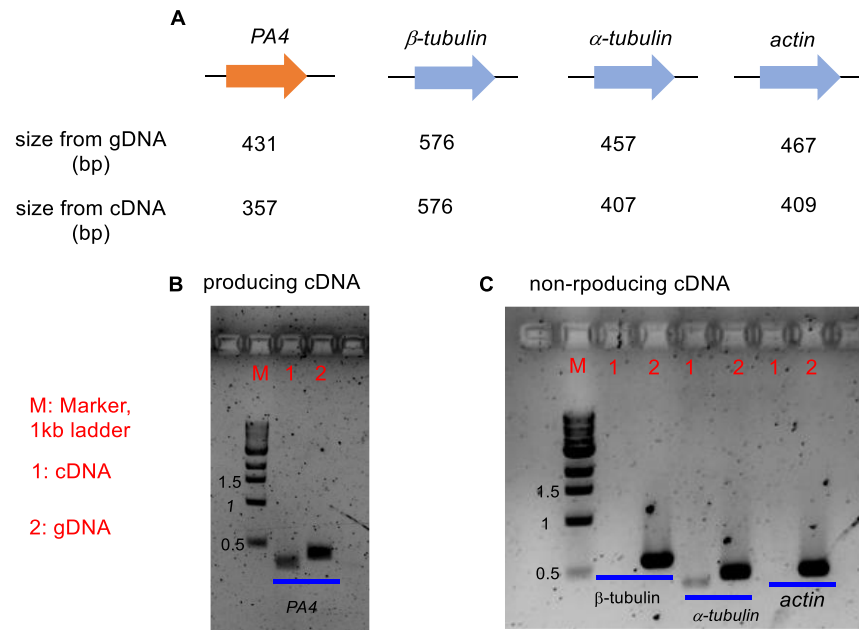


**Figure 3.5.1** DAD chromatograms of *P. cirsii* extracts grown under producing (upper) and non-producing (lower) conditions. Scytolide **158** is highlighted in color.

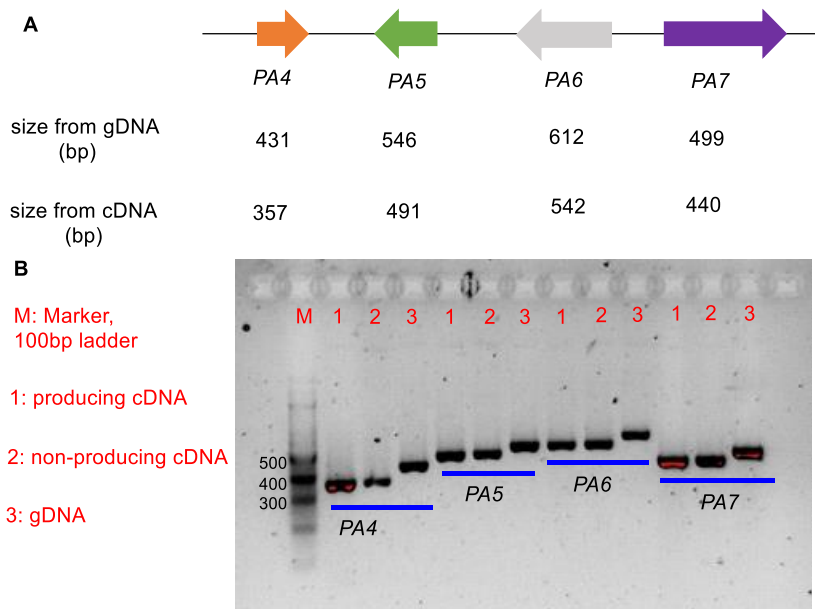
Liquid cultures for both conditions (PD medium, TOM medium) were inoculated with *P. cirsii* and cultured for 4 days. The messenger RNAs (mRNAs) *P. cirsii* mycelium in both PD medium and TOM medium were extracted using the ZR Fungal Bacterial RNA MiniPrep kit (Zymo Research) and converted into the corresponding complementary DNA (cDNA) with High Capacity RNA-to-cDNA Kit. In order to confirm that the cDNAs were made successfully, PCR experiments were carried out with the oligonucleotides targeting *PA4* from *scy PA* cluster and three housekeeping genes (Table 4.1). gDNA of *P. cirsii* was also analyzed with the same oligonucleotides as control group. Since the targeted genes contained introns, the amplification product from gDNA is longer than the product from cDNA (Figure 3.5.2.A). For cDNA from non-producing condition, PCR experiments were performed to amplify three housekeeping gene fragments that encodes respectively  $\beta$  tubulin,  $\alpha$  tubulin and actin, while for cDNA from producing condition, PCR experiment was performed to amplify *PA4* gene fragment. As expected, both producing cDNA and non-producing cDNA were obtained successfully (Figure 3.5.2.B).

To gain more insight into the *scy PA* BGC, RT-PCR experiments were started to check the expression of the 4 putative functional genes under both producing and non-producing condition. These genes encoded a methyltransferase, a DAHP synthetase, a non-heme iron  $\alpha$ -ketoglutarate dependent oxygenase and an acetyl-CoA synthetase-like protein (Figure 3.5.3.A). The results of RT-PCR in general show no significant differences in the expression levels of *PA4 - PA6* genes under two the different fermentation conditions. While the high expression level of all the genes indicated that *scy*

PA BGC is switched on under both conditions (Figure 3.5.3.B).

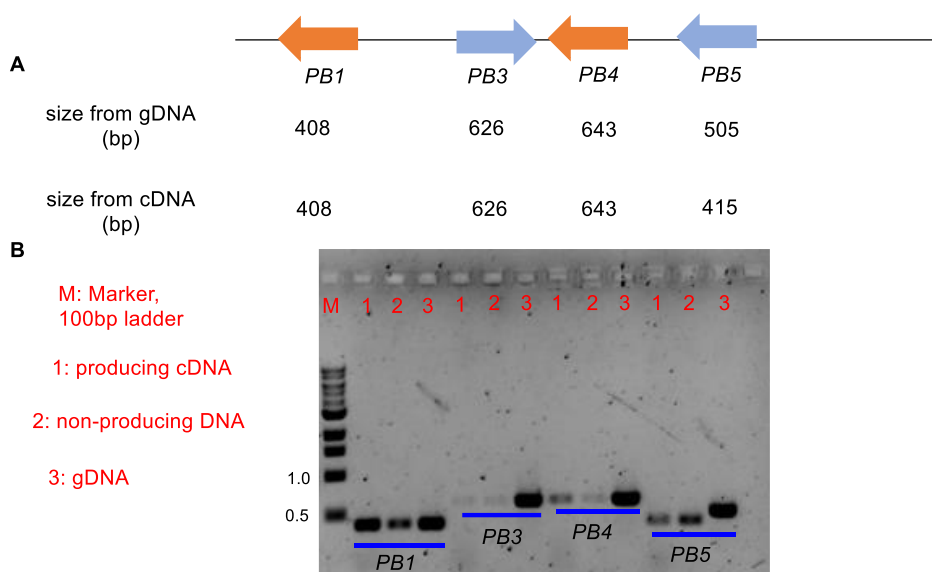


**Figure 3.5.2** PCR analysis of cDNA of *P. cirsi* in producing and non-producing condition: **A**, The expected sizes of PCR products from cDNA and gDNA; **B**, PCR analysis of producing cDNA from *P. cirsi*; **C**, PCR analysis of non-producing cDNA from *P. cirsi*.



**Figure 3.5.3** Reverse Transcriptase (RT) analysis of genes from scy PA BGC: **A**, The expected sizes of PCR products from cDNA and gDNA; **B**, RT-PCR analysis of targeted genes from producing cDNA, non-producing cDNA and gDNA respectively.





**Figure 3.5.4** Reverse Transcriptase (RT) analysis of genes from *scy PB* BGC: **A**, The expected sizes of PCR products from cDNA and gDNA; **B**, RT-PCR analysis of targeted genes from producing cDNA, non-producing cDNA and gDNA respectively.

In addition, expression of 4 putative functional genes from the *scy PB* BGC were also analyzed by comparing PCR product from producing cDNA and non-producing cDNA of *P. cirsii* (Figure 3.5.4). These genes encoded a dehydroquinase (*PB1*), a 3-dehydroshikimate dehydratase (*PB3*), a quinate permease (*PB4*) and a quinate dehydrogenase (*PB5*) respectively. However, no significant difference in expression was observed among all 4 targeted genes between producing and non-producing conditions. The low expression of *PB5* and no expression of *PB3* indicated that the quinate pathway was not activated under the tested conditions.

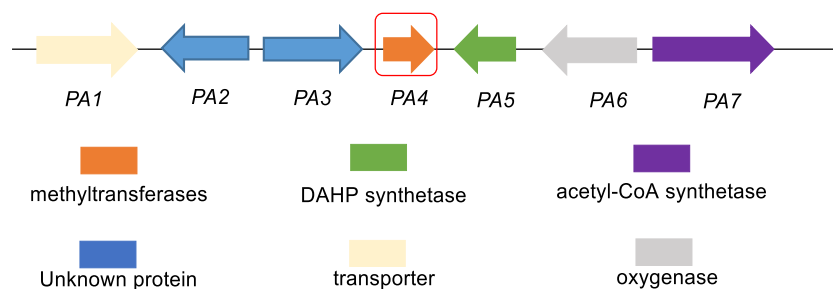
### 3.6 Results-Heterologous Expression of *PA4* in *A. oryzae*

Heterologous expression is a useful method for the characterisation of both enzyme function and biosynthetic gene clusters of secondary metabolites. Many natural product enzymes have been identified by heterologous expression, ranging from individual proteins to entire clusters.<sup>182–184</sup> Expression of a gene of interest in a heterologous organism provides a blank canvas to study the chemistry of a recombinant enzyme. *A. oryzae* is filamentous fungus that is commonly used as a heterologous host of secondary metabolites biosynthetic gene cluster, which are already applied

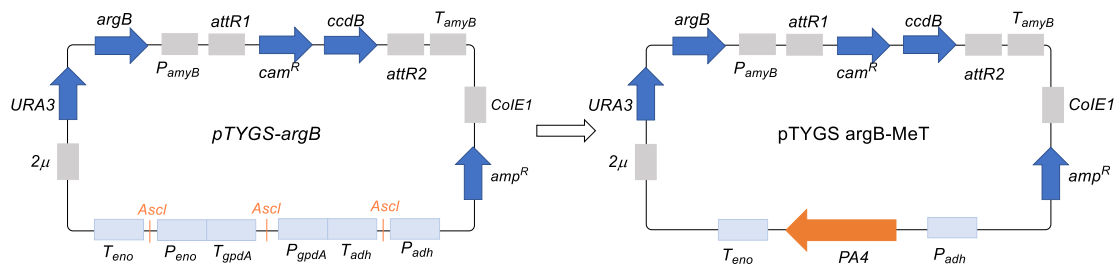
successfully in many examples.

Based on the hypothesis that scytolide **158** is a secondary metabolite that derives from the shikimic acid pathway and methyltransferase may possibly be the only necessary tailoring enzyme, an experiment of heterologous expression of methyltransferase (*PA4*) in *A. oryzae* NSAR1 was carried out. To verify this hypothesis, *PA4*, encoding a putative MeTase from the *scy PA* cluster, was first cloned into the pTYGSargB vector (Figure 3.6.1).

The pTYGSargB plasmid was used as the vector for *PA4* heterologous expression plasmid. *ArgB* was designed to complement the arginine-auxotrophy of *A. oryzae* NSAR1.<sup>138</sup> *PA4* was cloned to pTYGSargB plasmid by yeast recombination. Total mRNA was isolated from scytolide-producing cultures of *P. cirsi* and converted to single-stranded complementary DNA (cDNA) by reverse-transcriptase treatment. Primers that include 30 bp homologous overlap sequence from pTYGSargB vector were designed to amplify the intronless *PA4* sequence from total cDNA, while the pTYGSargB vector was digested with *AscI* enzyme; The vector and the DNA fragment were connected together in the specific site by homologous recombination in *S. cerevisiae*, which gave the desired pTYGSargB-*PA4* plasmid (Figure 3.6.2).<sup>185</sup> Finally, the plasmid was confirmed by DNA sequencing and transformed into *E. coli* CCDB cell for amplification.

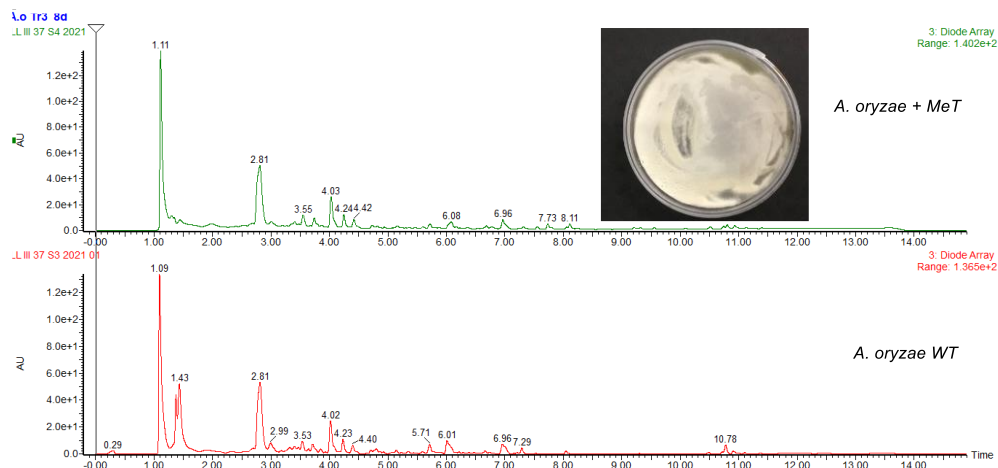


**Figure 3.6.1** *Scy PA* BGC, framed gene was selected for heterologous expression in *A. oryzae* NSAR1.

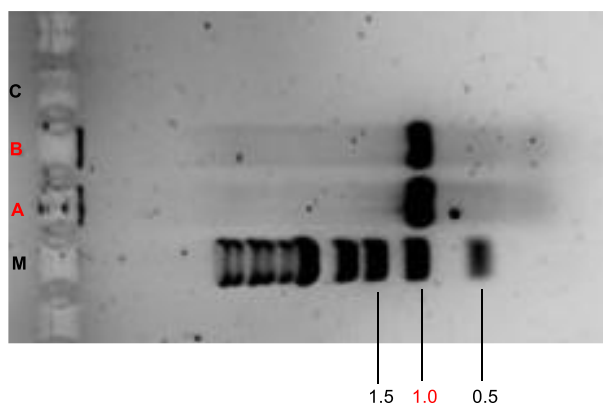


**Figure 3.6.2** Clone of *PA4* gene into *pTYGS-argB* vector by *S. cerevisiae*.

*A. oryzae* *NSARI* was transformed with the *PA4* expression-vector *pTYGS argB-MeT*. Seven resultant transformants were selected after growth on CZD/S agar 3 times, which were then transferred in DPY medium for metabolites analysis. After 8 days' cultivation, the obtained extracts from the metabolites of 7 different transformants and the wild type fungi were extracted respectively and submitted to analytical LCMS. In comparison to the UV chromatogram of *A. oryzae* *NSARI* wild type obtained under identical conditions, no new peaks were detected in 4 transformants (Figure 3.6.3). PCR analysis of extracted gDNA confirmed the presence of *PA4* in the genome of the *A. oryzae* transformants (Figure 3.6.4). These data indicated that only one methyltransferase together with common shikimate pathway gene were not enough for scytolide production.



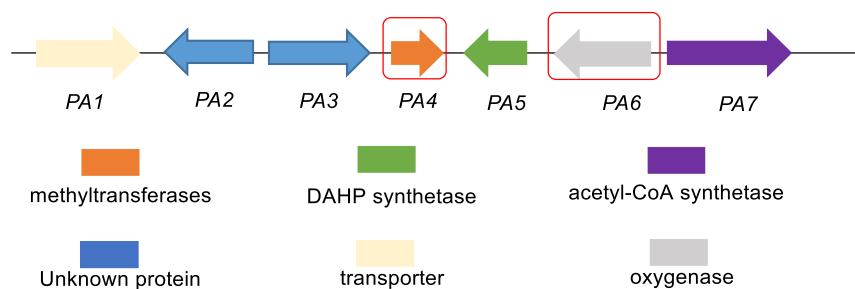
**Figure 3.6.3** DAD chromatogram of an *A. oryzae* WT and *A. oryzae + MeT* sample measured with analytical LCMS; New peaks are highlighted with red square.



**Figure 3.6.4** PCR analysis from gDNA of **A**, *A. oryzae-MeT* transformant 4; **B**, *A. oryzae-MeT* transformant 7 and **C**, *A. oryzae* WT. Expected size of amplicon in the presence of *PA4*, 900 bp.

### 3.7 Results-Targeted Gene Knockout in *P. cirsii*

Based on the bioinformatic findings, we designed two knock out cassettes to specifically target the *PA4* gene (encoding a putative methyltransferase) and *PA6* (encoding a non-heme iron  $\alpha$ -ketoglutarate dependent oxygenase) gene from *scy PA* BGC (Figure 3.7.1). The scytolide producing *P. cirsii* strain has been shown to efficiently produce protoplasts and was first successfully transformed with a  $\text{CaCl}_2/\text{PEG}$  method by Dr. Francesco Trenti from our group.<sup>142</sup> The strain was found to be sensitive to the antibiotic hygromycin B in concentrations of above 25  $\mu\text{g}/\text{mL}$ . Thus, a hygromycin resistance cassette with *hph* as selection marker was chosen for the transformation.

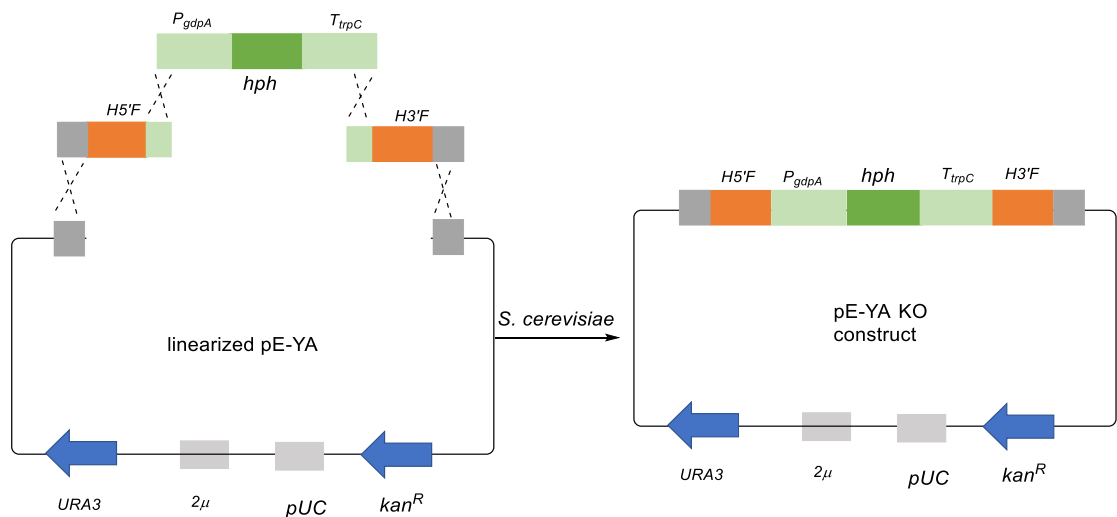


**Figure 3.7** *Scy PA* BGC, framed gene was selected for gene knockouts in *P. cirsii*.

#### 3.7.1 Plasmid Construction

For targeted gene disruption, KO cassette consisting of a selection marker (*hph*) flanked by arms homologous to the targeted gene was built exploiting yeast recombination. *S. cerevisiae* has been used extensively to build up vector constructs through PEG mediated co-transformation of individual

DNA fragments and a vector backbone (Figure 3.7.1). In this work, pE-YA was used as vector for the assembly of all KO constructs. This vector contains two selection markers: for revertant selection of uracil auxotroph yeast, and *kan<sup>R</sup>* for kanamycin selection during *E. coli* vector amplification. pE-YA was linearized by restriction hydrolysis with either *AscI* restriction enzyme or *NotI* restriction enzyme prior to PEG mediated co-transformation with individual DNA fragments. Three DNA fragments including the homologous 5' fragment (H5'F), the *hph* gene flanked by *A. nidulans* promoter/terminator *P<sub>gdpA</sub>/T<sub>trpC</sub>* and homologous 3' fragment (H3'F) were also necessary for the KO cassette constructions. H5'F and H3'F were generated by Q5 PCR from gDNA of *P. cirsii* using tailed primers with 30 bp overlapping to the up and downstream fragment separately, which were responsible for the site-specific homologous recombination in the gene knockout experiments. The selection marker *hph* was generated from high fidelity Q5 PCR, using pTH-GS-eGFP as the PCR template. Assembled vector DNA was isolated from *S. cerevisiae* and directly used to transform competent *E. coli*. The obtained *E. coli* colonies were screened by PCR and further used for plasmid amplification. A total of 2 constructs were generated by yeast recombination: namely pE-YA-PA4 KO and pE-YA-PA6 KO cassettes.

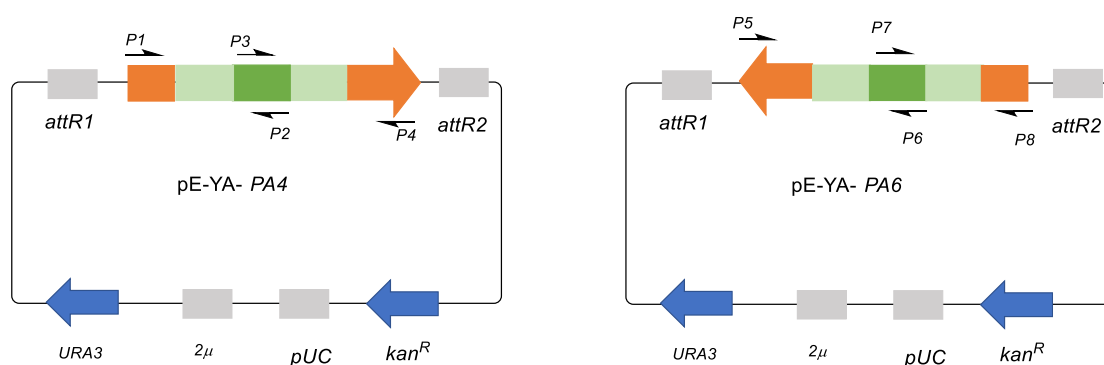


**Figure 3.7.1.** KO cassette construction into pE-YA vector by *S. cerevisiae* recombination.

### 3.7.2 Gene Knockout Targeting *PA4* and *PA6*

Next, the constructed pE-YA-*PA4* and pE-YA-*PA6* KO cassettes were separately used for gene

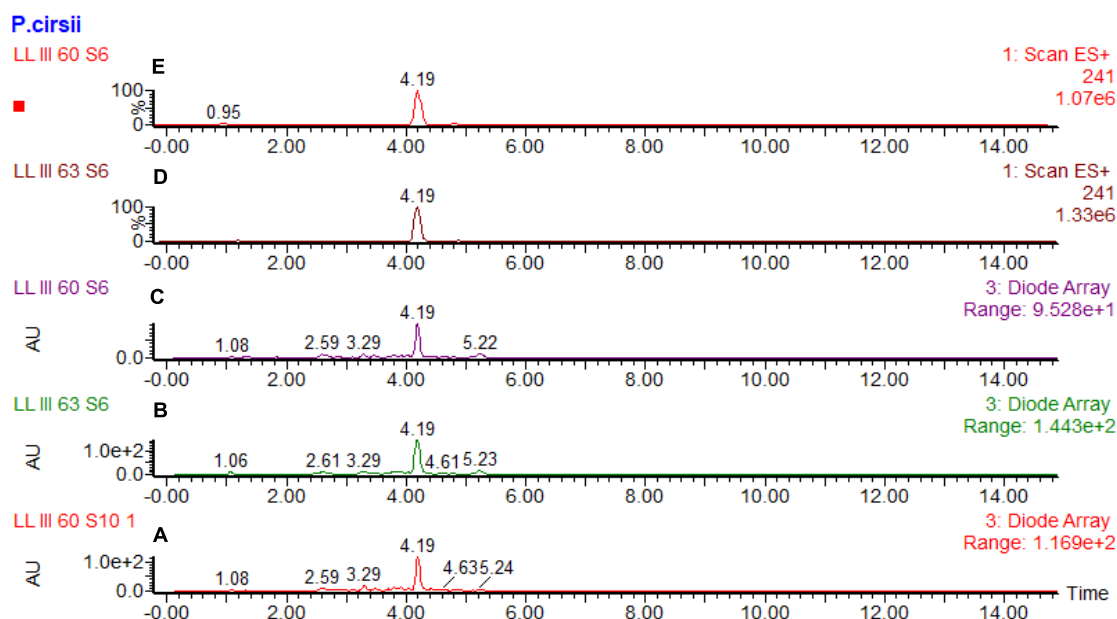
knockouts with the bipartite method mentioned before (chapter 1.5.2). The 5' and 3' bipartite substrates for *PA4* and *PA6* were amplified from pE-YA-*PA4* and pE-YA-*PA6* KO cassettes by Q5 PCR with oligonucleotides *P1+P2*, *P3+P4*, *P5+P6* and *P7+P8* (Figure 3.7.2.1). 2 µg of both parts were transformed together into freshly prepared protoplasts of *P. cirsii* WT following a protocol of PEG transformation procedure. In total 47 transformant targeting *PA4* and 31 transformants targeting *PA6* were obtained from three rounds of antibiotic selection. (Table 3.7.2). These transformants were grown individually on PD agar plates and scytolide production medium in order to be analyzed genetically and chemically.



**Figure 3.7.2.1** KO cassettes targeting *PA4* and *PA6* in pE-YA for bipartite PCR.

Target gene	Antibiotic selection	KO plasmid for PCR	Ttransformants	Chemically analyzed transformants	Genetically analyzed transformants
<i>PA4</i>	Hyg B	pE-YA- <i>PA4</i>	47	47	11
<i>PA6</i>	Hyg B	pE-YA- <i>PA6</i>	31	31	8

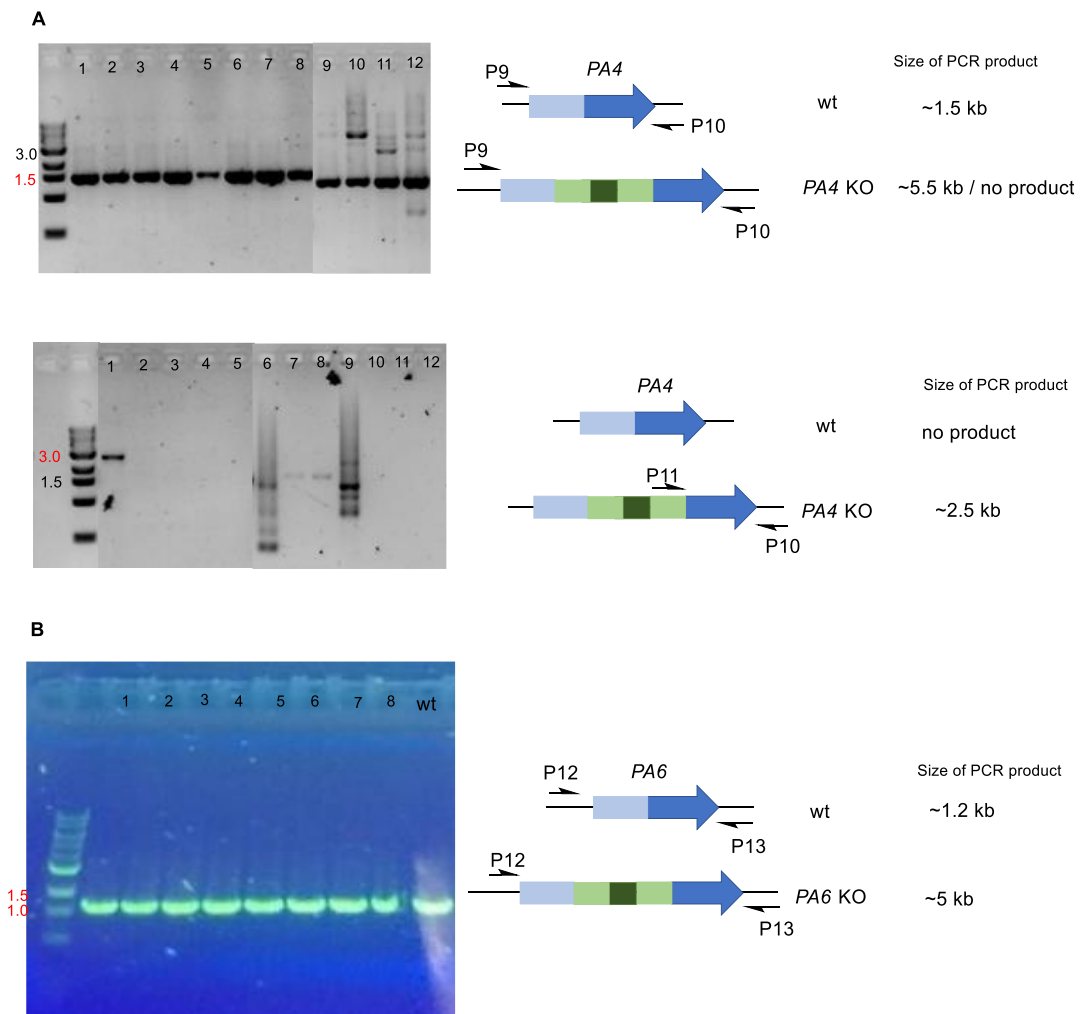
**Table 3.7.2** Overview of transformants obtained by bipartite substrate KO and chemical or genetic analysis.



**Figure 3.7.2.2** LCMS analysis of two transformants obtained with the bipartite marker strategy by analytical LCMS: **A**, UV chromatogram of *P. cirsií* WT; **B**, UV chromatogram of *PA4* KO transformant; **C**, UV chromatogram of *PA6* KO transformant; **D**, Extracted ion (ES+) chromatograms at  $m/z = 241$  for *PA4* KO transformant; **E**, Extracted ion (ES+) chromatograms at  $m/z = 241$  for *PA6* KO transformant;

For chemical analysis the 47 *PA4* KO transformants and 31 *PA6* KO transformants were grown under scytolide producing conditions. Metabolites were extracted following the protocol mentioned before (chapter 3.2) and analysed by analytical LCMS. Extracted ion chromatograms using 241 Da (ES+ spectrum) together with retention time led to the identification of scytolide **158** in all chromatograms (Figure 3.7.2.2).

For genetic analysis the gDNA of the 12 *PA4* KO transformants and 8 *PA6* KO transformants were extracted and used as Tag diagnostic PCR template. Disruption of genes was analysed by wPCR and extPCR (section 1.5.2) with oligonucleotides specific for the targeted genes. The results indicated that both non-disrupted *PA4* and *PA6* were amplified from all tested transformants while only one *PA4* KO transformant give extPCR product with expected size (Figure 3.7.2.3).



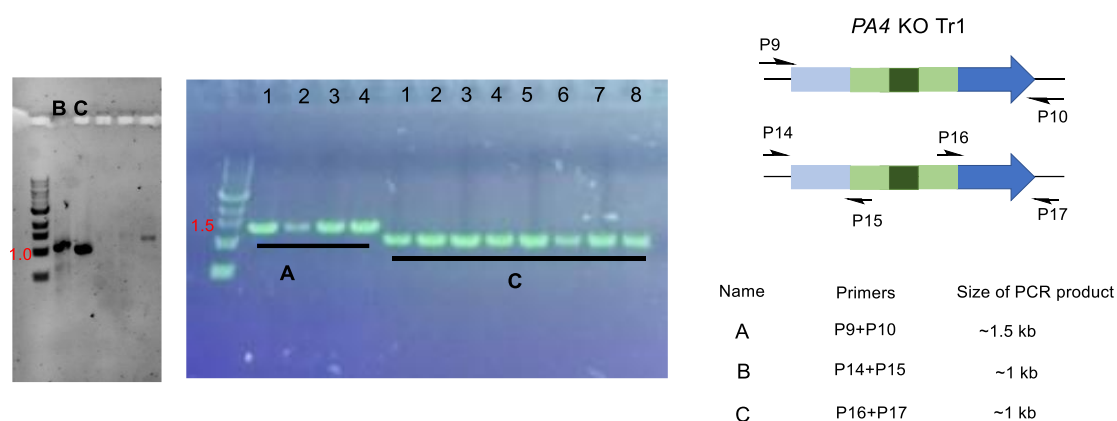
**Figure 3.7.2.3** Genetic analysis of the transformants obtained with bipartite gene knockout strategy: **A**, 12 *PA4* KO transformants analyzed with diagnostic PCR; **B**, 8 *PA6* KO transformants analyzed with diagnostic PCR.

Based on the genetic analysis of the transformants, the successful integration rate of the 5' and 3' bipartite substrates were low, with no true mutant observed except a mixture of wild type and real mutant, *PA4* KO Tr1. In order to obtain pure mutant, the spore solution of this transformants was diluted with water. 4 ml H<sub>2</sub>O was added to a fungi plate on PD agar to collect enough spores which was then filtered with miracloth to get clean spore solution. It was then diluted 20, 40, 80 and 100 times respectively. The diluted spore solutions were then cultivated on PD agar to get single colonies. 8 single colonies resulted from the dilution experiment of *PA4* KO Tr15 were further analyzed with diagnostic wPCR and exPCR, but no colonies were observed as a pure mutant without *PA4* (Figure 3.7.2.5).





**Figure 3.7.2.4** Single colonies of PA4 KO Tr1.



**Figure 3.7.2.5** Genetic analysis of the transformants with PA1 KO Tr1 after dilution experiments.

In summary, all chemically analysed transformants still produced scytolide and all genetically tested transformants possessed a non-disrupted *PA4* and *PA6*. This indicates that ectopic integration of bipartite substrate cassette based on non-homologous end joining is very common and effective in *P. cirsii* transformation. The previous attempts for gene disruption in *P. cirsii* within our group also had a very low success rate. Only two true KO mutants of the core *phyS* PKS-NRPS and the P450 cytochrome oxidase *phyL6* was identified from 101 transformants targeting 6 different genes from the BGC of phyllostictine A **196**.<sup>186</sup> This indicates that gene disruption in fungi is also locus dependent. In order to improve the successful KO rate in *P. cirsii*, knocking out the gene responsible for non-homologous end joining recombination and other 5 genes from *scy PA* cluster should be further investigated.

### 3.8 Results-Targeted Gene Knockout in *T. reesei* $\Delta$ *tmus53*

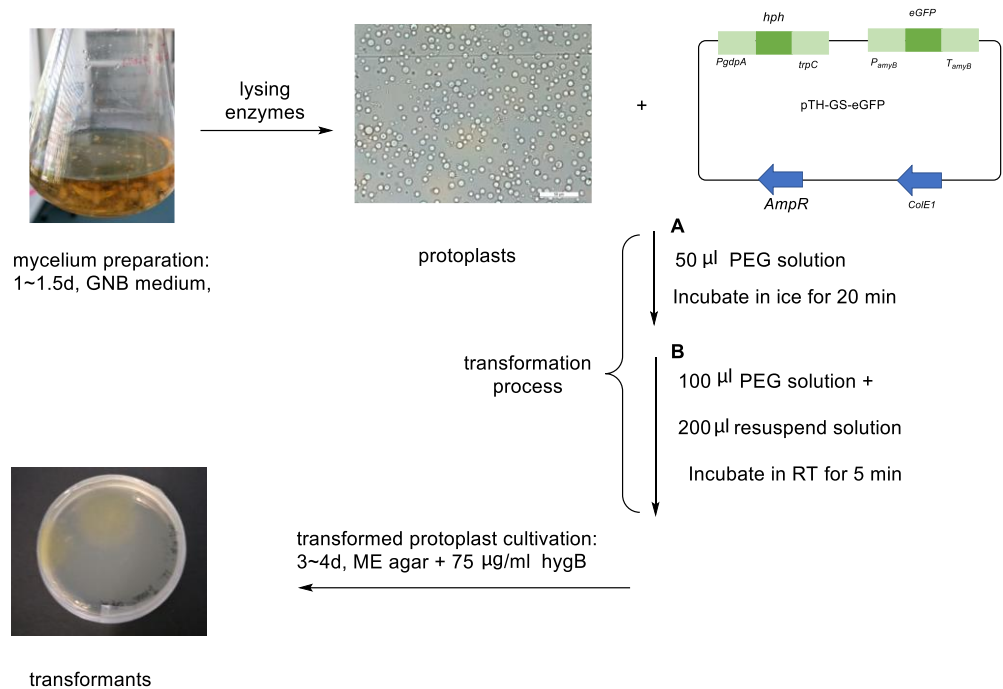
While gene knockout had previously failed in *P. cirsii*, which was found to have a prohibitively high rate of non-homologous recombination, it had not been attempted in *T. reesei*  $\Delta$ *tmus53*. The fungus is well known for efficient production of cellulases, which converts lignocellulosic biomass to monomeric sugars for biofuels or platform chemicals. Because of this enormous protein secretory capacity, this fungus appears to be an attractive host for genetic engineering. A number of *T. reesei* mutant strains secreting very large amounts of protein into the medium have been obtained using the direct approach of mutation and screening.<sup>187</sup> Practical basis of genetic engineering requires the development of a transformation system that must meet certain requirements: (i) a high rate of homologous integration (gene replacement) events; (ii) prototrophic strains for simple subsequent handling, and (iii) strains which are isogenic compared to the parental strain for subsequent investigations in a correct genetic background. Up to now, *T. reesei* can be transformed using different techniques, such as polyethylene glycol (PEG)-mediated transformation of protoplasts, particle bombardment, electroporation, and *Agrobacterium tumefaciens*-mediated transformation.<sup>188</sup>

Efficient PEG mediated transformation of *T. reesei* was first reported in 1987.<sup>189</sup> A plasmid carrying the dominant selectable marker *amdS* or the *argB* gene of *Aspergillus nidulans* was integrated into the gDNA of *T. reesei* complementing the respective *argB* mutation of *T. reesei*. After that, a novel vector containing the *E. coli* hygromycin B phosphotransferase gene (*hph*) fused between promoter of *P<sub>pki1</sub>* and terminator of *T<sub>cbh2</sub>* from *Trichoderma* was expressed successfully in *T. reesei* using a similar transformation protocol.<sup>190</sup> In 2011, a highly efficient gene targeting transformation system was developed using *T. reesei* without *tmus53* gene (*T. reesei*  $\Delta$ *tmus53*). *Tmus53* (a homolog of human Lig4) was shown to be required for non-homologous end joining (NHEJ) in *T. reesei*.<sup>168,191</sup> The deletion cassette targeting this gene contained the phosphinothricin resistance marker (*bar*) flanked by 1.5 kb fragments from up- and downstream of *tmus53*. Transformation of *T. reesei* with the linear KO cassette created a NHEJ-deficient *T. reesei* strain that exhibited efficient integration at the homologous site.<sup>168</sup> These developments facilitated not only producing the auxotrophic strain that allows multiple genomic manipulations but also site-directed gene deletion for identifying the function of targeted gene.

### 3.8.1 Transformation Protocol

*T. reesei*  $\Delta$ tmus53 was obtained from Cox group member Ms. Eman F. Bassiony. Transformation of this fungus followed the same protocol as that described in the literature.<sup>189</sup> The cell wall of young mycelia was digested by a mixture of lysing enzymes consisting of 10 mg/ml *Trichoderma* lysing enzyme and 5 mg/ml driselase in protoplast solution (1.2 M sorbitol, 100 mM potassium phosphate, pH 5.6). The plasmid transformed into *T. reesei*  $\Delta$ tmus53 is pTH-GS-eGFP, it contains the enhanced Green Fluorescent Protein reporter gene (*eGFP*) and the *hph* gene for selection with hygromycin B. Unfortunately, no colonies were obtained in the ME agar with 75  $\mu$ g/ml hygB from *T. reesei*  $\Delta$ tmus53 transformation with the reported protocol. The negative control (protoplasts without exogenous DNA incubated in selection plates) also failed to produce colonies while the positive control (protoplast without exogenous DNA incubated in the normal plates) grew well. This result probably indicated the quality of protoplast was good while the transformation process did not work properly. Therefore, several changes based on the original transformation process were made to generate colonies on selection plates, including testing different protoplast concentration and different volume of transformation solution (Table 3.8.1). Finally, *T. reesei*  $\Delta$ tmus53 PEG/CaCl<sub>2</sub> mediated protoplast transformation with pTH-GS-eGFP resulted in 8 colonies growing on tertiary selection plates based on the optimized transformation protocol indicated in scheme 3.8.1.

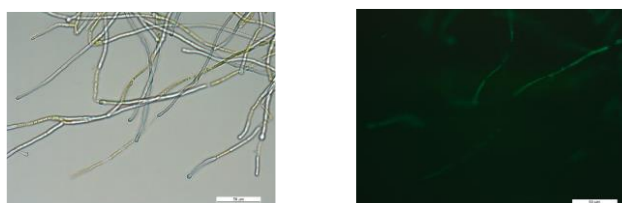
The *eGFP* gene is expressed under the control of the starch or maltose inducible *P<sub>amyB</sub>* promoter and terminator from *A. oryzae*, which causes green fluorescence on mycelia transformants when it is visualized under a fluorescence microscope. For fluorescence analysis eight transformants were sub-cultured in ME medium, which contains maltose as the carbon source for the induction of *P<sub>amyB</sub>*. The mycelia of 4d old transformants and *T. reesei*  $\Delta$ tmus53 WT were microscopically analyzed. Upon excitation (450 – 490 nm) the transformants displayed a green fluorescence correlating to the expression of *eGFP* and translation into protein (Figure 3.8.1), whereas wildtype control does not emit green light. This confirmed that an optimized transformation protocol was established, which enabled further genetic manipulation of *T. reesei*  $\Delta$ tmus53 by bipartite knock out strategy. It also showed that *P<sub>amyB</sub>* from *A. oryzae* is active in *T. reesei*  $\Delta$ tmus53



**Scheme 3.8.1** Optimized PEG-mediated transformation protocol of *T. reesei*  $\Delta$ tmus53.

Transformation protocol	Protoplast concentration	transformation solution added in the process B	Obtained Transformants
Standard protocol	$10^7$	2ml of PEG-buffer and 4 ml resuspension solution	0
Optimized protocol	$10^6$	100 μl of PEG-buffer and 200 μl resuspension solution	8

**Table 3.8.1** Difference between Standard transformation protocol and optimized transformation protocol for *T. reesei*  $\Delta$ tmus53.

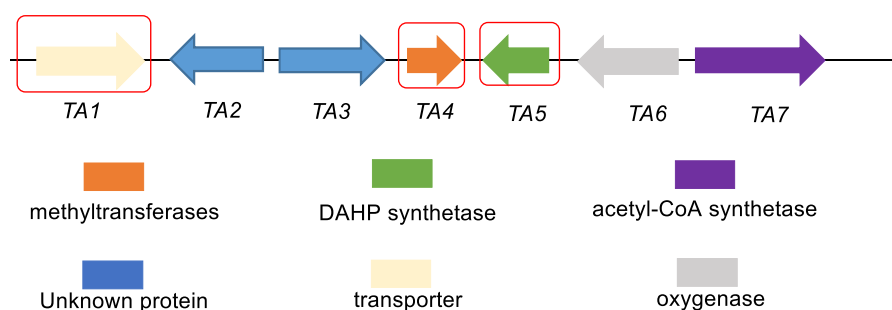


**Figure 3.8.1** Mycelia of *T. reesei*  $\Delta$ tmus53 with pGS-TH-eGFP under microscope with white light (left) and UV light (right).

### 3.8.2 Gene Knockout Targeting *TA1*, *TA4*, and *TA5*

*Scy TA* BGC from *T. reesei Atmus53* is the homologous gene cluster of *scy PA* BGC (Figure 3.8.2.1, Figure 3.4.2.2, Section 3.4.2). Putative annotations of all the genes from *scy TA* BGC were predicted by running protein blast in NCBI (Table 3.8.2), which were the same with that of the corresponding genes from *scy PA* BGC as expected (Table 3.4.2.3). Initially, *TA4*, a putative methyltransferase, was chosen for gene disruption. A knock-out (KO) cassette targeting *TA4* composed of *hph* gene flanked by two homologous DNA fragments was constructed by the protocol described before (Figure 3.7.1, section 3.7.1). If bipartite-knockout strategy was effective in *T. reesei Atmus53*, and *TA4* was found to be necessary for scytolide biosynthesis, then the function of the other clustered genes and a more persuasive picture of the biosynthetic pathway would be characterized.

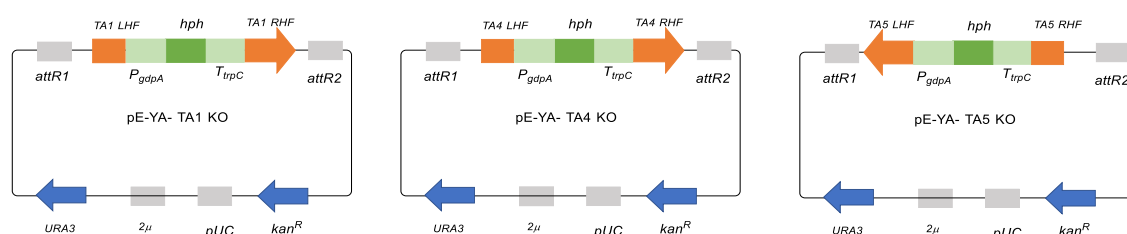
Protoplasts of *T. reesei Atmus53* were transformed with 5' and 3' bipartite DNA fragments based on both standard protocol and the optimized protocol for more than 3 times respectively. However, no transformant colonies were produced on ME agar with 75 µg/ml hygB. This was probably caused by short overlapping DNA fragments (500 bp) between 5' and 3' bipartite DNA fragments and the targeted gene. It was also possible that *TA4* was really difficult to destroy since gene disruption in fungi is sometimes locus dependent. To address this, three new KO cassettes targeting *TA1*, *TA4* and *TA5* with 1kb homologous arms were constructed using the yeast recombination (Figure 3.8.2.2). New fragments up- and downstream of *TA1*, *TA4* and *TA5* were respectively obtained by Q5 PCR from gDNA of *T. reesei Atmus53* and used as the homologous arms in the KO cassettes.



**Figure 3.8.2.1** *Scy TA* BGC, framed gene was selected for gene knockouts in *T. reesei Atmus53*

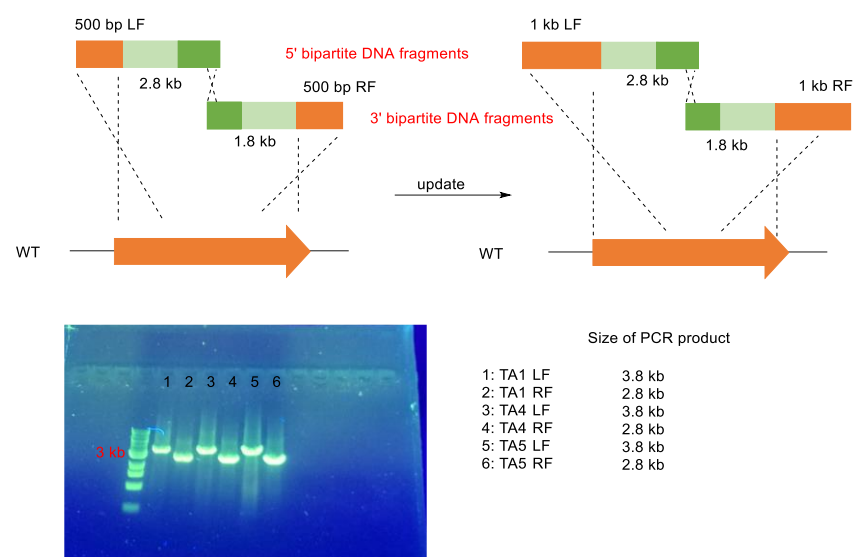
Name and Locus_tag	Annotation	Near Pblast hit
TA1 (TRIREDRAFT_80973)	putative ATP-binding multidrug cassette transport protein	<i>Fusarium napiforme</i> FNAPI_6401
TA2 (TRIREDRAFT_110658)	uncharacterized protein	<i>Trichoderma parareesei</i> A9Z42_0066530
TA3 (TRIREDRAFT_28185)	Hypothetical protein	<i>Trichoderma parareesei</i> A9Z42_0066770
TA4 (TRIREDRAFT_110660)	methyltransferase family protein	<i>Fusarium beomiforme</i> FBEOM_12189
TA5 (TRIREDRAFT_67290)	DAHPh synthetase	<i>Trichoderma longibrachiatum</i> ATCC 18648 M440DRAFT_1389792
TA6 (TRIREDRAFT_123431)	non-heme iron $\alpha$ -ketoglutarate dependent oxygenase	<i>Fusarium tjaetaba</i> FTJAE_10257
TA7 (TRIREDRAFT_110663)	acetyl-CoA synthetase-like protein, ATP-binding enzyme	<i>Trichoderma citrinoviride</i> BBK36DRAFT_1134399

**Table 3.8.2** Annotations of *scy* TA cluster in *T. reesei*  $\Delta$ mus53

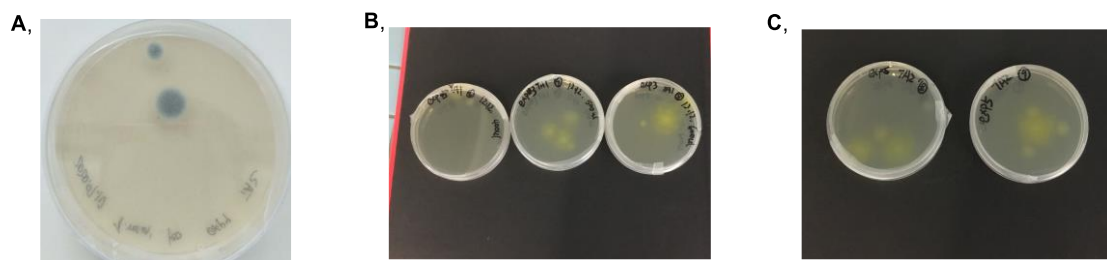


**Figure 3.8.2.2** Constructed pE-YA KO cassettes targeting TA1, TA4 and TA5.

Further bipartite knock-out in *T. reesei*  $\Delta$ mus53 with 5' and 3' bipartite DNA fragments from the three new KO cassette was carried out respectively (Figure 3.8.2.3). After growing the transformed protoplast on ME agar with 75  $\mu$ g/ml hygB for 4 days, attempted knockout of TA5 led to two unusual phenotype transformants (Figure 3.8.2.4A), which was then identified to be contaminating fungi *Penicillium chrysogenum* based on ITS sequence analysis. Besides, a few transformants were obtained in the knock out experiments targeting TA4 and TA5 (Figure 3.8.2.4B-C), but they did not grow on the second antibiotic selection plate. Thus, no *T. reesei*  $\Delta$ mus53 transformant was obtained from bipartite gene-knock targeting TA1, TA4 and TA5 with both standard protocol and the optimized protocol.



**Figure 3.8.2.3** Longer 5' and 3' bipartite DNA fragments for *T. reesei*  $\Delta$ *tmus53* knock-out targeting *TA1*, *TA4* and *TA5*.



**Figure 3.8.2.4** Transformants obtained from *T. reesei*  $\Delta$ *tmus53* bipartite knocking out experiments: **A,** *Penicillium chrysogenum*; **B,** *TA4* KO transformants; **C,** *TA5* KO transformants.

### 3.9 Conclusion and Outlook

The major focus of this project was the identification and investigation of a BGC involved in scytolide biosynthesis. For this reason, genome, transcriptome and secondary metabolites of *P. cirsii* and *T. reesei*  $\Delta$ *tmus53* were analysed in detail.

In our hands, *P. cirsii* produced the major compound scytolide **158** when cultured in PD medium. Scytolide **158** was not stable in methanol and was spontaneously converted to methylscytolide **198**. These two compounds were then isolated and elucidated. The production of scytolide **158** was also observed when growing *T. reesei* *Δtmus53* in ME medium at 4-5 days, and it was gradually consumed while the spirosorbicillinols A **160** and B **161** accumulated from day 6.

Isotope feeding experiments were carried out by supplementing *P. cirsii* with [2-<sup>13</sup>C]-glycerol **187** as carbon source in producing condition. The <sup>13</sup>C labelling positions in scytolide **158** indicated that it was an unusual secondary metabolite derived from PEP **120**, which is formed via glycolytic pathway starting from glycerol **187**. It was the first experiment confirming that scytolide **158** was derived from the shikimate pathway, not only just based on the structure similarities between scytolide **158** and EPSP **127**.

Next, bioinformatic analysis of the *T. reesei* *Δtmus53* and *P. cirsii* genomes had enabled identification of one promising gene cluster (*scy PA* cluster in *P. cirsii* or *scy TA* cluster in *T. reesei* *Δtmus53*) containing a DAHP synthase (*PA5*) and a methyltransferase (*PA4*) that possibly related to scytolide biosynthesis.

Looking for the homologous genes of pentafunctional aromatic polypeptide in *T. reesei* *Δtmus53* enabled the rapid identification of three putative BGCs that might encode scytolide biosynthesis. Further searching for the homologous genes of these three putative clusters in *P. cirsii* resulted in prioritization two putative gene clusters for scytolide biosynthesis, namely the *scy PA* and *scy PB* clusters. Putative annotations of all the genes from the *scy PA* and *scy PB* clusters were predicted by running protein BLAST in NCBI. The *scy PB* cluster could be easily ruled out as it contained all the expected quinate pathway genes, thus was identified as the *qut* BGC in *P. cirsii*. This left the *scy PA* cluster is the most promising candidate to encode scytolide biosynthesis based on protein blast and Artemis analysis between *scy PA* cluster in *P. cirsii* and *scy TA* cluster in *T. reesei* *Δtmus53*.

At the same time, looking for the homologous genes of methyltransferase resulted in 25 putative *O*-MeT genes in *T. reesei* *Δtmus53* and 19 putative *O*-MeT genes homologues *P. cirsii*, including an interesting gene, *TRIREDRAFT\_66059*, located near *scy PB* BGC. However, the promising methyltransferases encoded by *PA4/TA4* from *scy PA/TA* BGC were not included among these homologous genes. Further homology analysis among all the *O*-MeT genes including *PA4/TA4* were



carried out by multiple sequence alignment. The results show that *PA4/TA4* are the most closely-related pair of MeTs from two different fungi.

The expression level of the *scy PA* cluster, *scy PB* cluster were investigated using RT-PCR under both producing and non-producing conditions. It was strange that only one housekeeping gene was successfully amplified from non-producing cDNA of *P. cirsi* among 3 tested genes. The reason might be the small gene variations happened in the housekeeping gene after the longtime cultivation of the strain, even though the ITS sequence indicated that it was still *P. cirsi*. Besides, no obvious difference of the tested genes expression level was observed between producing condition and non-producing condition. However, all probed genes from *scy PA* cluster were confirmed to be transcribed at high levels, compared to the gene expressions from *scy PB* cluster and the three putative methyltransferases. This provided subtle evidence for its participation in the biosynthetic pathway of scytolide **158**. Our experiments were designed based on the proposal that production or non-production of scytolide was related only to gene expression. This assumption may be wrong as the production might be controlled by substrate availability or other factors.

Based on the proposal that scytolide **158** was a primary metabolite that directly derived from the shikimate pathway and only one methyltransferase was needed as the tailoring enzyme, the heterologous expression of *PA4* from *scy PA* BGC in *A. oryzae* NSARI was performed. LCMS chromatograms of the transformants showed no new peak compared to that of *A. oryzae* WT. This result indicated that scytolide was not a primary metabolite and clustered genes together with partial genes from the shikimate pathway were essential for scytolide production.

Both *T. reesei* *Δtmus53* and *P. cirsi* are hygromycin sensitive and can be transformed by PEG-mediated method according to the literature.<sup>102,142</sup> Based on the BGC prediction and the experimental results from RT-PCR and gene heterologous expression experiments, several bipartite knockout experiments targeting genes from *scy PA* and *scy TA* clusters were carried out. Unfortunately, the knockout of *PA4* and *PA6* in *P. cirsi* and *TA1*, *TA4* and *TA5* were not successful. In *P. cirsi* KO experiments, plenty of transformants was obtained, but none of them were real mutants with the deletion of targeted gene based on the genome analysis with wPCR and extPCR. This indicated that ectopic integration of bipartite substrate cassette based on non-homologous end joining was very common in *P. cirsi*. In order to improve KO rates, knocking out the gene

responsible for non-homologous end joining recombination and constructing the KO cassette with long homologous arms of the targeted gene should be tried in the future gene deletion experiments. Since the gene disruption in fungi is also locus dependent, knocking out the remaining five genes from *scy PA* cluster is also deserved to be further investigated.

The situation in *T. reesei Atmus53* KO was really different in that no transformant was obtained. The attempts for gene disruption with the constructed KO cassette containing long homologous arms did give some transformants. Unexpectedly, all these transformants failed to grow on the second selection agar with *hygB*, which might be caused by the low expression of *hph* gene. It was recommended that change the promoter of *hph* would give a better result in *T. reesei Atmus53* transformation, as the transformation frequencies was greatly improved when replacing the promoter of *hph* ( $P_{gdpA}$ ) with the promoter of *pkil* from *Trichoderma* ( $P_{pkil}$ ) according to literature.<sup>190</sup> The contamination of *Penicillium chrysogenum* colony during transformation process also indicated that exogenous DNA fragment was difficult to be transformed into *T. reesei Atmus53* protoplast.

Based on the labeling studies and the detailed analysis of the *scy PA* and *scy TA* biosynthetic gene cluster we could propose a possible biosynthetic pathway for the synthesis of scytolide **158**. (Scheme 3.4.2.1). In the first step, DAHP **122** is produced from PEP **120** and E4P **121** catalyzed by DAHP synthase encoded by *PA5*. The AROM protein from the primary metabolism shikimate pathway then catalyzes several consecutive chemical reactions to form EPSP **127** using DAHP **122** as substrate. Dephosphorylation of EPSP **127** and epimerization of hydroxyl group at C-8 then forms intermediate **206**, which is possibly catalyzed by the *PA2/PA3* encoded unknown proteins. Alternatively, epimerization of hydroxyl group at C-8 might follow after dephosphorylation of EPSP **127** and catalyzed by another enzyme. Formation of the second key intermediate **207** could involve an intramolecular lactone ring cyclization catalyzed by ATP dependent enzyme encoded by *PA7*. Finally, methylation of the carboxyl group would give scytolide **158**, possibly by methyltransferase *PA4*.

In conclusion, we were able to identify that scytolide **158** is a secondary metabolite most likely derived from the shikimate pathway. A promising BGC for scytolide and a novel biosynthetic pathway partly related to shikimate pathway (Scheme 3.4.2.1) were proposed and found in both *P. cirsii* and *T. reesei Atmus53*. However, the reliable evidence that link *scy PA/scy TA* cluster to scytolide production are still not available. In the future, identification of the right gene cluster for

scytolide will be investigated by co-expressing all the seven genes from *scy PA* cluster in *A. oryzae*. While knocking out experiments with bipartite fragments containing longer homologous arms in *P. cirsii* would be an alternative protocol to identify this gene cluster. Identification of the right BGC will enable clarifying the chemical reaction catalyzed by potentially functional enzyme in the biosynthetic pathway.

## 4 Experimental

All chemicals, materials and enzymes were purchased from one of the following companies: Sigma Aldrich (Steinheim), Thermo Fisher Scientific (Waltham, MA, USA), New England Biolabs (Beverly, MA, USA), Amresco (Solon, OH, USA), Applichem (Darmstadt), Bio-Rad (Munich), Roth (Karlsruhe) and Stratagene (Amsterdam, Netherland).

### 4.1 Growth Media, Buffers and Solutions

All media, buffers and solutions were prepared with millipore water (GenPure Pro UV/UF milipore device, Thermo Scientific) or distilled water. Growth media and transformation solutions were sterilised at 121 °C for 15 min (Autoclave 2100 Classic, Prestige Medical) or by disposable sterile filter (0.45 µm pore size, Roth).

**Table 4.1.1** Media and agar used in this work for *E. coli* and *S. cerevisiae*.

Media	Composition
LB	0.5 % (w/v) NaCl, 1 % (w/v) Tryptone, 0.5 % (w/v) Yeast extract
LB Agar	LB Medium, 1.5 % (w/v) Agar
SOC	0.5 % (w/v) Yeast extract, 2 % (w/v) Tryptone, 0.06 % (w/v) NaCl, 0.02 % (w/v) KCl
2TY	1 % (w/v) Yeast extract, 1.6 % (w/v) Tryptone, 0.5 % (w/v) NaCl
TB	2,40 % (w/v) Yeast extract, 1,20 % (w/v) Tryptone, 0,40 % (w/v) Glycerol (Roth), To be added (to 900 mL TB): 10 % (v/v) 10X KPI Puffer (100 mL)
YPAD	1 % (w/v) Yeast extract, 2 % (w/v) Tryptone, 2 % (w/v) D(+)-Glucose Monohydrate, 0.03 % (w/v) Adenine
YPAD Agar	YPAD Medium, 1.5 % (w/v) Agar
SM-URA Agar	0.17 % (w/v) Yeast nitrogen base, 0.5 % (w/v) (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> , 2 % (w/v) D(+)-Glucose Monohydrate, 0.077 % (w/v) Complete supplement mixture minus Uracil, 1.5 % (w/v) Agar

**Table 4.1.2** Media and agar used in this work for *A. oryzae*.

<b>Media</b>	<b>Composition</b>
CZD/S agar	3.5 % (w/v) Czapek Dox broth, 18.22 % (w/v) D-sorbitol, 0.1 % (w/v) (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> , 0.15 % (w/v) L-methionine, 0.05 % adenine, 1.5 % (w/v) agar
CZD/S soft agar	3.5 % (w/v) Czapek Dox broth, 18.22 % (w/v) D-sorbitol, 0.1 % (w/v) (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> , 0.15 % (w/v) L-methionine, 0.05 % adenine, 0.8 % (w/v) agar
CZD/S agar + argenine (w/o L-Methionine)	3.5 % (w/v) Czapek Dox broth, 18.22 % (w/v) D-sorbitol, 0.1 % (w/v) (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> , 0.15 % (w/v) argenine, 0.05 % adenine, 1.5 % (w/v) agar
CZD/S soft agar + argenine (w/o L-Methionine)	3.5 % (w/v) Czapek Dox broth, 18.22 % (w/v) D-sorbitol, 0.1 % (w/v) (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> , 0.15 % (w/v) argenine, 0.05 % adenine, 0.8 % (w/v) agar
DPY	2 % (w/v) Dextrin from potato starch, 1 % (w/v) Polypeptone, 0.5 % (w/v) Yeast extract, 0.5 % (w/v) KH <sub>2</sub> PO <sub>4</sub> , 0.05 % (w/v) MgSO <sub>4</sub> x H <sub>2</sub> O
DPY Agar	DPY Medium, 2.5 % (w/v) Agar
GN	2,00 % (w/v) D(+)-Glucose Monohydrate 1,00 % (w/v) Nutrient broth Nr. 2 from Oxoid

**Table 4.1.3** Buffers used in this work.

<b>Buffers</b>	<b>Composition</b>
Lysis buffer	50 mM Tris-HCl, pH 8, 150 mM NaCl, 10 mM imidazole, 10 % (v/v) glycerol
Elution buffer	50 mM Tris-HCl, pH 8, 150 mM NaCl, 500 mM imidazole, 10 % (v/v) glycerol
Storage buffer	50 mM Tris-HCl, 150 mM NaCl, 10 % (v/v) glycerol, pH 8
Coomassie dye	25 % (v/v) acetic acid, 10 % (v/v) isopropanol, 0,1 % (w/v) Coomassie
Coomassie bleach	25 % (v/v) acetic acid, 10 % (v/v) isopropanol
4x Lämmli	10 % (v/v) β-mercaptoethanol, 0.25 % (w/v) bromophenol blue, 30 % glycerol, 0.25 % xylene cyanol
10x SDS running	25 mM Tris-HCl, 192 mM glycine, 0.1 % (w/v) SDS, pH 8.3
50x TAE	2 M Tris acetate, 0.05 M EDTA, pH 8.3

**Table 4.1.4** Media and agar used in this work for other fungi.

<b>Media</b>	<b>Composition</b>
<i>CF236968</i> and <i>CF253087</i>	
PD agar	2.4 % (w/v) PDB powder, 1.5 % (w/v) agar
<i>CF150626</i>	
PD agar	2.4 % (w/v) PDB powder, 1.5 % (w/v) agar
TOM	1.0 % (w/v) oat meal, 4 % (w/v) tomato paste, 0.5 % (w/v) corn steep liquor, 1 % (w/v) D(+)-glucose monohydrate, 1 % (w/v) trace element solution
PM	5 % (w/v) yellow cornmeal, 0.1 % (w/v) yeast extract, 8 % (w/v) sucrose
½ TOM	0.5 % (w/v) oat meal, 2 % (w/v) tomato paste, 0.25 % (w/v) corn steep liquor, 0.5 % (w/v) D(+)-glucose monohydrate, 0.5 % (w/v) trace element solution
<i>T. stipitatus ΔtsL2</i>	
Sucrose	2,00 % (w/v) Sucrose, 2,00 % (w/v) Polypeptone (Roth), 5,00 % (v/v) Solution A, 5,00 % (v/v) Solution B
<i>P. cirsii</i>	
PD agar	2.4 % (w/v) PDB powder, 1.5 % (w/v) agar
PD	2,40 % (w/v) Potato Dextrose broth
TOM	1.0 % (w/v) oat meal, 4 % (w/v) tomato paste, 0.5 % (w/v) corn steep liquor, 1 % (w/v) D(+)-glucose monohydrate, 1 % (w/v) trace element solution
<i>T. reesei Δtmus53</i>	
ME agar	ME Medium, 1.5 % (w/v) Agar
ME	1,28 % (w/v) Malt extract (Roth), 0,08 % (w/v) Peptone ex Soya (Roth), 0,24 % (w/v) Glycerol (Roth), 0,28 % (w/v) Dextrin from potato starch (Sigma Aldrich), pH 4.7
GNB	2 % (w/v) D(+)-Glucose Monohydrate, 3 % (w/v) Nutrient broth Nr. 2 from Oxoid

**Table 4.1.5** Solutions used for fungal transformations.

<b>solutions</b>	<b>composition</b>
<i>A. oryzae</i>	
Solution 1	0.8 M NaCl, 10 mM CaCl <sub>2</sub> , 50 mM Tris-HCl, pH 7.5
Solution 2	60 % PEG 3350, 0.8 M NaCl, 10 mM CaCl <sub>2</sub> , 50 mM Tris-HCl, pH 7.5
<i>CF1506262</i>	
Solution 3	1.1 M sodium chloride solution, 10 mM CaCl <sub>2</sub> , 50 mM Tris-HCl pH 7.5
Solution 4	PEG 4000 (60% w/v) 1.1 M NaCl, 10 mM CaCl <sub>2</sub> , 50 mM Tris-HCl pH 7.5
<i>P. cirsii</i>	
Solution 5	1.3 M NaCl, 10 mM CaCl <sub>2</sub> , 50 mM Tris-HCl, pH 7.5
Solution 6	60 % PEG 3350, 1.3 M NaCl, 10 mM CaCl <sub>2</sub> , 50 mM Tris-HCl, pH 7.5
<i>T. reesei Δtmus53</i>	
T Protoplast solution	1.2 M sorbitol ,100 mM potassium phosphate pH 5.6
T Resuspend solution	1M sorbitol-10 mM CaCl <sub>2</sub> -10 mM Tris-HCl, pH 7.5
T PEG solution	25% PEG 6000- 50 mM CaCl <sub>2</sub> -10 mM Tris-HCl, pH 7.5

## 4.2 Enzymes and Antibiotics

Antibiotic stock solution were prepared in millipore water or ethanol and filter sterilized through 0.45 µm syringe filters and stored at - 20 °C. Stock and working concentrations are listed in Table 4.2. All enzymes used in this study were obtained from NEB or Thermo Scientific and used following the manufacturers' instructions

**Table 4.2** Used antibiotics in this work

<b>Antibiotic</b>	<b>stock concentration [mg/mL]</b>	<b>working concentration [µg/mL]</b>
carbenicillin (Carb <sup>50</sup> )	50 in H <sub>2</sub> O	50
chloramphenicol (CA <sup>30</sup> )	30 in EtOH	30
kanamycin (Kan <sup>50</sup> )	50 in H <sub>2</sub> O	50
Hygromycin B	50 in H <sub>2</sub> O	100

## 4.3 Microbiology methods

All bacterial strains, yeast and *A. oryzae* NSAR1 were obtained from the BMWZ inhouse strain collection while other fungal strains were kindly provided for this work.

**Table 4.3** Bacterial, yeast and fungal strains used in this work.

Strain	Origin
<i>E. coli</i> BL21 (DE3)	Thermo Fisher Scientific
<i>E. coli</i> OneShot Top10	Thermo Fisher Scientific
<i>E. coli</i> OneShot <i>ccdB</i>	Thermo Fisher Scientific
<i>A. oryzae</i> NSAR1	Lazarus group, Bristol
<i>S. cerevisiae</i> CEN.PK2	Hahn group, Hannover
CF236968, CF253087 and CF150626	Fundación MEDINA (Granada, Spain)
<i>T. reesei</i> <i>Atmus53</i>	Mach-Aigner group, Austria
<i>P. cirsii</i>	Maurizio Vurro group, Italy

#### 4.3.1 *E. coli* growth and transformation

*E. coli* strains were cultivated on LB agar. Cultures were incubated in LB medium at 37 °C for 12 to 18 h shaking at 200 rpm. For long term storage glycerol stocks (25 % glycerol) were stored in cryovial at -80 °C.

Frozen chemically competent *E. coli* cells (kindly provided by technical staff, 50 µL aliquot) were thawed on ice. 1 µl of purified plasmid DNA or 10 µl of ligation mixture were added and mixed gently with the cells. The mixture was cooled on ice for 30 min followed by a heat shock at 42 °C for 10 sec (BL21, Rossetta2 and BL21 codonPlus RP) or for 40 sec (*ccdB* and Top 10). Afterwards, the cell suspension was placed on ice for 2 min and 250 µL SOC-Medium was added, followed by an incubation of 45 min at 37 °C under shaking (250 rpm) conditions. 50 - 200 µL of bacteria suspension were spread on pre-warmed LB-Agar plates containing the appropriate antibiotic(s). The plates were incubated at 37 °C overnight.

#### 4.3.2 *Saccharomyces cerevisiae* growth and transformation

A stock of *S. cerevisiae* CEN.PK2 was spread on YPAD plates and incubated at 30 °C for three days or until the appearance of colonies. A single colony was used to inoculate 10 ml liquid YPAD, grown overnight at 30 °C and 200 rpm. *S. cerevisiae* CEN.PK2 transformed with *ura3* containing plasmids was grown on SM-URA plates at 30 °C for 2-4 days.



Preparation and transformation of competent *S. cerevisiae* cells was done using the LiOAc/SS carrier DNA/PEG protocol developed by Gietz and Woods.<sup>192</sup> To generate competent cells, a single colony was picked from a 3-5 days old YPAD plate, inoculated into 10 ml of YPAD medium and grown overnight at 30 °C with shaking at 200 rpm. The pre-culture was added to 40 ml fresh YPAD medium and incubated at 30 °C, 200 rpm, for 4–5 h. Yeast cells were then collected by centrifugation at 3000 rpm for 5 min and washed with 25 ml water before the cell pellet was resuspended in 1 ml water and transferred to a 1.5 ml tube. The cells were centrifuged at 20000 x g for 30 s and the supernatant removed. 50 µl aliquots were transferred into 1.5 mL tubes. A solution of ssDNA (Single- Strand Carrier DNA or salmon sperm DNA, 2 mg/mL) was boiled at 95 °C for 5 min and stored on ice until usage. For each transformation the following components were added to the 50 µl yeast suspension aliquot: 250 µl PEG solution (50% w/v polyethylene-glycol [PEG] 3350), 36 µl 1 M LiOAc, 50 µl ssDNA and up to 34 µl of DNA in water (with equimolar concentrations of DNA fragments). The transformation suspension was mixed by careful pipetting and incubated at 30 °C for 50 min. Afterwards the cells were pelleted at 3000 x g for 15 s and suspended in 1 mL distilled water. 200 µl of the mixture was spread over pre-warmed SM–URA plates and incubated at 30 °C for 3 to 4 days until the appearance of colonies.

#### **4.3.3 *Aspergillus oryzae* NSAR1 growth and transformation**

*A. oryzae* NSAR1 was grown on DPY agar plates or in DPY medium at 28 °C for 3-7 days (110 rpm). For long term storage *A. oryzae* spores were collected from plates by adding 3 ml distilled water, mixed with glycerol (final concentration 25 %) and stored in cryovials at –80 °C.

1 mL fresh spore suspension was used to inoculate 100 mL GN media (500 mL flask). The culture was incubated overnight at 28°C with shaking at 180 rpm. The resulting young mycelia were collected by filtering over a sterile miracloth filter, washed with 0.8 M NaCl (50-100 mL) and resuspended in 10 mL of filter-sterilised protoplast solution (10 mg/mL Trichoderma lysing enzyme in 0.8 M NaCl). After gentle shaking at 30°C for 2-3 h, protoplasts were released from hyphae strands by gently pipetting with a wide-bore pipette and filtered through a sterile miracloth filter. Protoplast were collected by centrifugation (3000 x g, 5 min) and resuspended in resuspension in fungal transformation solution 1 (100 µL per transformation). Up to 10 µl of purified plasmid DNA

was added to the 100  $\mu$ L protoplast suspension and incubated on ice for 2 min. After adding 1 mL fungal transformation solution 2, the transformation mixture was incubated 20 min at room temperature. 5 mL of molten (50°C) CZD/S1 soft agar was added to the transformation mixture and overlaid onto CZD/S1 plates. The plates were incubated for 3-5 days at 28°C until colonies became visible.

Transformants were then further selected twice on CZD/S1 plates (2<sup>nd</sup> and 3<sup>rd</sup> selection plate) to reduce the number of false-positive transformants. From the 3<sup>rd</sup> selection plate, a single colony was picked with a toothpick and grown on DPY agar for 1 week before they were used for fermentation and/or long-term storage. For fermentation *A. oryzae* transformants were grown on DPY plates for 5-7 days before spores were collected and used to inoculate 100 ml DPY medium (500 ml flasks). The inoculated flasks were incubated for 2-4 days at 28 °C shaking at 110 rpm followed by extraction and chemical analysis

#### **4.3.4 CF150626 growth and transformation**

*CF150626* was cultured on PD agar and at ambient temperature for 14 d. For long term storage mycelia was resuspended in a minimal amount of water, mixed with glycerol to a final concentration of 25 % (v/v) and stored at - 80 °C indefinitely.

1 mL fresh spore suspension was used to inoculate 100 mL PD media (500 mL flask). The culture was incubated 2 days at 22°C without shaking and light. The resulting young mycelia were collected by filtering over a sterile miracloth filter, washed with 1.1 M NaCl (50-100 mL) and resuspended in 10 mL of filter-sterilised protoplast solution (10 mg/mL *Trichoderma* lysing enzyme and 5 mg/mL driselase from Basidiomycetes in 1.1 M NaCl). After gentle shaking at 30°C for 3h, protoplasts were released from hyphae strands by gently pipetting with a wide-bore pipette and filtered through a sterile miracloth filter. Protoplasts were collected by centrifugation (3000 x g, 5 min) and resuspended in resuspension in 200ul fungal transformation solution 3 and 1/5 of volume of solution 4. Up to 10  $\mu$ l of purified plasmid DNA was added to 200  $\mu$ L of the protoplast suspension and incubated on ice for 30 min. After adding 1 mL fungal transformation solution 2, the transformation mixture was incubated 20 min at room temperature. 10 mL of molten PD soft agar (55 °C) was added to the transformation mixture and overlaid into empty plate. After incubation at 22 °C for 1

day the plates were overlaid with 10 mL PD agar containing 100 µg/mL hygromycin B. The plates were incubated for 8-10 days at 22°C until colonies became visible.

Transformants were then further selected twice on PD agar containing 100 µg/mL hygromycin B to reduce the number of false-positive transformants. From the 3<sup>rd</sup> selection plate a single colony was picked with a toothpick and grown alone on PD agar for 1 week before they were used for fermentation and/or long-term storage.

#### **4.3.5 *P. cirsii* growth and transformation**

*P. cirsii* was grown on PD agar plates or in PD medium at 28 °C for 7 days (110 rpm). For long term storage *P. cirsii* spores were collected from plates by adding 3 ml distilled water, mixed with glycerol (final concentration 50 %) and stored in cryovials at -80 °C.

1 mL fresh spore suspension was used to inoculate 100 mL PD media (500 mL flask). The culture was incubated 4-5 days at 28°C without shaking and light. The resulting young mycelia were collected by filtering over a sterile miracloth filter, washed with 1.3 M NaCl (50-100 mL) and resuspended in 10 mL of filter-sterilised protoplast solution (10 mg/mL Trichoderma lysing enzyme and 5 mg/mL driselase from Basidiomycetes in 1.3 M NaCl). After gentle shaking at 30°C for 2-3h, protoplasts were released from hyphae strands by gently pipetting with a wide-bore pipette and filtered through a sterile miracloth filter. Protoplasts were collected by centrifugation (4000 x g, 5 min) and resuspended in resuspension in 100ul fungal transformation solution 5. Up to 10 µg of purified DNA (bipartite KO fragments) was added to the protoplast suspension and incubated on ice for 2 min. After adding 1 mL fungal transformation solution 6 the transformation mixture was incubated 20 min at room temperature. 10 mL of molten PD soft agar (55 °C) was added to the transformation mixture and overlaid into empty plate. After incubation at 28 °C for overnight the plate was overlaid with 10 mL PD agar containing 100 µg/mL hygromycin B. The plate was incubated for 5-7 days at 28°C until colonies became visible.

Transformants were then further selected twice (2<sup>nd</sup> and 3<sup>rd</sup> selection plate) on PD agar containing 100 µg/mL hygromycin B to reduce the number of false-positive transformants. From the 3<sup>rd</sup> selection plate a single colony was picked with a toothpick and grown alone on PD agar for 1 week before they were used for fermentation and/or long-term storage. For fermentation *P. cirsii*

transformants were grown on PD plates for 8-10 days before spores were collected and used to inoculate 100 ml PD medium (500 ml flasks). The inoculated flasks were incubated for 5-7 days at 28 °C shaking at 110 rpm followed by extraction and chemical analysis.

#### **4.2.6 *T. reesei* $\Delta$ tmus53 growth and transformation**

*T. reesei*  $\Delta$ tmus53 was grown on ME agar plates or in ME medium at 28 °C for 7 days (110 rpm). For long term storage *T. reesei*  $\Delta$ tmus53 spores were collected from plates by adding 3 ml distilled water, mixed with glycerol (final concentration 25 %) and stored in cryovials at -80 °C.

1 mL fresh spore suspension was used to inoculate 100 mL GNB media (500 mL flask). The culture was incubated 1.5 days at 28°C with 150rpm shaking and light. The resulting young mycelia were collected and dried by filtering over a sterile miracloth filter and resuspended in 10 mL of filter-sterilised T protoplast solution. After gentle shaking at 30°C for 3 h, protoplasts were released from hyphae strands by gently pipetting with a wide-bore pipette and filtered through a sterile miracloth filter. Protoplasts were collected by centrifugation (4000 x g, 5 min) and resuspended in 200ul resuspend solutions. Up to 10 µg of purified DNA (bipartite KO fragments/plasmid DNA) and 50 ul PEG solution were added to the protoplast suspension and incubated on ice for 20 min. After adding 1ml of T PEG-buffer and 2ml T resuspension solution the transformation mixtures were incubated 5 min at room temperature. Aliquots of 200, 300, 500, 600 ul transformation mixture were plated on PD agars. After incubation at 28 °C for 3.5 h the plates were overlaid with 10 mL PD soft agar containing 100 µg/mL hygromycin B. The plate was incubated for 3-4 days at 28°C until colonies became visible.

Transformants were then further selected twice (2<sup>nd</sup> and 3<sup>rd</sup> selection plate) on PD agar containing 100 µg/mL hygromycin B to reduce the number of false-positive transformants. From the 3<sup>rd</sup> selection plate a single colony was picked with a toothpick and grown alone on PD agar for 4 days before they were used for fermentation and/or long-term storage. For fermentation *T. reesei*  $\Delta$ tmus53 transformants were grown on PD plates for 4 days before spores were collected and used to inoculate 100 ml PD medium (500 ml flasks). The inoculated flasks were incubated for 3-4 days at 28 °C shaking at 110 rpm followed by extraction and chemical analysis.

## 4.4 Molecular biology methods

### 4.4.1 Oligonucleotides

All oligonucleotides used in this work were purchased from Sigma Aldrich.

**Table 4.4.1.1** Oligonucleotides used in chapter 2

Lab ID	Name	Sequence
TS 1	Hyg-2-F	CTGTCGAGAAGTTTCTGATCG
TS 2	Hyg-2-R	CTAGAAAGAAGGATTACCTC
TS 3	eGFP-1R	GTCCATGCCGAGAGTGATCC
TS 4	eGFP-1F	TAAACGGCCACAAGTTCAGC
TS 5	P1asL4	TTTCTTTCAACACAAGATCCCAAAGTCAAATGCCGCAAC TAAAGGTTCT
TS 6	P2asL4	GGTTGGCTGGTAGACGTCATATAATCATACTCACTCCTTGA GAAGCTCTG
TS 7	P1asL6	TTTCTTTCAACACAAGATCCCAAAGTCAAATGACTGTGA AGATCCTTGT
TS 8	P2asL6	GGTTGGCTGGTAGACGTCATATAATCATACTCATGCCTCAA ACTCCAGCT
TS 9	P1 L1 PKS F	TACTTTGAAGACGGCCGGTC
TS 10	P2 L1 PKS R	TTGGGTCAATCGTGGCATCA
TS 11	R5 INTRON F	TCAGCAAGCCCTTGTGGAAA
TS 12	R5 INTRO R	ACGTGGCTGATCAATGCTCA
TS 13	P1 150tub F	CGCTCATCCATCCTCTCGAC
TS 14	P2 150tub R	GAAGTTGTCAGGGCGGAAGA
TS 15	1385	TTTCTTTCAACACAAGATCCCAAAGTCAAATGTCATGGG GCTCGACAGC
TS 16	1393	GGGCTCGTGTCTTTGCCTG
TS 17	P1 norL1 F	CACTCGTTCGGATCGGAGAG
TS 18	P2 norL1 R	GGGTACAACCAGCATCGTGA
TS 19	P1 norL2 F	ACAAAGCCAACCATCGTTGC
TS 20	P2 norL2 R	CTTACACCTCCCAACCCGAC
TS 21	P1 norL3 F	TTTGAACCCAACTGGCTCGT
TS 22	P2 norL3 R	CGAACCTCGGTCCTTTCTCC
TS 23	P1 norL4 F	CCATCATTGCGCAAGACTCG
TS 24	P2 norL4 R	TCGCCAAAGCTCTTTTTGCC
TS 25	P1 norL5 F	AAGACCGAGAGCCAGGATCT
TS 26	P2 norL5 R	GTCGAAAAGATGAGCGACGC
TS 27	P1 norL6 F	TCAACGTCAAGGACTGGTGC
TS 28	P2 norL6 R	GTTGGGTTTTGGTGACGAGC
TS 29	P1 norL7 F	GGATTGACAAACGGCCTTGG
TS 30	P2 norL7 R	TATGGGCCGAAAGCATGTGT
TS 31	P1 norL8 F	AGCTCGTTGGTGGATGATGG
TS 32	P2 norL8 R	TTTCCCTCATCCCAAGTCGC
TS 33	P1 norR1 F	CCTTGTCATGCTCGGTCCTT
TS 34	P2 norR1 R	ATGGTAACCTGCTGCTCACC
TS 35	P1 norR2 F	CGAACAAGACCAACCCTCGA
TS 36	P2 norR2 R	GCGGCGTTGTGACATTATCC
TS 37	P1 norR3 F	ATCCATCCGCGGACTTTAT
TS 38	P2 norR3 R	CGGAAGACGCATACCTCCTC
TS 39	P1 norR4 F	GGCTTCTGGATGTGCAGACT
TS 40	P2 norR4 R	CAGGACGTCGATGAGAGAGC
TS 41	P1 norR5 F	ACGAGAAAACCTTCGAGGCGT
TS 42	P2 norR5 R	CCATAGTCGGACCATCCACG

TS 43	P1 norR6 F	AGTGAACGCGCCATAACAAGA
TS 44	P2 norR6 R	ATCCGCAATTGCATAACCGC
TS 45	P1 norR7 F	CTCAGCGATCGTAACTCGCT
TS 46	P2 norR7 R	GCGAATCGACATTCTGCACC
TS 47	norL4 Padh F	TTTCTTTCAACACAAGATCCCAAAGTCAAAATGACGTCTT TTGGAACCAA
TS 48	norL4 Padh R	TTCATTCTATGCGTTATGAACATGTTCCCTCTACCAGGGTTG CACCATAG

**Table 4.4.1.2** Oligonucleotides used in chapter 3

Lab ID	Name	Sequence
S 1	P1 L2MT F	CAGAGGGGGTGTACATGCTG
S 2	P2 L2MT R	CAAGGAGACGCTGGTGGATT
S 3	SCYR 1 L F	AACCACTTTAGTGCCACG
S 4	HYB F	CTATCGGCGAGTACTTCTAC
S 5	SCYR 1 RR	GTCTTTAGTCTTCACGTCAT
S 6	Gpda R	ACAACACTACCAACATGGAGTA
S 7	SCY R3 F	ACAGCCAACGTCACTACTGG
S 8	SCY R3 R	GCTTGGGAGAGGATGTGCTT
S 9	scy R1 TL Fw	GCCACTGAGTTGTGACGTTT
S 10	HygR internal Fw	GCTTTCAGCTTCGATGTAGG
S 11	HygR internal Re	CGTCAGGACATTGTTGGAG
S 12	scyR1 TR Re	CTTCATTTCTAGCCGTGGAC
S 13	TA1 KO 5F	TTTGTACAAAAAAGCAGGCTCCGCGGTACCCAACAG CAGTAACCATTCTT
S 14	TA1 KO 5R	ACGTATTTTCAGTGTTCGAAAGATCCACTAGAAGAGGG TTGAGCCGATGTTCG
S 15	TA1 KO 3F	AGCGCCCACTCCACATCTCCACTCGACCTGGGCAGTC GTCCTTCGACGTC
S 16	TA1 KO 3R	GTACAAGAAAGCTGGGTTCGGCGCGAAGCTTTCAGG AATCCGTCGCTG
S 17	P1 b tubulin F	CGAAGCTTATCAAGTGGCGC
S 18	P2 b tubulin R	CGGTCAAGAATGTGGGTGGA
S 19	P1 PA2 F	TTCTTCCCACCTCAAGCACC
S 20	P2 PA2 R	GCAATGACGATCCATATTGCA
S 21	P1 PA3 F	TCTGCAAGTCAGCCAATCGT
S 22	P2 PA3 R	CGATGTGCAGCTCGTATTGC
S 23	P1 PA4 F	CAGGGTCGACATGCCTATCC
S 24	P2 PA4 R	GCGTTGAAAGTTGGGTCACC
S 25	P1 FR1 F	GGCTCGCCGAAAATGTTTT
S 26	P2 FR1 R	GCTATTGACTCGTCGCTCGA
S 27	p1 PB6 F	CTGGTGGGGTTCGTCATTGAA
S 28	P2 PB6 R	CCCAGATGCAATGGTTCCCT
S 29	P 56 R	CAGGTTCGAGTGGAGATGTG
S 30	P 671	TCTAGTGGATCTTTTCGACAC
S 31	TA1 KO H F	AGCTTTCCCCTTCATCGCA
S 32	TA1 KO R R	TTCCAGTCGCTGCATCAGTT
S 33	PLA1 F	CCATCAACTCTCCCACCACC
S 34	PLA1 R	TGCAGGTTACTACGCAACGT
S 35	PLA2 F	GAGCACGCGGAATCAATGAC
S 36	PLA2 R	CTCCATGGAAGGTCCCAACC
S 37	PLA3 F	TCCTTCCGGAAGAGGATCGT
S 38	PLA3 R	GCCGCATTCTGTCCGATTTC
S 39	PB1 F	TGCCTCGCAGCATATTCGAT
S 40	PB1 R	CATCAACGGTCCGAACCTCA
S 41	PB2 F	AACAGGCGACAATTCCGACT
S 42	PB2 R	TCTCTACACGAGCTCTGCCT

S 43	PB3 F	CCAGAACCAGTTGTTTGC GG
S 44	PB3 R	CAACAAGCGTGAACAAGCGA
S 45	PB4 F	GCGAGCGACTATTCATGGGA
S 46	PB4 R	ATCTCCAGGATGGCTCCCTT
S 47	PB5 F	AGTGCGACAGCTGGAGTAAC
S 48	PB5 R	TGAACTCGGACACGTCAAGG
S 49	1327	CGTCAGGACATTGTTGGAG
S 50	1328	GCTTTCAGCTTCGATGTAGG
S 51	P1 ITS	TCCGTAGGTGAACCTGCGG
S 52	P2 ITS	TCCTCCGCTTATTGATATGC
S 53	PA1 5F	TTTGTACAAAAAAGCAGGCTCCGCGGTACCACCATC ACTGTTCTAGGCC
S 54	PA1 5R	ACGTATTTTCAGTGTGCGAAAGATCCACTAGATCCAAGA GGCAGGGAAAGGT
S 55	PA1 3F	AGCGCCCACTCCACATCTCCACTCGACCTGAACTCCA GAAACAGAGCAAT
S 56	PA1 3R	GTACAAGAAAGCTGGGTTCGGCGCGAAGCTTAAGAAT GGTCGCCATGATGC
S 57	PA1 KO RR	AAGAATGGTTCGCCATGATGC
S 58	P1 TR b tubulin	TAACCAAGTTGGTTCTGCCTT
S 59	P2 TR b tubulin	CCCTGGAGGCATTCACAGTT
S 60	PA3 5F	TTTGTACAAAAAAGCAGGCTCCGCGGTACCGATTCTT CTTTAGTGTTTCA
S 61	PA3 5R	ACGTATTTTCAGTGTGCGAAAGATCCACTAGATCGCAGC ATCGTTCGAAGAA
S 62	PA3 3F	AGCGCCCACTCCACATCTCCACTCGACCTGGGTGGAT GTGGAGCCAATTC
S 63	PA3 3R	GTACAAGAAAGCTGGGTTCGGCGCGAAGCTTAGCCGC ATAATTCATGTCAG
S 64	b tubulin 1 F	AACAAC TGGGCAAAGGGTCA
S 65	b tubulin 1 R	GGTGAGCTCAGGAACGTTGA
S 66	P1 a tubulin F	ACGATATCTGCCGTCGGAAC
S 67	P2 a Tubulin R	TTCGACCAGGTTGAAGGACG
S 68	P1 Actin F	ATATGTTGCCTCGGAGCTCG
S 69	P2 Actin R	CTGCAGAAGTCTCCCGTTGT
S 70	PC MT1 F	GCAACAATTACCACGACGAA
S 71	PC MT1 R	TTATTCCACGGCCTTTCTTC
S 72	Y MT2 F	TTTCTTTCAACACAAGATCCCAAAGTCAAAGCAACAA TTACCACGACGAA
S 73	Y MT2 R	GGTTGGCTGGTAGACGTCATATAATCATACTTATTCC ACGGCCTTTCTTC
S 74	TA1 KO F1	GATCCTTCTTGCTTGGGGCT
S 75	TA1 KO R1	CAGGGCATCCAGAACTCGT
S 76	PA1 KO 1F	GGGTCAATTGTCCACGCAAC
S 77	PA1 KO 1R	ACTCTCTCATCGTCCCGTGA
S 78	PA3 KO F1	GACCCCAACGGATTGCCTAA
S 79	PA3 KO R1	AGTGGTGCAGTCCAAC TAGC
S 80	LTA3 P 5F	TTTGTACAAAAAAGCAGGCTCCGCGGTACCCTTCATA CACAAGTACGCCT
S 81	LTA3 P 5R	ACGTATTTTCAGTGTGCGAAAGATCCACTAGAGCAGGG ACATTGGCATATTC
S 82	LTA3 P 3F	AGCGCCCACTCCACATCTCCACTCGACCTGATTGGCT AGGTTCTTCTTGG
S 83	LTA3 P 3R	GTACAAGAAAGCTGGGTTCGGCGCGAAGCTTTTAGAA AAAGTAAAACAGCT
S 84	TA2 P 5F	TTTGTACAAAAAAGCAGGCTCCGCGGTACCATGCC GAGCTGGGACTCCT
S 85	TA2 P 5R	ACGTATTTTCAGTGTGCGAAAGATCCACTAGAAGCCGGT TGGGGTCCGGCTG

S 86	TA2 P 3F	AGCGCCCACTCCACATCTCCACTCGACCTGCAGCTCG ACGGCGCGCACGT
S 87	TA2 P 3R	GTACAAGAAAGCTGGGTCGGCGCGAAGCTTTCACGC TTCGTGCTGGTAGT
S 88	P419	CCACTTCATCGCAGCTTGAC
S 89	PA1 KO 2R	AGTTCTTCTCGGCGTTCTGG
S 90	PA1 KO 2F	CGAGACTGAGGAATCCGCTC
S 91	PA3 KO 2F	GAGCGGATTCTCAGTCTCG
S 92	PA3 KO 2R	CCCGGTGTATGAAACCGGAA
S 93	PA2 P 5F	TTTGTACAAAAAAGCAGGCTCCGCGGTACCCTGCGC GATGTGCCACAGAT
S 94	PA2 P 5R	ACGTATTTTCAGTGTTCGAAAGATCCACTAGATGTGGGC ACCACATCTCCGC
S 96	PA2 P 3F	AGCGCCCACTCCACATCTCCACTCGACCTGATTGCAC CTCCCTCTGGAGG
S 97	PA2 P 3R	GTACAAGAAAGCTGGGTCGGCGCGAAGCTTGGACGA TGAGAGAGTTTATT
S 98	PA4 P 5F	TTTGTACAAAAAAGCAGGCTCCGCGGTACCCATGGT ATTTTCGTTCCAAAC
S 99	PA4 P 5R	ACGTATTTTCAGTGTTCGAAAGATCCACTAGACGGTTGT GAGAGTATTGACG
S 100	PA4 P 3F	AGCGCCCACTCCACATCTCCACTCGACCTGTAGCCCT AGAGTGGGCGTTCG
S 101	PA4 P 3R	GTACAAGAAAGCTGGGTCGGCGCGAAGCTTGCTACT GCTACTCTAATGAC
S 102	P.C MeT 12 F	TGGGCAATCGACTTCGGTAC
S 103	P.C MeT 12 R	AACGGAAAACCTGCTTTGGC
S 104	P.C MeT 5 F	TCTTGGGTGCCGGCAAATAT
S 105	P.C MeT 5R	GCTCAAACACGCCGACTTTT
S 106	P. C MeT 10 F	TCTGGAGTCTGGCTTCTCGA
S 107	P. C MeT 10 R	CTCGATCCTCTGCTTCGTCC
S 108	MeT 9 F	AAGCGTTCCCGTCAATGTCT
S 109	MeT 9 R	CATCCGGTGTGAGATCCTCG
S 110	LTA1 LF P1	TTTGTACAAAAAAGCAGGCTCCGCGGTACCCTTCATA CACAAGTACGCCT
S 111	LTA1 LF P2	ACGTATTTTCAGTGTTCGAAAGATCCACTAGACAGTGCG GTGGAGGAGTCAA
S 112	LTA1 RF P1	AGCGCCCACTCCACATCTCCACTCGACCTGTTCTCAG CGCGTTTTTTTGT
S 113	LTA1 RF P2	GTACAAGAAAGCTGGGTCGGCGCGAAGCTTTTAGAA AAAGTAAAACAGCT
S 114	TA1 LF1 P1	TTTGTACAAAAAAGCAGGCTCCGCGGTACCGTTATGG TTGCCCCGAAGCG
S 115	TA1 LF1 P2	ACGTATTTTCAGTGTTCGAAAGATCCACTAGATGAGCCG ATGTCGACGCCGT
S 116	TA1 RF1 P1	AGCGCCCACTCCACATCTCCACTCGACCTGGCCGACG TCGATCTCCGCGA
S 117	TA1 RF1 P2	GTACAAGAAAGCTGGGTCGGCGCGAAGCTTACACGG TGATCTGCCTACCT
S 118	TA1 LF2 P1	TTTGTACAAAAAAGCAGGCTCCGCGGTACCGAATTGT AAAAGCGCCGGCT
S 119	TA1 LF2 P2	ACGTATTTTCAGTGTTCGAAAGATCCACTAGATGAGCCG ATGTCGACGCCGT
S 120	TA1 RF2 P1	AGCGCCCACTCCACATCTCCACTCGACCTGGCCGACG TCGATCTCCGCGA
S 121	TA1 RF2 P2	GTACAAGAAAGCTGGGTCGGCGCGAAGCTTTCGCTG GTAGTTCTGCGCAA
S 122	TA2 LF P1	TTTGTACAAAAAAGCAGGCTCCGCGGTACCCAGAA GCATGCAAGTCTCA



S 123	TA2 LF P2	ACGTATTTTCAGTGTTCGAAAGATCCACTAGATCGGCCA CCACCATCTGGCT
S 124	TA2 RF P1	AGCGCCCACTCCACATCTCCACTCGACCTGTGCAGGG CCTCGTGCACGT
S 125	TA2 RF P2	GTACAAGAAAGCTGGGTTCGGCGCGAAGCTTACAGGA CGCAATACCAGGAG
S 126	TA4 LF P1	TTTGTACAAAAAAGCAGGCTCCGCGGTACCCGAGAC CATCAAGGGGGCTC
S 127	TA4 LF P2	ACGTATTTTCAGTGTTCGAAAGATCCACTAGAGGATGTT CAGGCCGAGGTAC
S 128	TA4 RF P1	AGCGCCCACTCCACATCTCCACTCGACCTGACACCAC CCCTTGCGGCGGA
S 129	TA4 RF P2	GTACAAGAAAGCTGGGTTCGGCGCGAAGCTTCTCACC GATTGTTGCTACC
S 130	KO LTA1 LF P1	TTCCTATTAGAAGCCTCGACTCTT
S 131	KO LTA1 LF P2	AACCAGATTCGAAAGCGCCT
S 132	KO LTA1 RF P1	TTTACCCAGAATGCACAGGT
S 134	KO LTA1 RF P2	AAACACCGCAAACAGTGAAT
S 135	KO TA1 LF1 P1	TTCGACGAATGCCGAACTCT
S 136	KO TA1 LF1 P2	AACCAGATTCGAAAGCGCCT
S 137	KO TA1 RF1 P1	TTTACCCAGAATGCACAGGT
S 138	KO TA1 RF1 P2	AGACTTGCATGCTTCTGGAA
S 139	KO TA2 LF P1	CAGATGCATAAGACCGCAA
S 140	KO TA2 LF P2	AGGCGCTTTCGAATCTGGTT
S 141	KO TA2 RF P1	ACCTGTGCATTCTGGGTA
S 142	KO TA2 RF P2	TCGTTCGTGCAGGATCTTTGA
S 143	KO TA1 LF2 P1	ATCGTTCGGTATCAGGGATGT

#### 4.4.2 Obtained and constructed vectors in this thesis

Construction of fungal and bacterial expression vectors was performed using the described in vivo homologous recombination in *S. cerevisiae* (section 4.3.2).

**Table 4.4.2.1** Overview of obtained and purchased plasmids.

Lab ID	Name	Vector backbone	Origin
LL I 1	pTH-GS-eGFP	-	Lazarus Group, Bristol
LL I 2	pTYGSmet- <i>asL4-asL6</i>	pTYGS- <i>sC</i>	Dr. RAISSA SCHOR
LL II 3	pTYGSargB- <i>PA4</i>	pTYGS-arg	Self-constructed
LL II 4	pE-YA- <i>PA4</i>	pE-YA	Self-constructed
LL II 5	pE-YA- <i>PA6</i>	pE-YA	Self-constructed
LL II 6	pE-YA- <i>TA1</i>	pE-YA	Self-constructed
LL II 7	pE-YA- <i>TA4</i>	pE-YA	Self-constructed
LL II 8	pE-YA- <i>TA5</i>	pE-YA	Self-constructed
LL I 9	pET100- <i>eup2R1</i>	pET100/D-TOPO	Thermo Fischer
LL I 10	pET100- <i>eup2L4</i>	pET100/D-TOPO	Thermo Fischer
LL I 11	pET100- <i>pycR1</i>	pET100/D-TOPO	Thermo Fischer
LL I 12	pET100- <i>eup2R3</i>	pET100/D-TOPO	Thermo Fischer

**Table 4.4.2.2** Oligonucleotides used for plasmid construction in *Saccharomyces cerevisiae*

Lab ID	Name	Oligonucleotides
LL II 3	pTYGSargB-PA4	PA4: S72 + S73
LL II 4	pE-YA-PA4	LF: S53 + S54; RF: S55 + S56; <i>hph</i> : S29 + S30
LL II 5	pE-YA-PA6	LF: S60 + S61; RF: S62 + S63; <i>hph</i> : S29 + S30
LL II 6	pE-YA-TA1	LF: S110 + S111; RF: S112 + S113; <i>hph</i> : S29 + S30
LL II 7	pE-YA-TA4	LF: S114 + S115; RF: S116 + S117; <i>hph</i> : S29 + S30
LL II 8	pE-YA-TA5	LF: S122 + S123; RF: S124 + S125; <i>hph</i> : S29 + S30

#### 4.4.3 Polymerase chain reaction (PCR)

PCR was used to amplify DNA fragments from plasmid or genomic DNA templates. For cloning the proofreading Q5<sup>®</sup> High-Fidelity 2X Master Mix (New England Biolabs) was used following the manufacturers' instructions. Sequences were checked by DNA sequencing (Section 4.4.10). OneTaq<sup>®</sup> 2X Master Mix (New England Biolabs) with standard buffer was used, following the manufacturers' instructions, for diagnose.

#### 4.4.4 RNA Extraction and cDNA Preparation

Total RNA of *CF150626* was isolated from fungal mycelia (typically 100 mg) using the TRIzol<sup>®</sup> Reagent (Thermo Fischer Scientific) according to the manufacturer's instructions. Total RNA was frozen in liquid nitrogen and stored at - 80 °C or directly converted to cDNA.

Total RNA of *P. cirsii* was isolated from fungal mycelia using the ZR Fungal Bacterial RNA MiniPrep kit (Zymo Research) according to the provided protocol. RNase free tubes and plugged pipet tips were used to avoid degradation of RNA.

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

#### 4.4.5 Agarose gel electrophoresis

DNA or RNA samples (*e. g.* plasmid DNA, PCR fragments) were mixed with 6 x loading buffer and loaded on a 1 % agarose gel (w/v) containing Roti<sup>®</sup> -Safe GelStain (Roth). 1kb DNA Ladder (New England Biolabs) was used to determine the size of the DNA fragments. The gel ran at 120 V for 25 - 30 min in a Bio-Rad gel chamber containing 0.5 x TAE running buffer. Digital images of the gels were created with the associated software. Visualization and documentation were achieved with the Molecular Imager<sup>®</sup> Gel Doc<sup>™</sup> XR + (Bio-Rad) under UV light (312 nm).

#### **4.4.6 DNA purification from gel or PCR**

Amplified DNA was purified for sequencing using NucleoSpin® Gel and PCR Clean-up kit (Machery-Nagel) according to the manufacturer's instructions.

#### **4.4.7 Isolation of plasmid DNA from *E. coli***

The NucleoSpin® Plasmid kit (Machery-Nagel) was used to isolate plasmid DNA from an overnight culture of *E. coli*, following the manufacturer's protocol.

#### **4.4.8 Isolation of plasmid DNA from *S. cerevisiae***

The Zymoprep Yeast Plasmid Miniprep II kit purchased from Zymo Research was used to isolate plasmid DNA from transformed *S. cerevisiae* cells grown on SM-URA plates. All the cells from the plates were collected and added directly to the lysis buffer provided by the kit. Yeast DNA was isolated according to manufacturers' instructions. For final elution 10 µL ddH<sub>2</sub>O was used.

#### **4.4.9 Isolation of genomic DNA from Fungi**

For the genomic DNA extraction of fungi, the GeneElute™ Plant Genomic DNA Miniprep Kit (SIGMA Life Science) was used following manufacturer's instructions. Around 100 mg of vacuum filtered mycelium from a liquid culture was grinded in liquid nitrogen to obtain a very fine powder then used for genomic DNA extraction.

#### **4.4.10 DNA sequencing**

DNA samples were sequenced by Eurofins Genomics (Ebersberg) and analyzed using the software Geneious.

#### **4.4.11 Determination of DNA Concentration**

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

## 4.5 Heterologous Protein Production and Purification

### 4.5.1 Production of Protein in *E. coli*

For the heterologous expression of fungal genes in *E. coli* the *E. coli* BL21 (DE3) strain was used as an expression host. *E. coli* BL21 (DE3) is a common laboratory expression strain and deletion of the *lon* and *ompT* proteases makes it suitable for the production of high-level recombinant protein (recombinant protein can account for up to 50 % of total biomass).<sup>193,194</sup> Genes of interest were typically test-expressed on an analytical scale to find suitable expression conditions. For this 250 mL baffled flasks containing 50 mL medium (either 2TY or LB) were inoculated with 1 mL of an overnight bacterial seed culture. The culture was then cultivated at 37 °C and 225 rpm until the optical density OD<sub>600</sub> reached values between 0.4 and 0.6. Protein production was induced by addition of IPTG to a final concentration of 0.1 - 1.0 mM and the cultures were kept at 10 - 25 °C for 16 - 20 h. After cultivation the cells were collected by centrifugation (4000 x g, 20 min, 4 °C). The obtained pellet was resuspended in 15 mL loading buffer, shock-frozen in liquid nitrogen and stored at - 20 °C until further usage.

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

### 4.5.2 Cell Lysis and Protein Extraction

The harvested cell pellets were resuspended in 10 mL lysis buffer. Lysis of the cells was done by sonication using a SonoPlus Ultrasonic homogenizer from Bandelin. The cells were intermittently sonicated on ice for 30 s with 30 s allowed for cooling. The total sonication time was 5-10 min. The cell lysate was centrifuged at 12000 rpm for 45 min at 4 °C. The supernatant and the pellet of samples were used for SDS PAGE. The protein of interest was purified using immobilized metal ion affinity chromatography. During the whole process the protein samples were kept on ice.

### 4.5.3 Histidine-tagged Protein Purification by Ni-NTA

Purification of histidine-tagged recombinant proteins were done by immobilized metal ion affinity chromatography using Ni-NTA agarose (Qiagen). All buffers, washing solutions and samples were filtered through a 0.45 µm filter prior to usage. The total lysate was mixed with 1.5 mL PureCube

Ni-NTA agarose and incubated at 4 °C for 1 h under light rotation. The mixture was then applied to a CHROMABOND® empty column with PE filter element (15 mL; Macherey-Nagel). Manual pressure was used to increase the flow rate. The crude lysate was passed through a Ni-NTA column (2 mL bed volume). Firstly, the column was washed with 10 mL loading buffer. Elution buffers were added with a stepwise gradient of imidazole (20 - 500 mM). 3 mL fractions were collected over the course of elution and analysed using SDS-PAGE. The eluate was concentrated and the elution buffer was replaced by storage buffer using amicon ultra centrifugal filter units, cut off 10000 MW, from Merck. The purified enzyme was stored at -80 °C or immediately used for enzyme assays.

#### 4.5.4 Protein Separation Using SDS - Polyacrylamide Gel Electrophoresis

Separation and analysis of proteins was performed using 12 % polyacrylamide gels (Table 4.5.4). A gel electrophoresis and gel casting system from Bio-Rad was used. Protein samples were prepared by mixing 5 µL 4x Lämmli buffer with 15 µL of protein solution before boiling the mixture at 95 °C for 10 minutes. 10 to 20 µL of the samples were loaded onto the gel. The Colour Prestained Protein Standard (Broad Range, 11 - 245 kDa; New England Biolabs) was used as a reference. Protein electrophoresis occurred at 40 mA for 45 - 50 min. After that, gels were incubated in 30 mL Coomassie staining solution for 2-3 h. Destaining was achieved using coomassie bleach solution and fresh 20 mL aliquots were changed every 20 - 30 min. Analysis and documentation was performed with the Molecular Imager Gel Doc XR+ system (Bio-Rad).

**Table 4.5.4** Composition of a 12 % separating gel and a 5% stacking gel.

Separating gel volume / mL	Stacking gel volume / mL	Composition
3.00	0.54	30 % acrylamide/bisacrylamide (Rotiphorese® Gel 30 [37,5:1])
2.45	1.70	ddH <sub>2</sub> O
1.90		1.5 M Tris-HCl, pH 8.8
	0.25	0.5 M Tris-HCl, pH 6.8
0.075	0.02	10 % (w/v) SDS
0.075	0.02	10 % (w/v) APS
0.003	0.002	TEMED

#### **4.5.5 Determination of Protein Concentration**

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

### **4.6 Chemical Analysis**

#### **4.6.1 Extraction of *A. oryzae* cultures**

The mycelia and solution of grown *A. oryzae* cultures were separated by vacuum filtration. The solution part was extracted twice with equal volume of ethyl acetate respectively. Organic layers were dried with anhydrous MgSO<sub>4</sub> and filtered with filter paper. Next, the solvent was removed with rotary evaporator to give dry crude extracts. They were stored at - 20 °C or dissolved in methanol to a concentration of 10 mg/ml. The sample need to be filtered over glass wool before LCMS analysis.

#### **4.6.2 Extraction of *CF236968* and *CF253087***

Grown *CF236968* or *CF253087* cultures (on PD agar) were cut into small pieces and transfer to a clean beaker. 200 ml ethyl acetate (EtOAc) was added to extract the metabolites with 4 hour's stirring. The extract was separated from mycelia by filtration. Organic layers were dried with anhydrous MgSO<sub>4</sub> and filtered with filter paper. Next, the solvent was removed with rotary evaporator to give dry crude extracts. They were stored at - 20 °C or dissolved in dichloromethane:methanol (1:1) to a concentration of 10 mg/ml. The sample need to be filtered over glass wool before LCMS analysis.

#### **4.6.3 Extraction of *CF150626***

Grown *CF150626* cultures were mixed with equal volumes of acetone and stirred at ambient temperature for 1 h. After filtration, the crude extract solution was concentrated to aqueous under reduced pressure with rotary evaporator. The remaining aqueous phase was extracted two times with equal volumes EtOAc respectively. Combined organic layers were dried with anhydrous MgSO<sub>4</sub> and the solvent was removed with rotary evaporator. Extracts were dissolved in dichloromethane:methanol (1:1) to a concentration of 10 mg/ml. The sample need to be filtered over glass wool before LCMS analysis.

#### 4.6.4 Extraction of *P. cirsii* and *T. reesei Δtmus53*

The mycelia and solution of grown *P. cirsii* or *T. reesei Δtmus53* cultures were separated by vacuum filtration. The solution part was extracted twice with equal volume of ethyl acetate respectively. Organic layers were dried with anhydrous MgSO<sub>4</sub> and filtered with filter paper. Next, the solvent was removed with rotary evaporator to give dry crude extracts. They were stored at - 20 °C or dissolved in acetonitrile to a concentration of 10 mg/ml. The sample need to be filtered over glass wool before LCMS analysis.

#### 4.6.5 Extraction of *T. stipitatus ΔtsL2*

The mycelia and solution of grown *T. stipitatus ΔtsL2* cultures were separated by vacuum filtration. The aqueous phase was acidified with 2 M HCl to pH 3 and extracted twice with equal volume of ethyl acetate. Combined organic layers were dried with anhydrous MgSO<sub>4</sub> and filtered with filter paper. Next, the solvent was removed with rotary evaporator to give dry crude extracts. They were stored at - 20 °C or dissolved in methanol to a concentration of 10 mg/ml. The sample need to be filtered over glass wool before LCMS analysis.

#### 4.6.6 Liquid Chromatography Mass Spectrometry (LCMS) Analysis

Analytical LCMS data was obtained using a Waters LCMS system comprising of a Waters 2767 autosampler, Waters 2545 pump, a Phenomenex Kinetex column (2.6 μm, C<sub>18</sub>, 100 Å, 4.6 x 100 mm) equipped with a Phenomenex Security Guard precolumn (Luna, C<sub>5</sub>, 300 Å) with a flow rate of 1 ml/min. Data detection was carried out by a diode array detector (Waters 2998) in the range 210 to 600 nm and an ELSD detector (Waters 2424) together with a mass spectrometer, Waters SQD-2 mass detector, operating simultaneously in ES<sup>+</sup> and ES<sup>-</sup> modes between 150 and 1000 *m/z*. A solvent gradient was run over 15 min starting at 10 % acetonitrile(0.045% formic acid)/90 % HPLC grade water (0.05 % formic acid) and ramping to 90 % acetonitrile/10 % HPLC grade water.

Preparative LCMS consisted of a Waters mass-directed auto-purification system including a Waters 2767 autosampler, Waters 2545 pump system, a Phenomenex Kinetex Axia column (5 μm, C<sub>18</sub>, 100 Å, 21.2 x 250 mm) equipped with a Phenomenex Security Guard precolumn (Luna, C<sub>5</sub>, 300 Å) eluted at 20 ml/min at ambient temperature. The same solvent gradient used for analytical LCMS

was chosen to purify compounds from a crude extract. 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 for simultaneous analysis by diode array detector (Waters 2998), an ELSD detector (Waters 2424) and a mass spectrometry, Waters SQD-2 mass detector, operating simultaneously in ES+ and ES- modes between 150 and 1000 *m/z*. Detected peaks were collected into glass test tubes. Combined peak samples were concentrated to aqueous phase and frozen to ice by nitrogen for drying.

#### **4.6.7 Nuclear Magnetic Resonance (NMR) Analysis**

NMR measurements were carried out by the NMR department of Organic Chemistry Department, Leibniz University Hannover. <sup>1</sup>H and <sup>13</sup>C NMR spectra were recorded on either on a Bruker DRX 400, 500 or 600 MHz. Chemical shifts are given in parts per million (ppm) relative to the TMS (tetramethylsilane) standard and the coupling constants *J* were shown in Hz. 2D experiments were obtained for complete structural elucidation including Correlation Spectroscopy (COSY), Heteronuclear Single Quantum Coherence (HSQC), and Heteronuclear Multiple Bond Correlation (HMBC).



## References

- 1 D. J. Newman and G. M. Cragg, *J. Nat. Prod.*, 2016, **79**, 629–661.
- 2 M. Nett, H. Ikeda and B. S. Moore, *Nat. Prod. Rep.*, 2009, **26**, 1362–1384.
- 3 H. Raistrick, *Proc. R. Soc. London.*, 1950, **136**, 481–508.
- 4 A. Fleming, *Br. Med. J.*, 1941, **2**, 386.
- 5 N. P. Keller, G. Turner and J. W. Bennett, *Nat. Rev. Microbiol.*, 2005, **3**, 937–947.
- 6 R. J. Cox, *Org. Biomol. Chem.*, 2007, **5**, 2010–2026.
- 7 H. L. Conductor and S. D. Bruner, *Nat. Prod. Rep.*, 2012, **29**, 1099–1110.
- 8 M. B. Quin, C. M. Flynn and C. Schmidt-Dannert, *Nat. Prod. Rep.*, 2014, **31**, 1449–1473.
- 9 P. Tudzynski, K. Hölter, T. Correia, C. Arntz, N. Grammel and U. Keller, *Mol. Gen. Genet.*, 1999, **261**, 133–141.
- 10 Y. Matsuda and I. Abe, *Nat. Prod. Rep.*, 2015, **33**, 26–53.
- 11 R. Geris and T. J. Simpson, *Nat. Prod. Rep.*, 2009, **26**, 1063–1094.
- 12 H. Tomoda, N. Tabata, Y. Nakata, H. Nishida, T. Kaneko, R. Obata, T. Sunazuka and S. Omura, *J. Org. Chem.*, 1996, **61**, 882–886.
- 13 S. Omura, H. Tomoda, Y. K. Kim and H. Nishida, *J. Antibiot. (Tokyo)*, 1993, **46**, 1168–1169.
- 14 H. Tomoda, H. Nishida, Y. K. Kim, R. Obata, T. Sunazuka and S. Omura, *J. Am. Chem. Soc.*, 1994, **116**, 12097–12098.
- 15 F. A. Macías, R. M. Varela, A. M. Simonet, H. G. Cutler, S. J. Cutler, S. A. Ross, D. C. Dunbar, F. M. Dugan and R. A. Hill, *Tetrahedron Lett.*, 2000, **41**, 2683–2686.
- 16 F. A. Macías, R. M. Varela, A. M. Simonet, H. G. Cutler, S. J. Cutler, F. M. Dugan and R. A. Hill, *J. Org. Chem.*, 2000, **65**, 9039–9046.
- 17 G. H. Harris, K. Hoogsteen, K. C. Silverman, S. L. Raghoobar, G. F. Bills, R. B. Lingham, J. L. Smith, H. W. Dougherty, C. Cascales and F. Peláez, *Tetrahedron*, 1993, **49**, 2139–2144.
- 18 F. Mayerl, Q. Gao, S. Huang, S. E. Klohr, J. A. Matson, D. R. Gustavson, D. M. Pirnik, R. L. Berry, C. Fairchild and W. C. Rose, *J. Antibiot. (Tokyo)*, 1993, **47**, 1082–1088.
- 19 A. M. Bailey, R. J. Cox, K. Harley, C. M. Lazarus, T. J. Simpson and E. Skellam, *Chem. Commun.*, 2007, **39**, 4053–4055.
- 20 R. Schor, C. Schotte, D. Wibberg, J. Kalinowski and R. J. Cox, *Nat. Commun.*, 2018, **9**, 1–9.
- 21 W. L. Muth and C. H. Nash, *Antimicrob. Agents Chemother.*, 1975, **8**, 321–327.
- 22 T. S. Bugni, D. Abbanat, V. S. Bernan, W. M. Maiese, M. Greenstein, R. M. Van Wagoner and C. M. Ireland, *J. Org. Chem.*, 2000, **65**, 7195–7200.
- 23 D. K. Holm, L. M. Petersen, A. Klitgaard, P. B. Knudsen, Z. D. Jarczynska, K. F. Nielsen, C. H. Gotfredsen, T. O. Larsen and U. H. Mortensen, *Chem. Biol.*, 2014, **21**, 519–529.
- 24 A. W. Dunn, R. A. W. Johnstone, B. Sklarz and T. J. King, *J. Chem. Soc. Chem. Commun.*, 1976, **7**, 270–270.
- 25 M. Cueto, J. B. MacMillan, P. R. Jensen and W. Fenical, *Phytochemistry*, 2006, **67**, 1826–1831.
- 26 E. Okuyama, M. Yamazaki and Y. Katsube, *Tetrahedron Lett.*, 1984, **25**, 3233–3234.
- 27 C. Rank, R. K. Phipps, P. Harris, P. Fristrup, T. O. Larsen and C. H. Gotfredsen, *Org. Lett.*, 2008, **10**, 401–404.
- 28 K. K. Chexal, J. P. Springer, J. Clardy, R. J. Cole, J. W. Kirksey, J. W. Dorner, H. G. Cutler and B. J. Strawter, *J. Am. Chem. Soc.*, 1976, **98**, 6748–6750.
- 29 D. B. Stierle, A. A. Stierle, J. D. Hobbs, J. Stokken and J. Clardy, *Org. Lett.*, 2004, **6**, 1049–1052.

- 30 K. Shiomi, R. Uchida, J. Inokoshi, H. Tanaka, Y. Iwai and S. Omura, *Tetrahedron Lett.*, 1996, **37**, 1265–1268.
- 31 K. F. Nielsen, P. W. Dalsgaard, J. Smedsgaard and T. O. Larsen, *J. Agric. Food Chem.*, 2005, **53**, 2908–2913.
- 32 Y. Matsuda, T. Awakawa, T. Itoh, T. Wakimoto, T. Kushiro, I. Fujii, Y. Ebizuka and I. Abe, *ChemBioChem*, 2012, **13**, 1738–1741.
- 33 P. Cai, D. Smith, B. Cunningham, S. Brown-Shimer, B. Katz, C. Pearce, D. Venables and D. Houck, *J. Nat. Prod.*, 1998, **61**, 791–795.
- 34 D. P. Overy, F. Berrue, H. Correa, N. Hanif, M. Lanteigne, K. Mquilian, S. Duffy, P. Boland, R. Jagannathan, G. S. Carr, M. Vansteeland, R. G. Kerr, D. P. Overy, F. Berrue, H. Correa, N. Hanif, M. Lanteigne, K. Mquilian, S. Duffy, P. Boland, R. Jagannathan, G. S. Carr, M. Vansteeland and R. G. Kerr, *Mycology*, 2016, **5**, 130–144.
- 35 C. Schotte, L. Li, D. Wibberg, J. Kalinowski and R. J. Cox, *Angew. Chemie - Int. Ed.*, 2020, **59**, 23870–23878.
- 36 P. Pittayakhajonwut, M. Theerasilp, P. Kongsaree, A. Rungrod, M. Tanticharoen and Y. Thebtaranonth, *Planta Med.*, 2002, **68**, 1017–1019.
- 37 S. Ayers, D. L. Zink, J. S. Powell, C. M. Brown, A. Grund, G. F. Bills, G. Platas, D. Thompson and S. B. Singh, *J. Nat. Prod.*, 2008, **71**, 457–459.
- 38 T. Bunyapaiboonsri, S. Veeranondha, T. Boonruangprapa and S. Somrithipol, *Phytochem. Lett.*, 2008, **1**, 204–206.
- 39 J. Zhang, L. Liu, B. Wang, Y. Zhang, L. Wang, X. Liu and Y. Che, *J. Nat. Prod.*, 2015, **78**, 3058–3066.
- 40 J. Zhang, Y. Li, F. Ren, Y. Zhang, X. Liu, L. Liu and Y. Che, *J. Nat. Prod.*, 2019, **82**, 1678–1685.
- 41 T. El-Elimat, H. A. Raja, S. Ayers, S. J. Kurina, J. E. Burdette, Z. Mattes, R. Sabatelle, J. W. Bacon, A. H. Colby, M. W. Grinstaff, C. J. Pearce and N. H. Oberlies, *Org. Lett.*, 2019, **21**, 529–534.
- 42 S. Chen, Z. Liu, H. Tan, Y. Chen, S. Li, H. Li, S. Zhu, H. Liu and W. Zhang, *Org. Chem. Front.*, 2020, **7**, 557–562.
- 43 J. E. Baldwin, A. V. W. Mayweg, K. Neumann and G. J. Pritchard, *Org. Lett.*, 1999, **1**, 1933–1935.
- 44 R. M. Adlington, J. E. Baldwin, A. V. W. Mayweg and G. J. Pritchard, *Org. Lett.*, 2002, **4**, 3009–3011.
- 45 R. M. Adlington, J. E. Baldwin, G. J. Pritchard, A. J. Williams and D. J. Watkin, *Org. Lett.*, 1999, **1**, 1937–1939.
- 46 P. J. Li, G. Dräger and A. Kirschning, *Org. Lett.*, 2019, **21**, 998–1001.
- 47 C. Y. Bemis, C. N. Ungarean, A. S. Shved, C. S. Jamieson, T. Hwang, K. S. Lee, K. N. Houk and D. Sarlah, *J. Am. Chem. Soc.*, 2021, **143**, 6006–6017.
- 48 S. Pornpakakul, S. Roengsumran, S. Deechangvipart, A. Petsom, N. Muangsin, N. Ngamrojnavanich, N. Sriubolmas, N. Chaichit and T. Ohta, *Tetrahedron Lett.*, 2007, **48**, 651–655.
- 49 U. Höller, A. D. Wright, G. F. Mathee, G. M. König, S. Draeger, H.-J. Aust and B. Schulz, *Mycol. Res.*, 2000, **104**, 1354–1365.
- 50 M. Kaneko, D. Matsuda, M. Ohtawa, T. Fukuda, T. Nagamitsu, T. Yamori and H. Tomoda, *Biol. Pharm. Bull.*, 2012, **35**, 18–28.
- 51 C. J. Hsiao, S. H. Hsiao, W. L. Chen, J. H. Guh, G. Hsiao, Y. J. Chan, T. H. Lee and C. L. Chung, *Chem. Biol. Interact.*, 2012, **197**, 23–30.

- 52 M. E. Raggatt, T. J. Simpson and M. I. Chicarelli-Robinson, *Chem. Commun.*, 1997, 2245–2246.
- 53 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, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 7642–7647.
- 54 Y. Zhai, Y. Li, J. Zhang, Y. Zhang, F. Ren, X. Zhang, G. Liu, X. Liu and Y. Che, *Fungal Genet. Biol.*, 2019, **129**, 7–15.
- 55 Q. Chen, J. Gao, C. Jamieson, J. Liu, M. Ohashi, J. Bai, D. Yan, B. Liu, Y. Che, Y. Wang, K. N. Houk and Y. Hu, *J. Am. Chem. Soc.*, 2019, **141**, 14052–14056.
- 56 H. Birkinshaw, J. H. and Chambers, A. R. and Raistrick, *Biochem. J.*, 1942, **36**, 242.
- 57 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, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 7642–7647.
- 58 A. Al Fahad, A. Abood, T. J. Simpson and R. J. Cox, *Angew. Chemie - Int. Ed.*, 2014, **53**, 7519–7523.
- 59 N. L. Brock, K. Huss, B. Tudzynski and J. S. Dickschat, *ChemBioChem*, 2013, **14**, 311–315.
- 60 D. B. Amby, T. Manczak, M. A. Petersen, T. Sundelin, C. Weitzel, M. Grajewski, H. T. Simonsen and B. Jensen, *Microbiol. (United Kingdom)*, 2016, **162**, 1773–1783.
- 61 M. B. Quin, C. M. Flynn, G. T. Wawrzyn, S. Choudhary and C. Schmidt-Dannert, *ChemBioChem*, 2013, **14**, 2480–2491.
- 62 C. Schotte, P. Lukat, A. Deuschmann, W. Blankenfeldt and R. J. Cox, *Angew. Chemie Int. Ed.*, 2021, **60**, 1–6.
- 63 D. Arigoni, *Pure Appl. Chem.*, 1975, **41**, 219–245.
- 64 D. E. Cane, R. Iyengar and M. S. Shiao, *J. Am. Chem. Soc.*, 1978, **100**, 7122–7125.
- 65 J. Rinkel and J. S. Dickschat, *Beilstein J. Org. Chem.*, 2019, **15**, 789–794.
- 66 A. R. Hawkins, H. K. Lamb, J. D. Moore, I. G. Charles and C. F. Roberts, *J. Gen. Microbiol.*, 1993, **139**, 2891–2899.
- 67 K. M. Herrmann, *Plant Cell*, 1995, **7**, 907–919.
- 68 A. R. Knaggs, *Nat. Prod. Rep.*, 2003, **20**, 119–136.
- 69 H. G. Floss, *Nat. Prod. Rep.*, 1997, **14**, 433–452.
- 70 R. B. McQualter, B. F. Chong, K. Meyer, D. E. Van Dyk, M. G. O’Shea, N. J. Walton, P. V. Viitanen and S. M. Brumbley, *Plant Biotechnol. J.*, 2005, **3**, 29–41.
- 71 S. Wang, M. Bilal, H. Hu, W. Wang and X. Zhang, *Appl. Microbiol. Biotechnol.*, 2018, **102**, 3561–3571.
- 72 J. L. Barker and J. W. Frost, *Biotechnol. Bioeng.*, 2001, **76**, 376–390.
- 73 N. Kallscheuer, M. Vogt and J. Marienhagen, *ACS Synth. Biol.*, 2017, **6**, 410–415.
- 74 S. Sengupta, S. Jonnalagadda, L. Goonewardena and V. Juturu, *Appl. Environ. Microbiol.*, 2015, **81**, 8037–8043.
- 75 S. Q. E. Lee, T. S. Tan, M. Kawamukai and E. S. Chen, *Microb. Cell Fact.*, 2017, **16**, 1–16.
- 76 Y. Bai, H. Yin, H. Bi, Y. Zhuang, T. Liu and Y. Ma, *Metab. Eng.*, 2016, **35**, 138–147.
- 77 E. S. E. Habib, J. N. Scarsdale and K. A. Reynolds, *Antimicrob. Agents Chemother.*, 2003, **47**, 2065–2071.
- 78 Q. Zhu, Y. Song, H. Huang, Q. Li and J. Ju, *Org. Biomol. Chem.*, 2019, **17**, 3760–3764.
- 79 K. C.-C. Chou, H.-L. Wu, P.-Y. Lin, S.-H. Yang, T.-L. Chang, F. Sheu, K.-H. Chen and B.-H. Chiang, *Phytochemistry*, 2019, **161**, 97–106.
- 80 K. Pei, J. Ou, J. Huang and S. Ou, *J. Sci. Food Agric.*, 2016, **96**, 2952–2962.
- 81 A. Rodriguez, K. R. Kildegaard, M. Li, I. Borodina and J. Nielsen, *Metab. Eng.*, 2015, **31**, 181–188.

- 82 L. Morelli, K. Zór, C. B. Jendresen, T. Rindzevicius, M. S. Schmidt, A. T. Nielsen and A. Boisen, *Anal. Chem.*, 2017, **89**, 3981–3987.
- 83 J. Wu, T. Zhou, G. Du, J. Zhou and J. Chen, *PLoS One*, 2014, **9**, 1–9.
- 84 S. M. Yang, G. Y. Shim, B. G. Kim and J. H. Ahn, *Microb. Cell Fact.*, 2015, **14**, 1–12.
- 85 M. Berner, D. Krug, C. Bihlmaier, A. Vente, R. Müller and A. Bechthold, *J. Bacteriol.*, 2006, **188**, 2666–2673.
- 86 O. Choi, C. Z. Wu, S. Y. Kang, J. S. Ahn, T. B. Uhm and Y. S. Hong, *J. Ind. Microbiol. Biotechnol.*, 2011, **38**, 1657–1665.
- 87 E. G. Lee, S. H. Yoon, A. Das, S. H. Lee, C. Li, J. Y. Kim, M. S. Choi, D. K. Oh and S. W. Kim, *Biotechnol. Bioeng.*, 2009, **102**, 200–208.
- 88 G. F. Bills and J. B. Gloer, in *The Fungal Kingdom*, 2017, pp. 1087–1119.
- 89 K. K. Wallace, K. A. Reynolds, K. Koch, H. A. I. McArthur, R. G. Wax and B. S. Moore, *J. Am. Chem. Soc.*, 1994, **116**, 11600–11601.
- 90 Juan F. Martin and P. Liras, *J. Antibiot. (Tokyo)*, 1976, **29**, 1306–1309.
- 91 Y. Shen, F. Sun, L. Zhang, Y. Cheng, H. Zhu, S. Wang, W. Jiao, P. F. Leadlay, X. Y. Zhou and H. Lin, *J. Biol. Chem.*, 2020, **295**, 5509–5518.
- 92 J. N. Andexer, S. G. Kendrew, M. Nur-E-Alam, O. Lazos, T. A. Foster, A. S. Zimmermann, T. D. Warneck, D. Suthar, N. J. Coates, F. E. Koehn, J. S. Skotnicki, G. T. Carter, M. A. Gregory, C. J. Martin, S. J. Moss, P. F. Leadlay and B. Wilkinson, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 4776–4781.
- 93 Y. Lin, X. Sun, Q. Yuan and Y. Yan, *Metab. Eng.*, 2019, **23**, 62–69.
- 94 F. Shah, D. Schwenk, C. Nicolás, P. Persson, D. Hoffmeister and A. Tunlid, *Appl. Environ. Microbiol.*, 2015, **81**, 8427–8433.
- 95 J. Braesel, S. Götze, F. Shah, D. Heine, J. Tauber, C. Hertweck, A. Tunlid, P. Stallforth and D. Hoffmeister, *Chem. Biol.*, 2015, **22**, 1325–1334.
- 96 M. Gill and W. Steglich, in *Fortschritte der Chemie organischer Naturstoffe / Progress in the Chemistry of Organic Natural Products*, Springer Vienna, Vienna, 1987, pp. 1–297.
- 97 E. Stocker-Wörgötter, *Nat. Prod. Rep.*, 2008, **25**, 188–200.
- 98 W.-W. Sun, C.-J. Guo and C. C. C. Wang, *Fungal Genet. Biol.*, 2016, **89**, 84–88.
- 99 W. A. Ayer, Y. Fukazawa and H. Orszanska, *Nat. Prod. Lett.*, 1993, **2**, 77–82.
- 100 G. Mazzeo, E. Santoro, A. Andolfi, A. Cimmino, P. Troselj, A. G. Petrovic, S. Superchi, A. Evidente and N. Berova, *J. Nat. Prod.*, 2013, **76**, 588–599.
- 101 A. Evidente, *Nat. Prod. Rep.*, 2015, **32**, 1629–1653.
- 102 L. Kahlert, E. F. Bassiony, R. J. Cox and E. J. Skellam, *Angew. Chemie Int. Ed.*, 2020, **132**, 5865–5871.
- 103 A. Al Fahad, A. Abood, K. M. Fisch, A. Osipow, J. Davison, M. Avramović, C. P. Butts, J. Piel, T. J. Simpson and R. J. Cox, *Chem. Sci.*, 2014, **5**, 523–527.
- 104 E. Kitamura, A. Hirota, M. Nakagawa and M. Nakayama, *Tetrahedron Lett.*, 1990, **31**, 4605–4608.
- 105 A. Arnone, R. Cardillo, G. Nasini and O. V. de Pava, *Tetrahedron*, 1993, **49**, 7251–7258.
- 106 R. Meier and B. Ganem, *Tetrahedron*, 1994, **50**, 2715–2720.
- 107 C. Tu and G. A. Berchtold, *J. Org. Chem.*, 1993, **58**, 6915–6916.
- 108 S. C. Schuster, *Nat. Methods*, 2008, **5**, 16–18.
- 109 E. R. Mardis, *Trends Genet.*, 2008, **24**, 133–141.
- 110 M. L. Metzker, *Nat. Rev. Genet.*, 2010, **11**, 31–46.

- 111 J. J. Kozich, S. L. Westcott, N. T. Baxter, S. K. Highlander and P. D. Schloss, *Appl. Environ. Microbiol.*, 2013, **79**, 5112–5120.
- 112 T. Ni, D. L. Corcoran, E. A. Rach, S. Song, E. P. Spana, Y. Gao, U. Ohler and J. Zhu, *Nat. Methods*, 2010, **7**, 521–527.
- 113 N. Kono and K. Arakawa, *Dev. Growth Differ.*, 2019, **61**, 316–326.
- 114 B. Y. Guo, T. Zeng and H. C. Wu, *Sci. Bull.*, 2015, **60**, 287–295.
- 115 S. Goodwin, J. D. McPherson and W. R. McCombie, *Nat. Rev. Genet.*, 2016, **17**, 333–351.
- 116 L. Hamer, K. Adachi, M. V. Montenegro-Chamorro, M. M. Tanzer, S. K. Mahanty, C. Lo, R. W. Tarpey, A. R. Skalchunes, R. W. Heiniger, S. A. Frank, B. A. Darveaux, D. J. Lampe, T. M. Slater, L. Ramamurthy, T. M. DeZwaan, G. H. Nelson, J. R. Shuster, J. Woessner and J. E. Hamer, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 5110–5115.
- 117 S. M. Bowman and S. J. Free, *BioEssays*, 2006, **28**, 799–808.
- 118 D. J. Adams, *Microbiology*, 2004, **150**, 2029–2035.
- 119 A. Sivan and I. Chet, *Microbiology*, 1989, **135**, 675–682.
- 120 V. Ramamoorthy, L. Govindaraj, M. Dhanasekaran, S. Vetrivel, K. K. Kumar and E. Ebenezar, *J. Microbiol. Methods*, 2015, **111**, 127–134.
- 121 M. L. Nielsen, L. Albertsen, G. Lettier, J. B. Nielsen and U. H. Mortensen, *Fungal Genet. Biol.*, 2006, **43**, 54–64.
- 122 H. L. Wang, B. L. Postier and R. L. Burnap, *Biotechniques*, 2002, **33**, 26–32.
- 123 X. Zhang, R. Henriques, S. S. Lin, Q. W. Niu and N. H. Chua, *Nat. Protoc.*, 2006, **1**, 641–646.
- 124 R. Liu, L. Chen, Y. Jiang, Z. Zhou and G. Zou, *Cell Discov.*, 2015, **1**, 1–11.
- 125 G. Meister and T. Tuschl, *Nature*, 2004, **431**, 343–349.
- 126 W. B. Frommer and O. Ninnemann, *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 1995, **46**, 419–444.
- 127 Y. He and R. J. Cox, *Chem. Sci.*, 2016, **7**, 2119–2127.
- 128 Y. M. Chiang, C. E. Oakley, M. Ahuja, R. Entwistle, A. Schultz, S. L. Chang, C. T. Sung, C. C. C. Wang and B. R. Oakley, *J. Am. Chem. Soc.*, 2013, **135**, 7720–7731.
- 129 D. Mumberg, R. Müller and M. Funk, *Gene*, 1995, **156**, 119–122.
- 130 M. N. Heneghan, A. A. Yakasai, L. M. Halo, Z. Song, A. M. Bailey, T. J. Simpson, R. J. Cox and C. M. Lazarus, *ChemBioChem*, 2010, **11**, 1508–1512.
- 131 K. Williams, A. J. Szwalbe, N. P. Mulholland, J. L. Vincent, A. M. Bailey, C. L. Willis, T. J. Simpson and R. J. Cox, *Angew. Chemie - Int. Ed.*, 2016, **55**, 6784–6788.
- 132 K. E. Lebe and R. J. Cox, *Chem. Sci.*, 2019, **10**, 1227–1231.
- 133 F. J. Jin, J. I. Maruyama, P. R. Juvvadi, M. Arioka and K. Kitamoto, *FEMS Microbiol. Lett.*, 2004, **239**, 79–85.
- 134 K. Gomi, Y. Iimura and S. Hara, *Agric. Biol. Chem.*, 1987, **51**, 2549–2555.
- 135 O. Yamada, B. Rho Lee and K. Gomi, *Biosci. Biotechnol. Biochem.*, 1997, **61**, 1367–1369.
- 136 K. Sakai, H. Kinoshita and T. Nihira, *Appl. Microbiol. Biotechnol.*, 2012, **93**, 2011–2022.
- 137 T. Christensen, H. Woeldike, E. Boel, S. B. Mortensen, L. Thim and M. T. Hansen, *Bio/Technology*, 1988, **6**, 1419–1422.
- 138 C. M. Lazarus, K. Williams and A. M. Bailey, *Nat. Prod. Rep.*, 2014, **31**, 1339–1347.
- 139 S. Sahdev, S. K. Khattar and K. S. Saini, *Mol. Cell. Biochem.*, 2008, **307**, 249–264.
- 140 J. Kaur, A. Kumar and J. Kaur, *Int. J. Biol. Macromol.*, 2018, **106**, 803–822.
- 141 T. J. Simpson, *Chem. Soc. Rev.*, 1987, **16**, 123–160.
- 142 F. Trenti and R. J. Cox, *J. Nat. Prod.*, 2017, **80**, 1235–1240.

- 143 J. C. Vederas, *Can. J. Chem.*, 1982, **60**, 1637–1642.
- 144 D. S. Tian, E. Kuhnert, J. Ouazzani, D. Wibberg, J. Kalinowski and R. J. Cox, *Chem. Sci.*, 2020, **11**, 12477–12484.
- 145 Y. Takenaka, T. Tanahashi, N. Nagakura and N. Hamada, *Chem. Pharm. Bull.*, 2003, **51**, 794–797.
- 146 L. J. Fremlin, A. M. Piggott, E. Lacey and R. J. Capon, *J. Nat. Prod.*, 2009, **72**, 666–670.
- 147 T. Bunyapaiboonsri, S. Yoiprommarat, K. Intereya and K. Kocharin, *Chem. Pharm. Bull.*, 2007, **55**, 304–307.
- 148 C. L. Schoch, K. A. Seifert, S. Huhndorf, V. Robert, J. L. Spouge, C. A. Levesque, W. Chen, E. Bolchacova, K. Voigt, P. W. Crous, A. N. Miller, M. J. Wingfield, M. C. Aime, K. D. An, F. Y. Bai, R. W. Barreto, D. Begerow, M. J. Bergeron, M. Blackwell, T. Boekhout, M. Bogale, N. Boonyuen, A. R. Burgaz, B. Buyck, L. Cai, Q. Cai, G. Cardinali, P. Chaverri, B. J. Coppins, A. Crespo, P. Cubas, C. Cummings, U. Damm, Z. W. de Beer, G. S. de Hoog, R. Del-Prado, B. Dentinger, J. Diéguez-Uribeondo, P. K. Divakar, B. Douglas, M. Dueñas, T. A. Duong, U. Eberhardt, J. E. Edwards, M. S. Elshahed, K. Fliegerova, M. Furtado, M. A. García, Z. W. Ge, G. W. Griffith, K. Griffiths, J. Z. Groenewald, M. Groenewald, M. Grube, M. Gryzenhout, L. D. Guo, F. Hagen, S. Hambleton, R. C. Hamelin, K. Hansen, P. Harrold, G. Heller, C. Herrera, K. Hirayama, Y. Hirooka, H. M. Ho, K. Hoffmann, V. Hofstetter, F. Högnabba, P. M. Hollingsworth, S. B. Hong, K. Hosaka, J. Houbraken, K. Hughes, S. Huhtinen, K. D. Hyde, T. James, E. M. Johnson, J. E. Johnson, P. R. Johnston, E. B. G. Jones, L. J. Kelly, P. M. Kirk, D. G. Knapp, U. Kõljalg, G. M. Kovács, C. P. Kurtzman, S. Landvik, S. D. Leavitt, A. S. Liggenstoffer, K. Liimatainen, L. Lombard, J. J. Luangsa-ard, H. T. Lumbsch, H. Maganti, S. S. N. Maharachchikumbura, M. P. Martin, T. W. May, A. R. McTaggart, A. S. Methven, W. Meyer, J. M. Moncalvo, S. Mongkolsamrit, L. G. Nagy, R. H. Nilsson, T. Niskanen, I. Nyilasi, G. Okada, I. Okane, I. Olariaga, J. Otte, T. Papp, D. Park, T. Petkovits, R. Pino-Bodas, W. Quaedvlieg, H. A. Raja, D. Redecker, T. L. Rintoul, C. Ruibal, J. M. Sarmiento-Ramírez, I. Schmitt, A. Schübler, C. Shearer, K. Sotome, F. O. P. Stefani, S. Stenroos, B. Stielow, H. Stockinger, S. Suetrong, S. O. Suh, G. H. Sung, M. Suzuki, K. Tanaka, L. Tedersoo, M. T. Telleria, E. Tretter, W. A. Untereiner, H. Urbina, C. Vágvölgyi, A. Vialle, T. D. Vu, G. Walther, Q. M. Wang, Y. Wang, B. S. Weir, M. Weiß, M. M. White, J. Xu, R. Yahr, Z. L. Yang, A. Yurkov, J. C. Zamora, N. Zhang, W. Y. Zhuang and D. Schindel, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 6241–6246.
- 149 M. S. Blouin, *Int. J. Parasitol.*, 2002, **32**, 527–531.
- 150 S. Guindon, J. F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk and O. Gascuel, *Syst. Biol.*, 2010, **59**, 307–321.
- 151 D. Wibberg, L. Andersson, G. Tzelepis, O. Rupp, J. Blom, L. Jelonek, A. Pühler, J. Fogelqvist, M. Varrelmann, A. Schlüter and C. Dixelius, *BMC Genomics*, 2016, **17**, 1–12.
- 152 M. Stanke, O. Schöffmann, B. Morgenstern and S. Waack, *BMC Bioinformatics*, 2006, **7**, 1–11.
- 153 T. Carver, M. Berriman, A. Tivey, C. Patel, U. Böhme, B. G. Barrell, J. Parkhill and M. A. Rajandream, *Bioinformatics*, 2008, **24**, 2672–2676.
- 154 G. Zhang, V. Gurtu and S. R. Kain, *Biochem. Biophys. Res. Commun.*, 1996, **227**, 707–711.
- 155 Y. Kallberg, U. Oppermann, H. Jörnvall and B. Persson, *Eur. J. Biochem.*, 2002, **269**, 4409–4417.
- 156 E. M. Stocking and R. M. Williams, *Angew. Chemie - Int. Ed.*, 2003, **42**, 3078–3115.
- 157 A. Al Fahad, A. Abood, T. J. Simpson and R. J. Cox, *Angew. Chemie - Int. Ed.*, 2014, **53**, 7519–7523.
- 158 R. Schor, 2018. Ph.D. Thesis, Leibniz University of Hannover, Biosynthesis of Xenovulenes

in *Acremonium strictum*

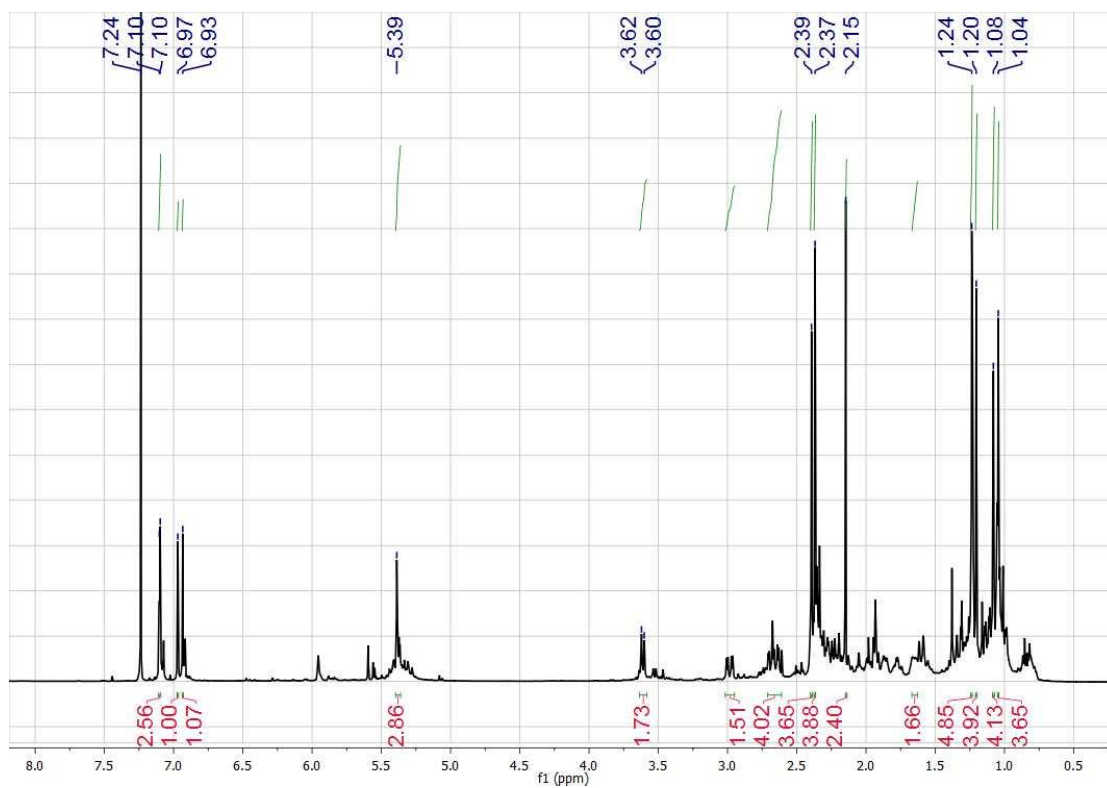
- 159 T. J. Doyon, 2020. Ph.D. Thesis, University of Michigan, Development and Characterization of Non-heme Iron Biocatalysts for Complex Molecule Synthesis
- 160 C. Feng, Q. Wei, C. Hu and Y. Zou, *Org. Lett.*, 2019, **21**, 3114–3118.
- 161 M. E. Raggatt, T. J. Simpson and M. I. Chicarelli-Robinson, *Chem. Commun.*, 1997, **22**, 2245–2246.
- 162 A. Evidente, A. Cimmino, A. Andolfi, M. Vurro, M. C. Zonno, C. L. Cantrell and A. Motta, *Tetrahedron*, 2008, **64**, 1612–1619.
- 163 A. Evidente, A. Cimmino, A. Andolfi, M. Vurro, M. C. Zonno and A. Motta, *J. Agric. Food Chem.*, 2008, **56**, 884–888.
- 164 P. M. Chouinard and P. A. Bartlett, *J. Org. Chem.*, 1986, **51**, 75–78.
- 165 C. Derntl, A. Rassinger, E. Srebotnik, R. L. Mach and A. R. Mach-Aigner, *Appl. Environ. Microbiol.*, 2016, **82**, 6247–6257.
- 166 C. Derntl, F. Guzmán-Chávez, T. M. Mello-de-Sousa, H. J. Busse, A. J. M. Driessen, R. L. Mach and A. R. Mach-Aigner, *Front. Microbiol.*, 2017, **08**, 2037–2048.
- 167 K. Washida, N. Abe, Y. Sugiyama and A. Hirota, *Biosci. Biotechnol. Biochem.*, 2009, **73**, 1355–1361.
- 168 M. G. Steiger, M. Vitikainen, P. Uskonen, K. Brunner, G. Adam, T. Pakula, M. Penttilä, M. Saloheimo, R. L. MacH and A. R. Mach-Aigner, *Appl. Environ. Microbiol.*, 2011, **77**, 114–121.
- 169 I. Häuslein, F. Cantet, S. Reschke, F. Chen, M. Bonazzi and W. Eisenreich, *Front. Cell. Infect. Microbiol.*, 2017, **7**, 1–15.
- 170 M. Hochuli, H. Patzelt, D. Oesterheld, K. Wüthrich and T. Szyperski, *J. Bacteriol.*, 1999, **181**, 3226–3237.
- 171 M. Thongchuang, P. Pongsawasdi, Y. Chisti and K. Packdibamrung, *World J. Microbiol. Biotechnol.*, 2012, **28**, 2937–2943.
- 172 K. Gottlieb, C. Albermann and G. A. Sprenger, *Microb. Cell Fact.*, 2014, **13**, 1–16.
- 173 A. R. Hawkins, J. D. Moore and H. K. Lamb, *Biochem. Soc. Trans.*, 1993, **21**, 181–186.
- 174 A. R. Hawkin, *Curr. Genet.*, 1987, **11**, 491–498.
- 175 M. Hartmann, G. Heinrich and G. H. Braus, *Arch. Microbiol.*, 2001, **175**, 112–121.
- 176 S. Grant, C. F. Roberts, H. Lamb, M. Stout and A. R. Hawkins, *J. Gen. Microbiol.*, 1988, **134**, 347–358.
- 177 J. Guo, Y. Carrington, A. Alber and J. Ehlting, *J. Biol. Chem.*, 2014, **289**, 23846–23858.
- 178 A. R. Hawkins, H. K. Lamb, M. Smith, J. W. Keyte and C. F. Roberts, *MGG Mol. Gen. Genet.*, 1988, **214**, 224–231.
- 179 A. R. Hawkins, J. D. Moore and A. M. Adeokun, *Biochem. J.*, 1993, **296**, 451–457.
- 180 M. R. Bennett, S. A. Shepherd, V. A. Cronin and J. Micklefield, *Curr. Opin. Chem. Biol.*, 2017, **37**, 97–106.
- 181 J. Deen, C. Vranken, V. Leen, R. K. Neely, K. P. F. Janssen and J. Hofkens, *Angew. Chemie - Int. Ed.*, 2017, **56**, 5182–5200.
- 182 K. E. Lebe and R. J. Cox, *RSC Adv.*, 2019, **9**, 31527–31531.
- 183 C. A. Uchima, G. Tokuda, H. Watanabe, K. Kitamoto and M. Arioka, *Appl. Microbiol. Biotechnol.*, 2011, **89**, 1761–1771.
- 184 K. Sakai, H. Kinoshita, T. Shimizu and T. Nihira, *J. Biosci. Bioeng.*, 2008, **106**, 466–472.
- 185 S. Bing Hua, M. Qiu, E. Chan, L. Zhu and Y. Luo, *Plasmid*, 1997, **38**, 91–96.
- 186 F. Trenti, 2018. Ph.D. Thesis, Leibniz University of Hannover, The Biosynthesis of Phyllostictine

- A and Sch-642305 from *Phyllosticta cirsii* and *Phomopsis CMU-LMA*
- 187 B. S. Montenecourt, *Trends Biotechnol.*, 1983, **1**, 156–161.
- 188 C. Derntl, D. P. Kiesenhofer, R. L. Mach and A. R. Mach-Aigner, *Appl. Environ. Microbiol.*, 2015, **81**, 6314–6323.
- 189 M. Penttilä, H. Nevalainen, M. Rättö, E. Salminen and J. Knowles, *Gene*, 1987, **61**, 155–164.
- 190 R. L. Mach, M. Schindler and C. P. Kubicek, *Curr. Genet.*, 1994, **25**, 567–570.
- 191 K. Ishibashi, K. Suzuki, Y. Ando, C. Takakura and H. Inoue, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 14871–14876.
- 192 R. D. Gietz and R. H. Schiestl, *Nat. Protoc.*, 2007, **2**, 31–34.
- 193 K. Graumann and A. Premstaller, *Biotechnol. J.*, 2006, **1**, 164–186.
- 194 L. Imamovic, M. A. Misiakou, E. van der Helm, G. Panagiotou, M. Muniesa and M. O. A. Sommer, *Genome Announc.*, 2018, **6**, 10816.

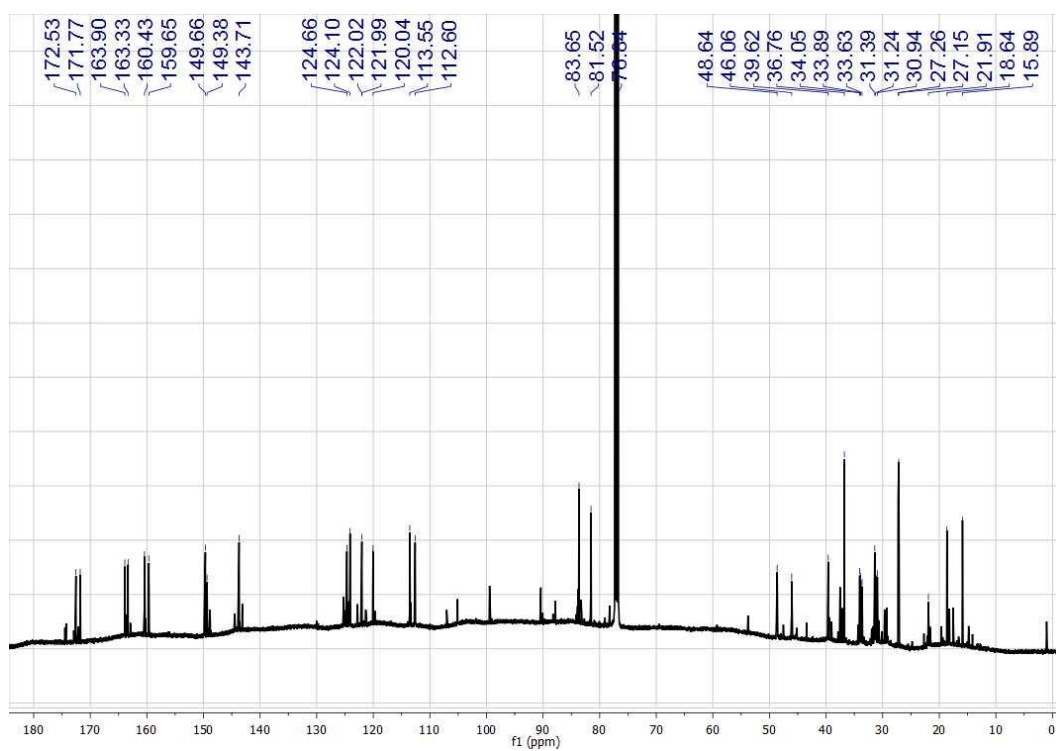


## Appendix

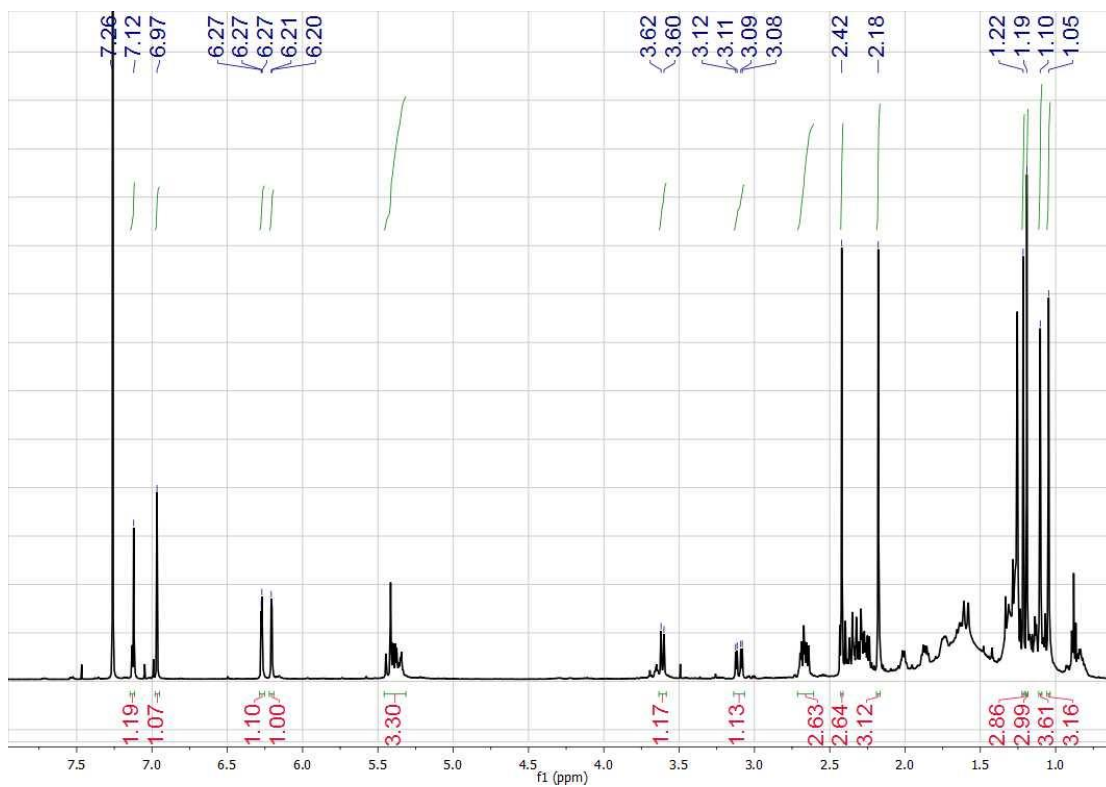
$^1\text{H}$  NMR of pycnidione **22** in  $\text{CDCl}_3$  (500 MHz) referenced to  $\text{CDCl}_3$



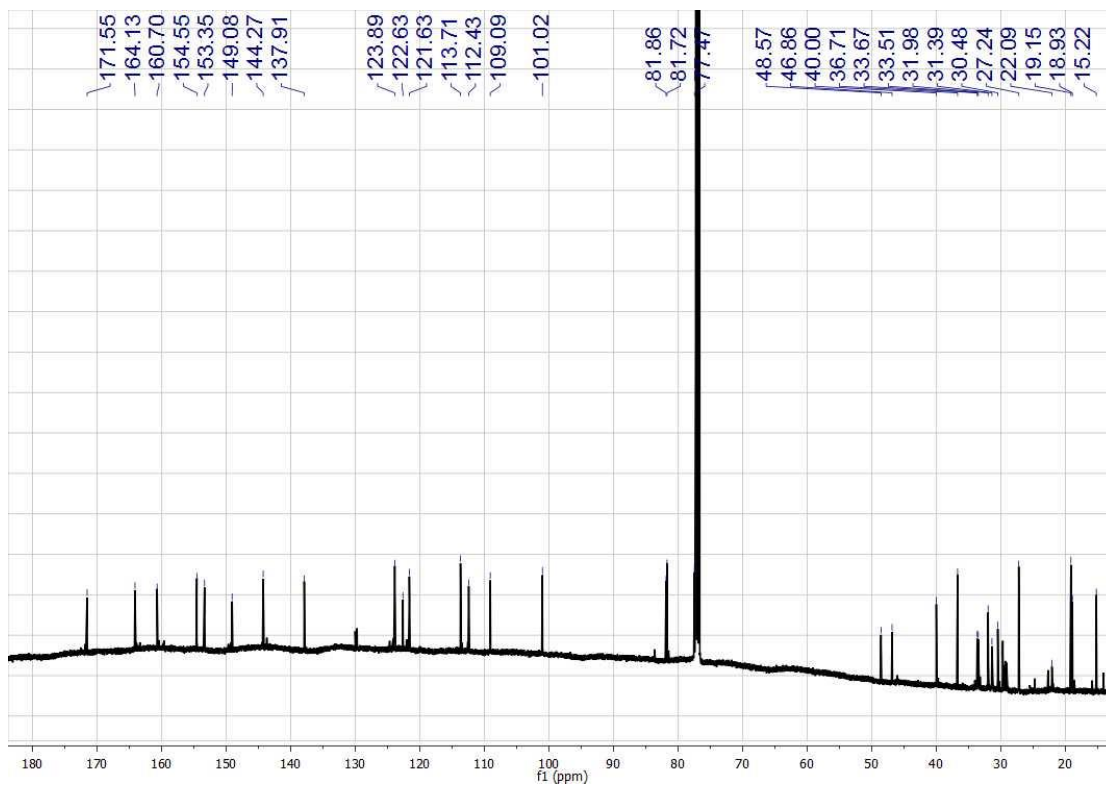
$^{13}\text{C}$  NMR of pycnidione **22** in  $\text{CDCl}_3$  (125 MHz) referenced to  $\text{CDCl}_3$



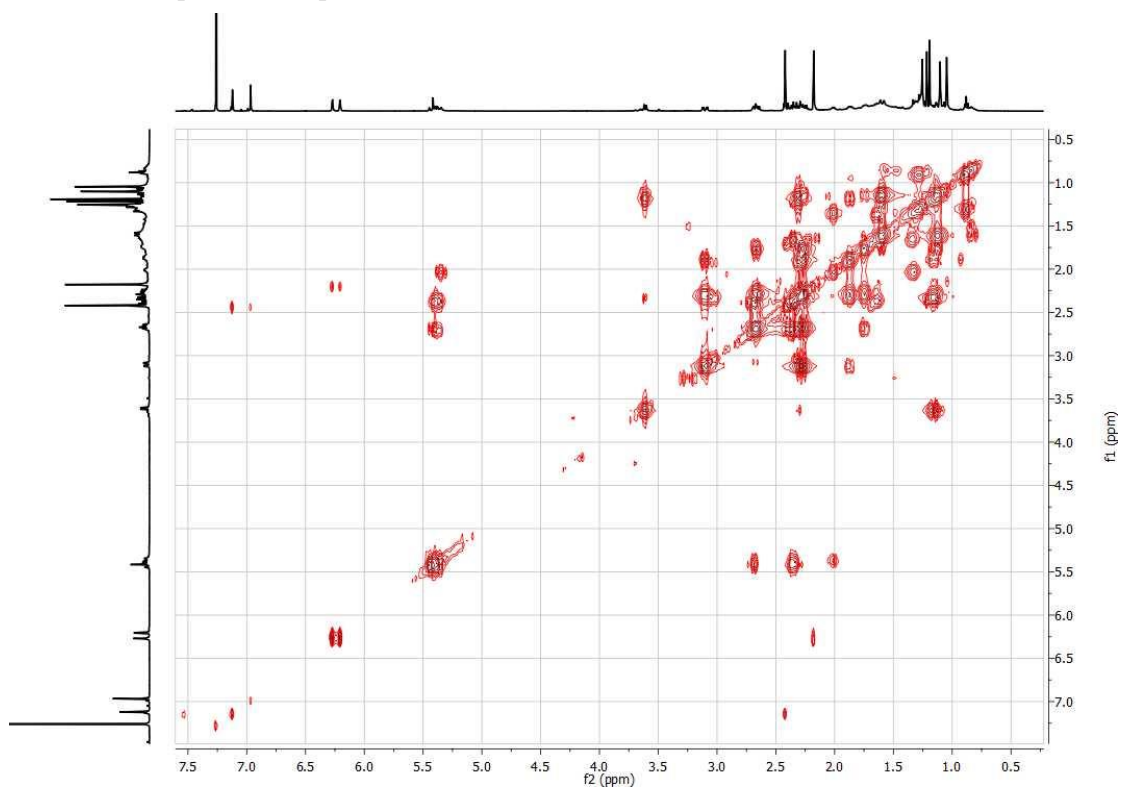
$^1\text{H}$  NMR of epolone C **42** in  $\text{CDCl}_3$  (500 MHz) referenced to  $\text{CDCl}_3$



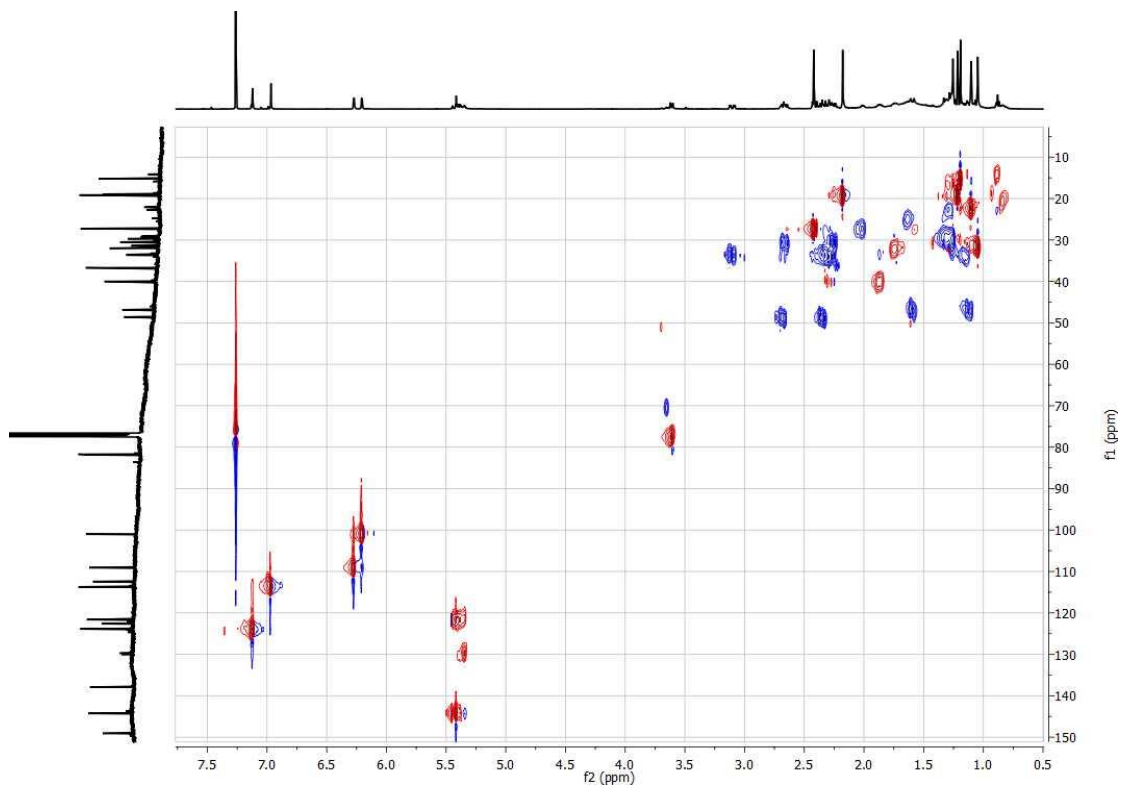
$^{13}\text{C}$  NMR of epolone C **42** in  $\text{CDCl}_3$  (125 MHz) referenced to  $\text{CDCl}_3$



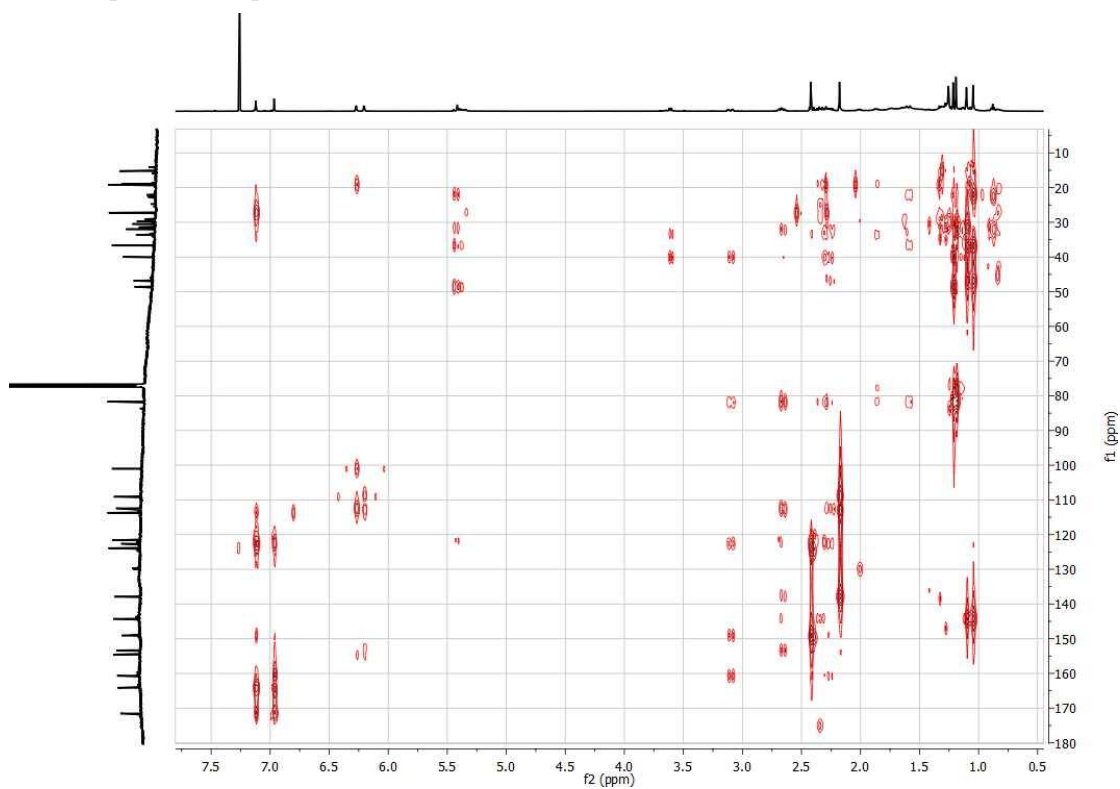
$^1\text{H}$ - $^1\text{H}$  COSY spectrum of epolone C 42



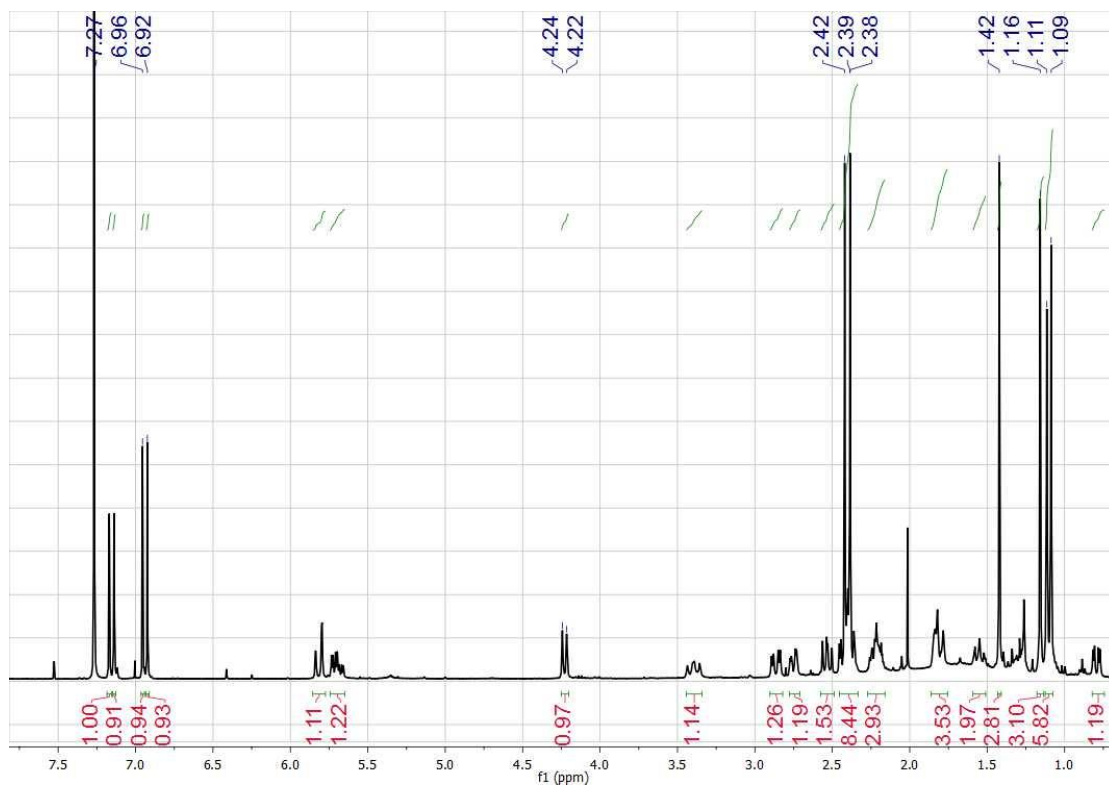
HSQC spectrum of epolone C 42



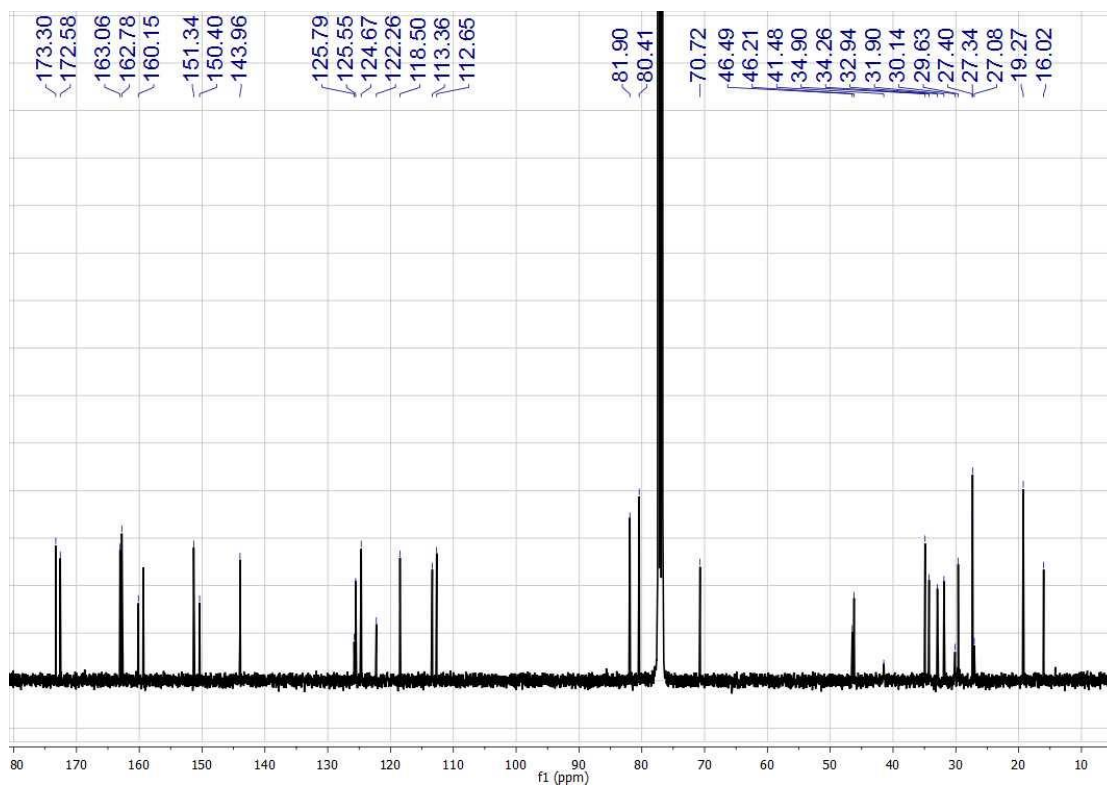
HMBC spectrum of epolone C 42.



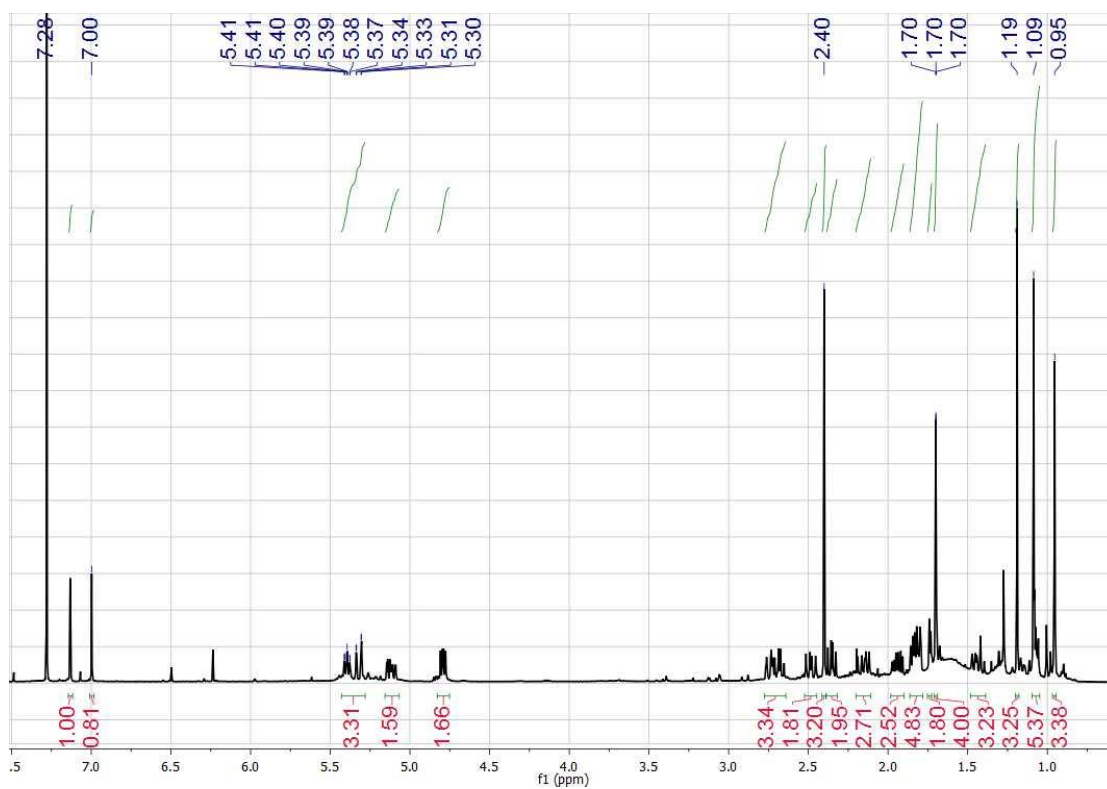
<sup>1</sup>H NMR of eupenifeldin 23 in CDCl<sub>3</sub> (500 MHz) referenced to CDCl<sub>3</sub>



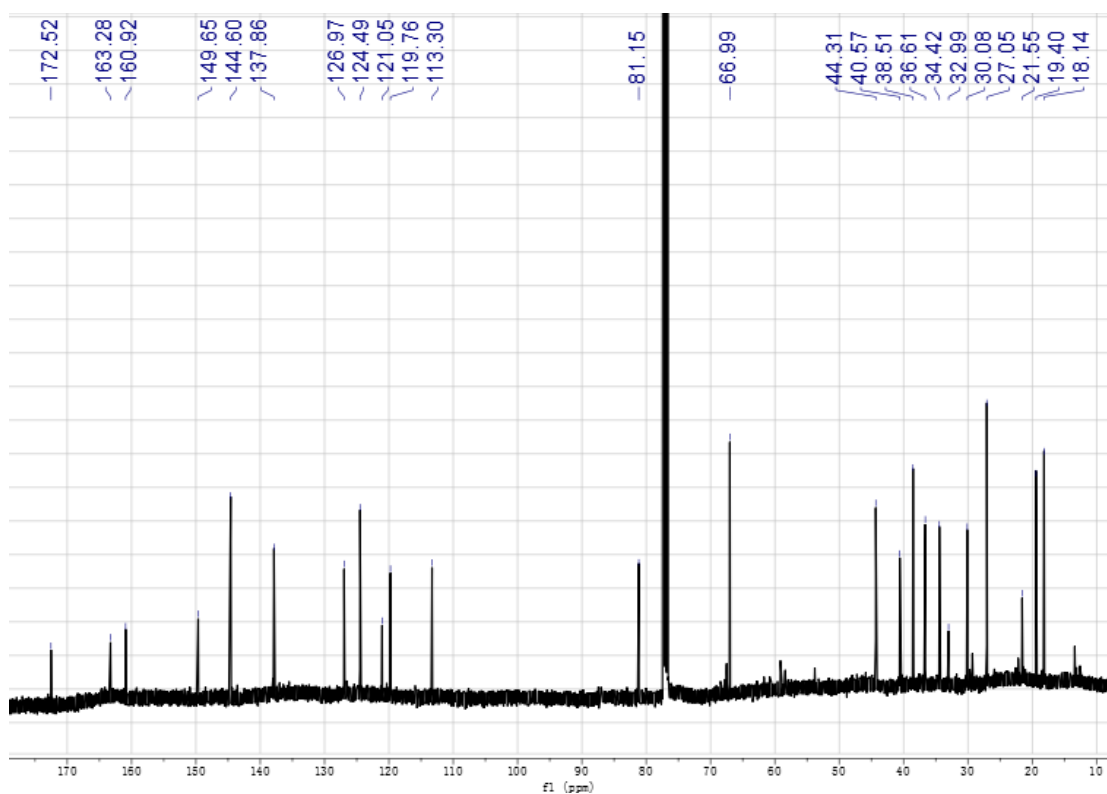
$^{13}\text{C}$  NMR of eupenifeldin **23** in  $\text{CDCl}_3$  (125 MHz) referenced to  $\text{CDCl}_3$



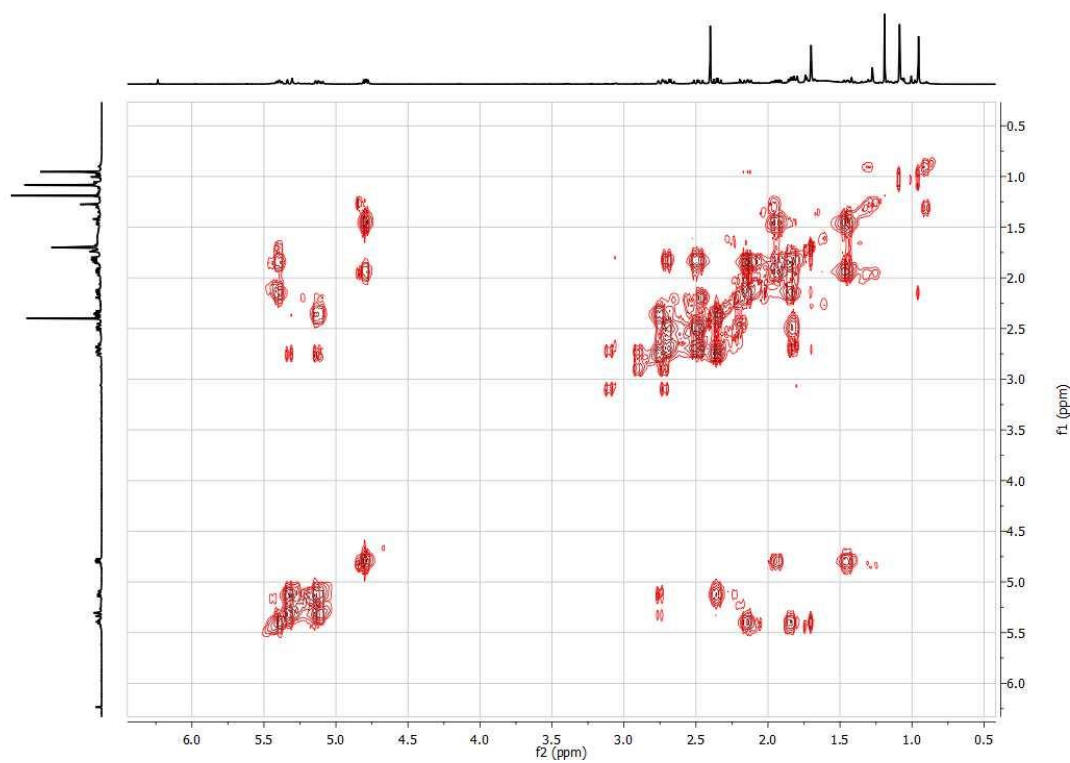
$^1\text{H}$  NMR of neosetophomone B **61** in  $\text{CDCl}_3$  (500 MHz) referenced to  $\text{CDCl}_3$



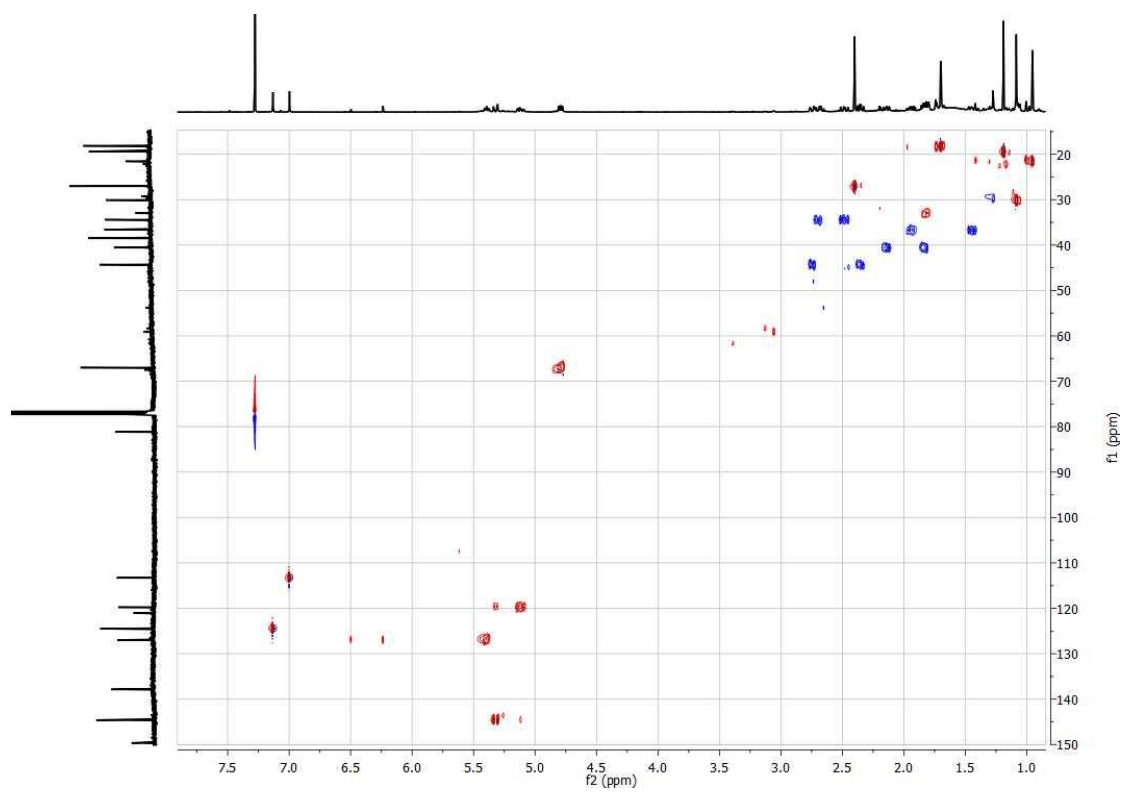
$^{13}\text{C}$  NMR of neosetophomone B **61** in  $\text{CDCl}_3$  (125 MHz) referenced to  $\text{CDCl}_3$



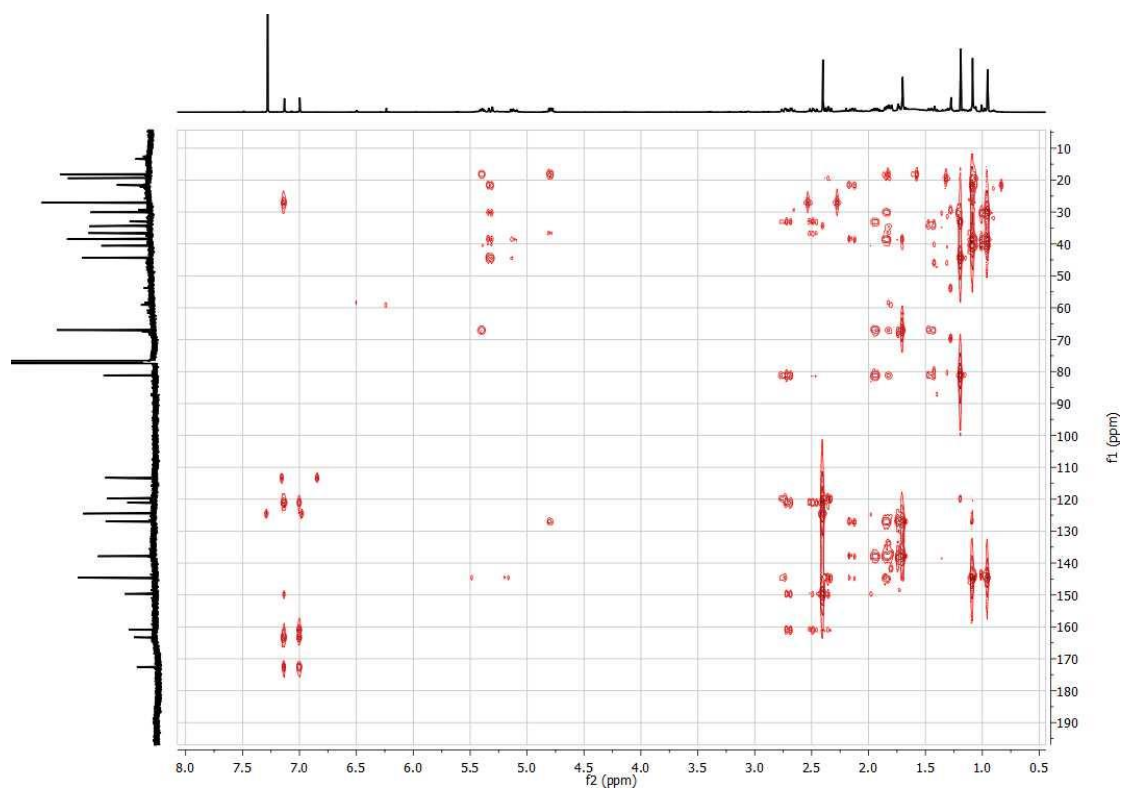
$^1\text{H}$ - $^1\text{H}$  COSY spectrum of neosetophomone B **61**.



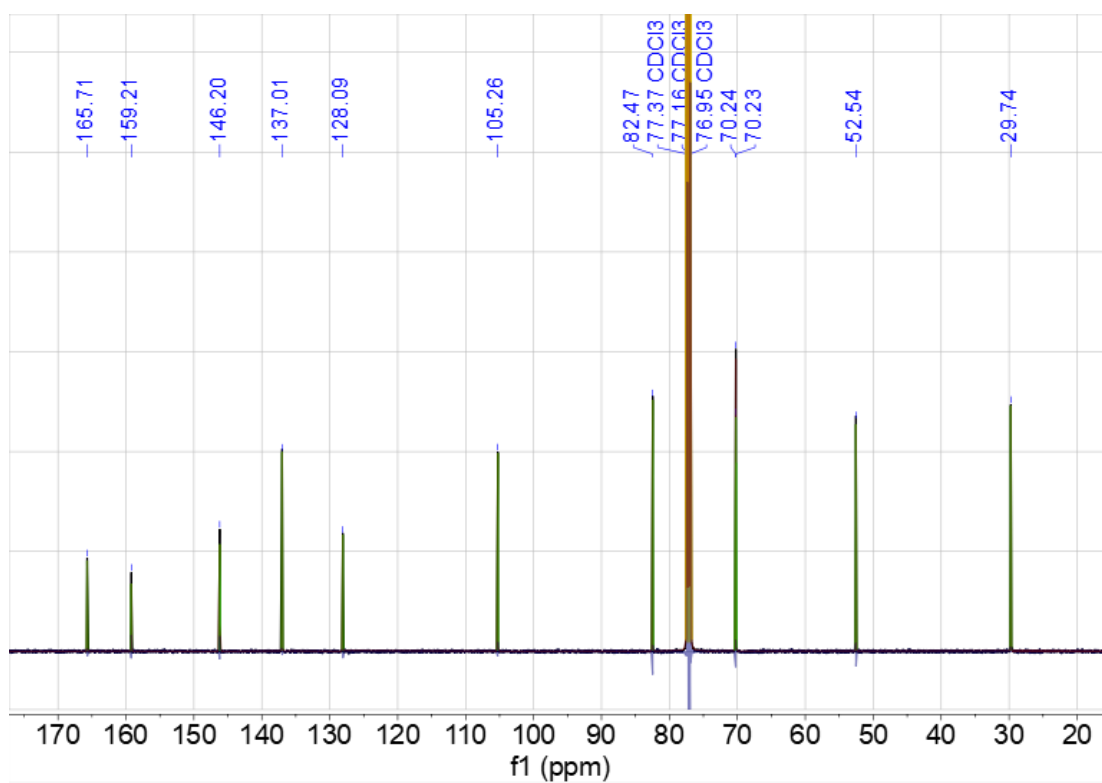
HSQC spectrum of neosetophomone B **61**.



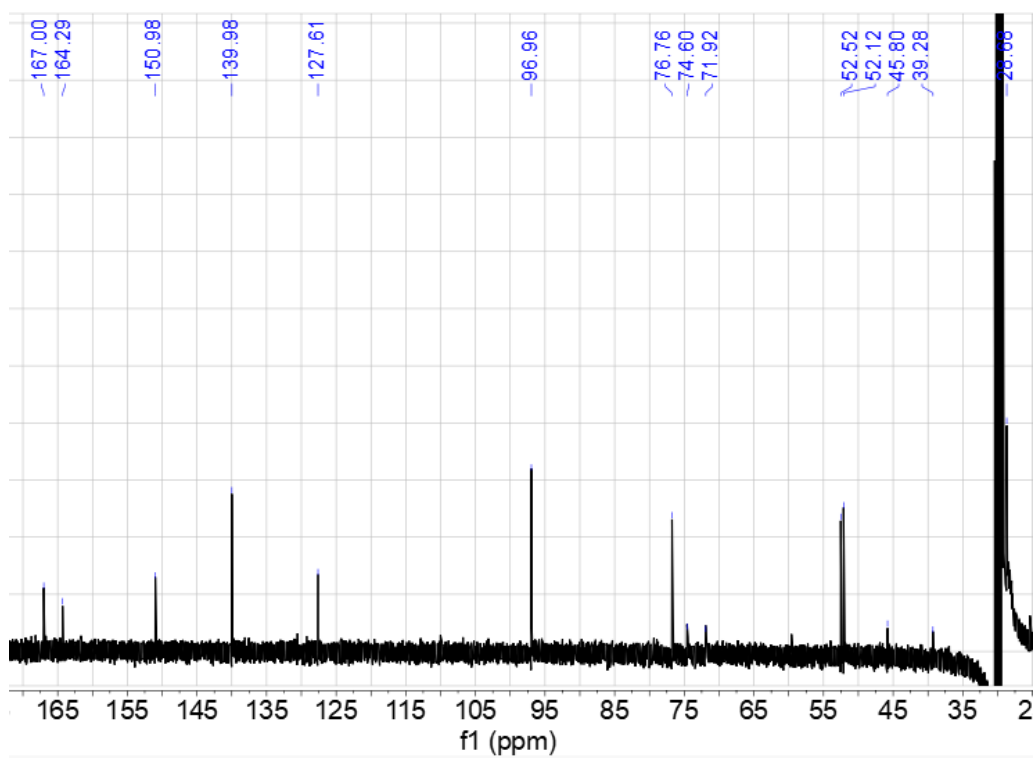
HMBC spectrum of neosetophomone B **61**.



$^{13}\text{C}$  NMR of scytolide **158** in  $\text{CDCl}_3$  (125 MHz) referenced to  $\text{CDCl}_3$ .

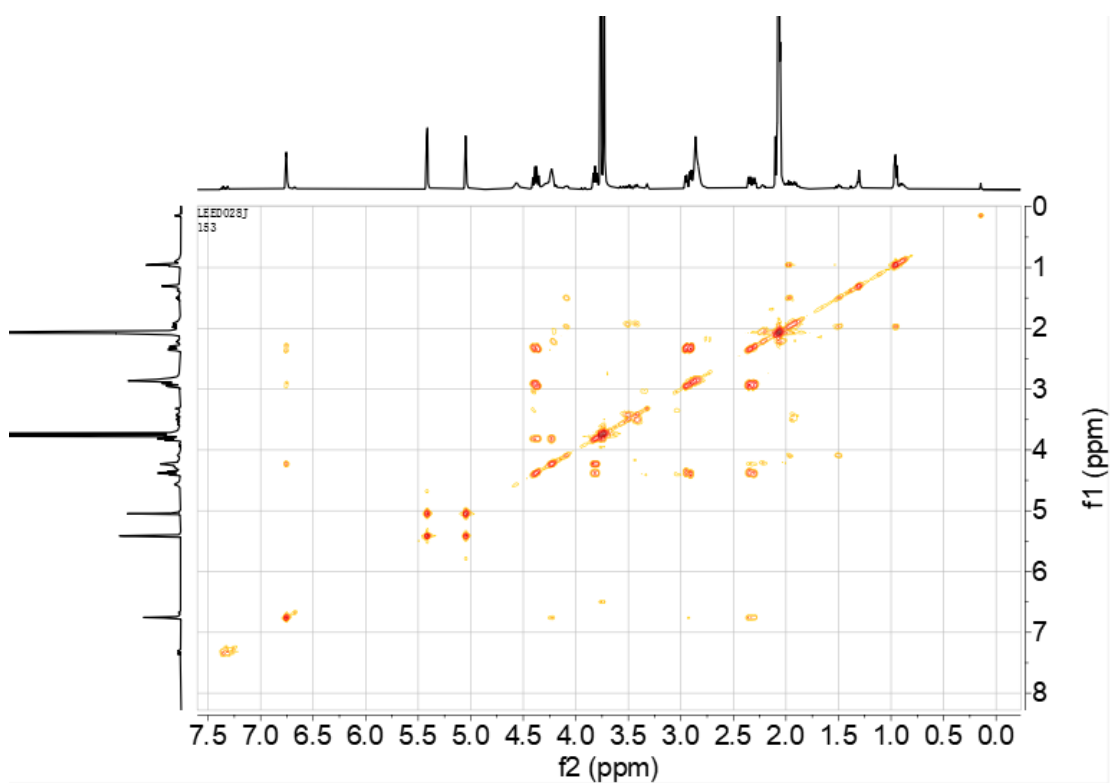


$^{13}\text{C}$  NMR of methylscytolide **198** in  $\text{CDCl}_3$  (125 MHz) referenced to  $\text{CDCl}_3$ .

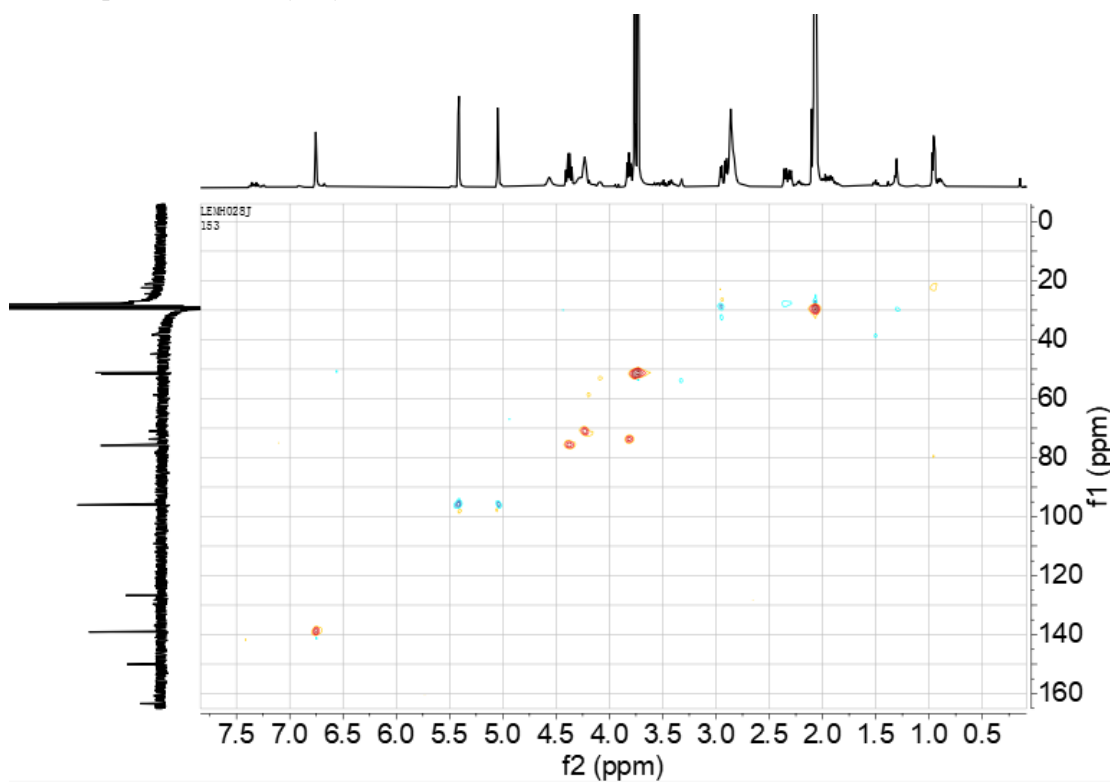




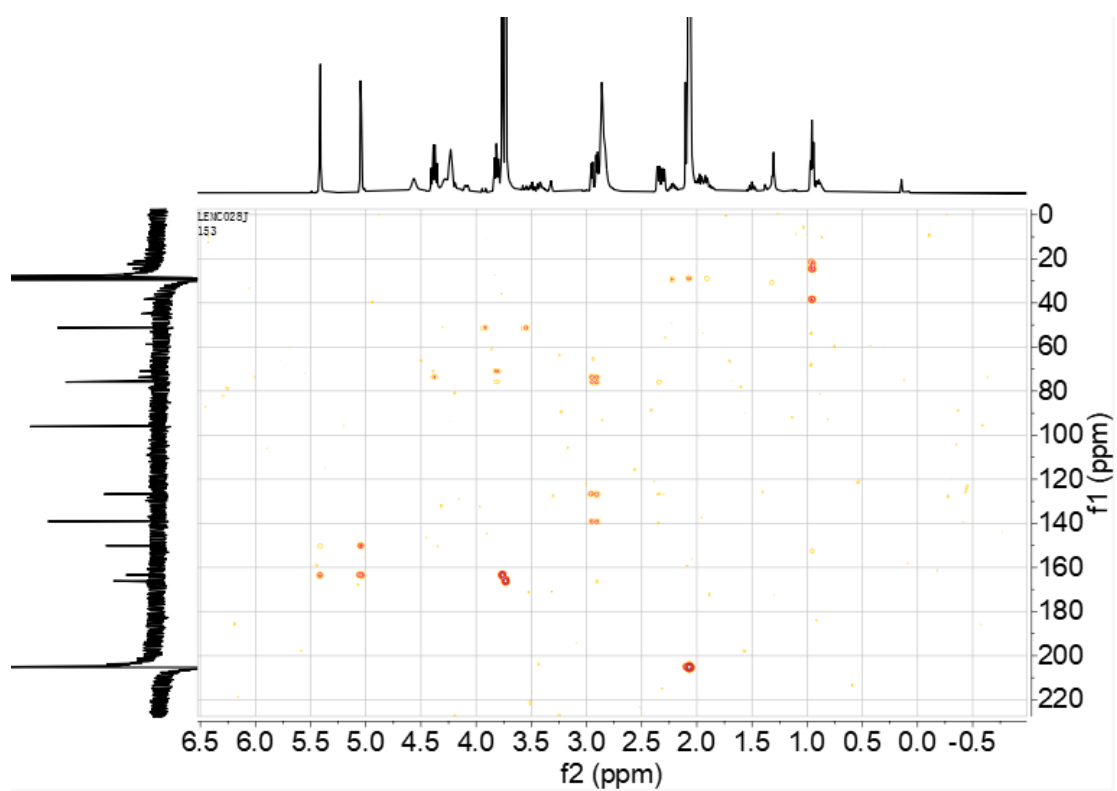
$^1\text{H}$ - $^1\text{H}$  COSY spectrum of methylscytolide **198**.



HSQC spectrum of methylscytolide **198**.



HMBC spectrum of methylscytolide **198**.



## Curriculum vitae

Oct 2017- Present	PhD student, <b>Organic Chemistry</b> , Gottfried Wilhelm Leibniz Universität Hannover, Hannover, Germany
Sep 2014- Jun 2017	Master of Science, <b>Pharmaceutical chemistry</b> , Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, China
Sep 2010- Jun 2014	Bachelor of Science, <b>Resources and development of traditional Chinese medicine</b> , China Pharmaceutical University, Nanjing, China

## List of Publications

- 1 C. Schotte, L. Li, D. Wibberg, J. Kalinowski and R. J. Cox, Synthetic Biology Driven Biosynthesis of Unnatural Tropolone Sesquiterpenoids, *Angew. Chemie - Int. Ed.*, 2020, **59**, 23870–23878.
- 2 X. R. Peng, L. Li, J. R. Dong, S. Y. Lu, J. Lu, X. N. Li, L. Zhou and M. H. Qiu, Lanostane-type triterpenoids from the fruiting bodies of *Ganoderma applanatum*, *Phytochemistry*, 2019, **157**, 103–110.
- 3 L. Li, X. R. Peng, J. R. Dong, S. Y. Lu, X. N. Li, L. Zhou and M. H. Qiu, Rearranged lanostane-type triterpenoids with anti-hepatic fibrosis activities from *Ganoderma applanatum*, *RSC Adv.*, 2018, **8**, 31287–31295.
- 4 L. Li, H. Li, X. R. Peng, B. Hou, M. Y. Yu, J. R. Dong, X. N. Li, L. Zhou, J. Yang and M. H. Qiu, (±)-Ganoapplanin, a Pair of Polycyclic Meroterpenoid Enantiomers from *Ganoderma applanatum*, *Org. Lett.*, 2016, **18**, 6078–6081.
- 5 X. Peng, L. Li, X. Wang, G. Zhu, Z. Li and M. Qiu, Antioxidant farnesylated hydroquinones from *Ganoderma capense*, *Fitoterapia*, 2016, **111**, 18–23.